# Modern techniques and applications for (imaginary-)time evolution of quantum spin systems

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

von
Maurits S. J. Tepaske
aus
Zaandam, Netherlands

Bonn, November 2023

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

|  |  |
|---|---|
| 1. Gutachter/Betreuer: | Prof. Dr. David J. Luitz |
| 2. Gutachterin: | Jun.-Prof. Dr. Lena Funcke |
| Tag der Promotion: | 27.06.2024 |
| Erscheinungsjahr: | 2024 |

# Acknowledgements

For my parents, for their unconditional love and support.

I would like to thank David Luitz for giving me the freedom to thrive and the guidance to achieve my goals, for granting me the oppertunity to develop as a researcher and as a person. I would like to thank Dominik Hahn for being a collaborator and a friend. I would like to thank the Max Planck Institute for the Physics of Complex Systems in Dresden for the wonderful time I spent there, and for the many clock cycles on its supercomputer. I would like to thank the University of Bonn for its hospitality.

# Contents

**1 Introduction** **1**
1.1 Genesis of quantum mechanics . . . 1
1.2 Classical spin systems . . . 1
1.3 Quantum spin systems . . . 3
1.4 Exponential complexity . . . 4
1.5 Imaginary-time evolution operator . . . 5
1.6 Local approximations . . . 6

**2 Quantum Monte Carlo** **7**
2.1 Worldline Quantum Monte Carlo . . . 7
2.2 Stochastic Series Expansion . . . 8
2.3 Topological phase transition in the XY model . . . 9
2.4 Commentary on Article 1 . . . 13
2.5 Commentary on Article 2 . . . 14

**3 Iterative algorithms** **16**
3.1 Approximate time evolution . . . 16
3.2 First-order decomposition . . . 18
3.3 Variational compiling . . . 18
3.4 Computational complexity . . . 20
3.5 Higher order circuits . . . 20
3.6 Symmetric circuits . . . 21
3.7 Designing efficient algorithms . . . 22
3.8 Improved first-order bound . . . 23
3.9 Recursive construction of higher-order circuits . . . 23
3.10 Computational complexity and no fast-forwarding . . . 24
3.11 Commentary on Article 3 . . . 25

**4 Tensor networks** **27**
4.1 Manifest locality . . . 27
4.2 Isometric tensor networks . . . 28
4.3 Circuit application . . . 31
4.4 States and operators of local Hamiltonians . . . 32
4.5 Efficient time evolution . . . 33
4.6 Higher-dimensional isometric tensor networks . . . 34

4.7 Algorithmic decoherence . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35
4.8 Long-range couplings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36
4.9 Non-isometric tensor networks . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36
4.10 Commentary on Article 4 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37

**5 Quantum computing** **39**
5.1 Early days of quantum computing . . . . . . . . . . . . . . . . . . . . . . . . . 39
5.2 Quantum circuits . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 40
5.3 Experimental implementation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 41
5.4 Efficient circuits . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42
5.5 Effects of decoherence . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43
5.6 Alternatives to product formulas . . . . . . . . . . . . . . . . . . . . . . . . . . 44
5.7 Randomization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44
5.8 Variational quantum compiling . . . . . . . . . . . . . . . . . . . . . . . . . . . 45
5.9 Mitigating errors . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46
5.10 Commentary on Article 5 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46

**Bibliography** **48**

**A Article 1** **69**

**B Article 2** **80**

**C Article 3** **89**

**D Article 4** **118**

**E Article 5** **131**

## CHAPTER 1

# Introduction

## 1.1 Genesis of quantum mechanics

If intuition for quantum mechanics would be contained a priori in the human understanding, the wave function and its implications would come naturally, without the need for analogy. But evolution has driven the human brain to develop a natural capacity for understanding causality as it appears in the classical world. As such, it is only natural that the effort of reason to grasp nature begins here, and works its way to modern quantum many-body physics through a sequence of discoveries powered by analogy.

Centuries of scattered research and speculation in mechanics and optics culminated into the classical mechanics and geometrical optics of Newton [1, 2] and the undulatory optics of Huygens [3], in the seventeenth century, providing a coherent and intuitive description of classical matter and waves. This was brought to even greater heights by Maupertuis, Euler, and Lagrange, among others, with a prime motivator being the potential unification of physics through a principle of least action [4]. Building on this, Hamilton unveiled a remarkable correspondence between mechanics and optics [5], with geometrical optics describing mechanical trajectories of light and undulatory optics offering a complementary wave perspective on mechanics.

This opto-mechanical correspondence motivated Schrödinger to develop his wave mechanics [6–11], under the impression that the discrepancies in classical microscopic theories, which were prevalent in the early days of quantum mechanics, could be explained as a regime where the mechanical trajectories become comparable in size to a wavelength that involves the Planck constant $h$.[1] Some months earlier, Heisenberg arrived at the same theory in a different formalism [13–15]. He already had the quantum many-body problem in mind, that he would introduce a couple years later [16].

## 1.2 Classical spin systems

Before quantum mechanics was formalized by Heisenberg and Schrödinger, some of its phenomena were already well known. An example is the peculiar behavior of magnetic moments in crystals [17]. Guided

---

[1] Essentially, Schrödinger introduces the wave function by interpreting Hamilton's principal function as the phase of a wave (constituted by trajectories that solve the dynamical equations). Then the Schrödinger equation follows from the Hamilton-Jacobi equation under the assumption that the Planck constant $h$ is the relevant energy×time scale. Many flavors of this were already present in de Broglie's work [12].

by emperical findings and their extrapolations, it seemed reasonable to assume that each molecule in a crystal effectively has a magnetic moment that points either up or down [18]. Out of this idea the Ising model was born [19], where binary-valued magnetic moments $Z_i$ (i.e. *spins*) are arranged on a lattice with $N$ sites, and nearest-neighbors (nn) are coupled classically according to the Hamiltonian

$$H = -\sum_{\langle ij \rangle} Z_i Z_j. \tag{1.1}$$

Here $\langle i, j \rangle$ contains all nn pairs.

The local coupling $-Z_i Z_j$ is chosen to promote ferromagnetism, which was experimentally observed by Weiss [17]. At finite temperature, the tendency to form a global ferromagnet competes with thermal fluctuations, leading to a phase transition in two dimensions (2D) between order and disorder [20]. It was the first many-body model for which large scale numerical simulations were extensively used, especially for 3D lattices [21, 22]. Although the model is classical and does not require any quantum mechanical calculations, it was targeted at quantum mechanical phenomena and became a blueprint for modern quantum many-body physics.

We can also consider a more general classical spin than $Z$, by defining it as a canonical coordinate in Hamilton's formulation of classical mechanics [23]. A classical 3D spin is denoted by $S^{x,y,z}$, consisting of three real numbers that are constrained such that $\sum_\alpha |S^\alpha|^2 = S_0$ is constant. We get the binary spin $Z$ by restricting $S$ to a single axis. The magnitude of the spin $S_0$ is largely irrelevant, since the spin variables are generally continuous functions.

We put a spin $S_j^\alpha$ at every site $j$ of a lattice, and for concreteness we couple neighboring spins ferromagnetically through their inner product, i.e. we consider the Heisenberg Hamiltonian

$$H = -\sum_{\langle i,j \rangle} S_i^x S_j^x + S_i^y S_j^y + S_i^z S_j^z. \tag{1.2}$$

In the Hamiltonian formulation of classical mechanics, we can get a corresponding equation for the time evolution of the spin variables, as long as we can endow the set of spins with a canonical structure. To this end we define a Poisson bracket for spin variables, i.e.

$$\{A, B\} = \sum_{j\alpha\beta\gamma} \epsilon_{\alpha\beta\gamma} \frac{\partial A}{\partial S_j^\alpha} \frac{\partial B}{\partial S_j^\beta} S_j^\gamma, \tag{1.3}$$

where $\epsilon_{\alpha\beta\gamma}$ is the Levi-Civita symbol. It satisfies the usual Poisson bracket algebra, and it contains $\{S_j^\alpha, S_k^\beta\} = \delta_{jk} \sum_\gamma \epsilon_{\alpha\beta\gamma} S_j^\gamma$, where $\delta_{jk}$ is the Kronecker symbol. This structure is familiar from the Pauli algebra that governs a quantum spin-1/2 particle.

With this we get the Hamilton evolution equation

$$\frac{\mathrm{d}S_j^\alpha}{\mathrm{d}t} = \{S_j^\alpha, H\}. \tag{1.4}$$

To solve it formally, we simply Taylor expand $S_j^\alpha$ around $t = 0$ and use (1.4) to reduce all derivates to nested Poisson brackets, i.e.

$$S_j^\alpha(t) = S_j^\alpha + t\{S_j^\alpha, H\} + \frac{t^2}{2!}\{\{S_j^\alpha, H\}, H\} + \dots. \tag{1.5}$$

By defining the linear operator $\{\cdot, H\}$, which maps a function $O$ to another function $\{O, H\}$, we can write the closed form

$$S_j^\alpha(t) = e^{t\{\cdot,H\}} S_j^\alpha. \tag{1.6}$$

Here the operator exponential is defined as

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}, \tag{1.7}$$

where $A^k$ signifies $k$ sequential applications of the linear operator $A$.

From (1.3) it is clear that that every order in (1.5) encodes increasingly long-ranged spin correlations. For example, with the Heisenberg Hamiltonian (1.2) we have only nn contributions at first order:

$$\{S_j^\alpha, H\} = \sum_{k\in\mathrm{nn}(j)} \sum_{\beta\gamma} \epsilon_{\alpha\beta\gamma} S_k^\beta S_j^\gamma. \tag{1.8}$$

Then at the next order also the nearest-neighbors of the $k$ sites contribute, and so on. Hence the exponential (1.6) boils down to a system of polynomials for the spin variables. If we split the evolution into $n$ segments, i.e.

$$e^{t\{\cdot,H\}} = \left(e^{\frac{t}{n}\{\cdot,H\}}\right)^n, \tag{1.9}$$

and choose $n$ such that $t/n << 1$, then we only need to consider a finite amount of terms in (1.5) to reach machine precision.

The evolution (1.6) leaves the spin magnitude $S_0$ invariant, as well as the fundamental Poisson bracket $\{S^\alpha, S^\beta\}$ [23]. Hence it is a canonical transformation, and the time evolution of a spin system can thus be considered as a sequence of canonical transformations of its phase space.

## 1.3 Quantum spin systems

Quantum spin was discovered around the same time as the Ising model and the formalization of quantum mechanics, in an effort to explain experimentally-observed quantum phenomena [24, 25]. Soon after, a theory of magnetism appeared [16]. It was found that the quantum version of the continuous spin-1/2 variable $S^\alpha$ is in fact more similar to the binary variable $Z$, being fundamentally discrete, but that quantum fluctuations couple different binary configurations and thereby create a *superposition*. The spin is now a state vector in a Hilbert space, which is spanned by $Z = \pm 1$. In contrast to the classical spin $S^\alpha$, the spin magnitude is intimately tied into the physics, with e.g. a spin-1 particle requiring a larger space with $Z = -1, 0, 1$. When the spin magnitude is measured along one of the space axes, and the experiment is repeated many times, it yields a probability distribution over the binary states, with measured spins pointing either up or down [26]. We call a quantum spin-1/2 particle a *qubit*, since it is the quantum version of a bit $Z$.

To quantize the classical time evolution of a many-qubit system (1.4) we replace the $3N$ real-valued binary variables $S_j^\alpha$ with a single $2^N$-dimensional complex-valued state $|\psi\rangle$, living in a Hilbert space spanned by all possible bit configurations $|Z_1, Z_2, ..., Z_N\rangle$. To determine the properties of a single qubit we generally need to consider all bit configurations of the entire system. Concretely, to extract local properties from the global state we assign a set of $2 \times 2$ Pauli operators $\sigma_j^{I,x,y,z}$ to every spin, forming a Pauli algebra (i.e. a two-dimensional fundamental representation of $SU(2)$) [25, 27, 28]. The inner product $m_j^{x,y,z} = \langle\psi|\,\sigma_j^{x,y,z}\,|\psi\rangle$ then gives the average $x, y, z$ spin components of the $j$th qubit, as it would be determined in an experiment. Because we are now dealing with a superposition instead of a single bit configuration, we have to do with probabilistic statements when concerned with measurable properties of the spin system. When the qubits are also coupled to their environment, which seems inevitable in practice, their quantum state is instead described by a density matrix, which encodes a probability distribution of states. For simplicity we consider isolated states $|\psi\rangle$ in most of this thesis.

Accordingly, instead of a classical Hamiltonian, which is a continuous function over the $3N$ spin variables, we get a quantum Hamiltonian that encodes the couplings between the $2^N$ bit configurations as a complex matrix. Consequently, while the classical Hamiltonian can be computed in $O(N)$ time and resources, for the quantum Hamiltonian we generally have $O(2^N)$. For example, we obtain the quantum spin-1/2 Heisenberg model by replacing the variables $S$ in (1.2) with Pauli operators $\sigma$. Schrödinger has argued that an isolated state evolves according to

$$i\hbar\frac{\partial\,|\psi\rangle}{\partial t} = H\,|\psi\rangle\,. \tag{1.10}$$

It is formally solved by

$$|\psi(t)\rangle = e^{-itH}\,|\psi(0)\rangle\,, \tag{1.11}$$

which boils down to a matrix-vector product of a unitary matrix with a normalized vector, which is therefore again normalized. The exponential operator in (1.11) is called the *time-evolution operator*, and it is a centerpiece of quantum dynamics since it can map any state to a state evolved according to $H$. It is the quantum analog of the classical time-evolution operator (1.6). Now the powers in the exponential (1.7) are calculated with matrix-matrix multiplication, or if we want to apply the exponential to a state we only need matrix-vector multiplication.

The time-evolution operator is diagonal when $H$ is diagonal, in which case it cannot generate a superposition, such that the time evolution is classical. An example is the Ising Hamiltonian (1.1) with the binary variables $Z$ upgraded to $\sigma^z$ operators. Classical spin systems such as (1.2) were only studied much later than their quantum counterparts [29–33]. This is likely because spin was conceived in the context of quantum mechanics, such that the quantum version of (1.2) was immediately considered. Notwithstanding, it has led to the discovery of many interesting phenomena, such as the topological Berezinskii-Kosterlitz-Thouless (BKT) phase transition that was unveiled in the classical two-dimensional XY model [34–37]. Afterwards, it was realized that this phase transition also occurs in quantum spin systems [38–41]. We will consider it in detail in Sec. 2.3. We have studied aspects of BKT physics in great detail in the first two Articles that feature in this thesis [42, 43], which we will describe at the end of Ch. 2.

## 1.4 Exponential complexity

By designing algorithms that construct or apply the time-evolution operator (1.11) we can in principle simulate any quantum many-body system. The difference in computational complexity between the evolution of quantum and classical many-body systems is clear: We now generally need to do linear algebra in a $2^N$-dimensional Hilbert space, spanned by the bit configurations, instead of solving a system of $N$ equations in a $N$-dimension phase space, spanned by the spin variables. Therefore, time evolution of quantum many-body systems generally requires $O(2^N)$ memory and floating point operations, whereas classical systems only require $O(N)$.

Consequently, on a classical computer we can simulate quantum time evolution only for small systems, highlighting a major bottleneck of classical simulation methods. Notwithstanding, we will see in Ch. 2 of this thesis, on Quantum Monte Carlo, and in Ch. 4 on tensor networks, that for specific applications there are classical algorithms with only polynomial scaling. To achieve this in general requires a more drastic approach, for which we have to replace the classical computer with a quantum computer. We will discuss this in Ch. 5. Except for modern Quantum Monte Carlo algorithms, these methods rely heavily on the circuit decompositions that we will discuss in great detail in Ch. 3

## 1.5 Imaginary-time evolution operator

Besides dynamics, the Schrödinger equation can also be used to directly determine steady states at zero temperature. These are the eigenstates $|n\rangle$ of the Hamiltonian, which satisfy $H|n\rangle = E_n|n\rangle$ with energies $E_0 < E_1 < \ldots$, and upon applying the real-time evolution operator they only gain a phase. Consequently, their quantum fluctuations average out and hence their expectation values are time independent. In other words, they describe a system in equilibrium.

We transform the Schrödinger equation (1.10) to imaginary time, i.e. we subtitute $t \longrightarrow -i\tau$, yielding the imaginary-time evolution operator

$$P = e^{-\tau H}. \tag{1.12}$$

It is no longer unitary, and instead it now serves as a projector. Specifically, if we have an arbitrary wavefunction $|\psi_i\rangle$, we first expand it into the eigenbasis $|n\rangle$ and then apply $P$, yielding

$$P|\psi_i\rangle = \sum_n \langle n|\psi_i\rangle\, e^{-\tau E_n} |n\rangle\,. \tag{1.13}$$

We now see that the excited states decay quicker with $\tau$ than the ground state. If we shift the energy such that $E_0 = 0$, we are left with only the ground state $|0\rangle$ as $\tau \longrightarrow \infty$, provided that $\langle 0|\psi_i\rangle \neq 0$. Having determined $|0\rangle$, we can now determine $|1\rangle$ by performing imaginary-time evolution on $|\psi_i'\rangle \propto (1 - |0\rangle\langle 0|)|\psi_i\rangle$. This enforces $\langle 0|\psi_i'\rangle = 0$, such that the slowest-decaying component in (1.13) is now the first excited eigenstate. In principle, we can obtain the entire spectrum of $H$ in this manner.

The imaginary-time evolution operator (1.12) also encodes the thermal equilibrium of $H$. Whereas quantum fluctuations can be confined to an eigenstate, thermal fluctuations couple different eigenstates. Hence we need to use density matrices, which describe a probability distribution over the eigenstates. Concretely, we project the maximally mixed state to find

$$\rho_s = P \sum_n |n\rangle\langle n| = \sum_n e^{-\tau E_n} |n\rangle\langle n|\,. \tag{1.14}$$

Here we neglected normalization, so we must divide by the partition function $Z \equiv \langle \rho_s \rangle$ to obtain a physical state. We immediately recognize this as the canonical ensemble at temperature $k_B T = 1/\tau$, with $k_B$ the Boltzmann constant, for which the eigenstates are distributed according to the Boltzmann distribution

$$p_n = \frac{e^{-\tau E_n}}{\sum_n e^{-\tau E_n}}. \tag{1.15}$$

To determine the expectation value of an arbitrary operator $O$ we evaluate

$$\langle O \rangle = \mathrm{Tr}\left[O\rho_s\right]/Z = \mathrm{Tr}\left[OP\right]/Z. \tag{1.16}$$

This corresponds to the average value of $O$ when the system is in thermal equilibrium at temperature $1/\tau$, where both the quantum and thermal fluctuations average out, such that $\langle O \rangle$ is again time independent.

## 1.6 Local approximations

To design an algorithm that is more efficient than the explicit construction of the entire time-evolution operator, we need to utilize the specific structure of the system that we want to simulate. In our case, we can use the properties of the spin Hamiltonian, e.g. the Heisenberg model (1.2). This model is local, meaning that spins only interact when they are physically close, thereby encoding the mechanism for generating long-ranged correlations that exists in nature and condensed matter system in particular. As such, it is a powerful tool, both for simplification of the model and of the simulation method. For instance, in Article 1 of this thesis [42], which we will discuss in Sec. 2.4, we are able to qualitatively capture the phenomenology of an experimental 3D quantum magnet with a simple anisotropic Heisenberg model.

A natural approach to simulating the global time-evolution of a local Hamiltonian is to approximate it as a sequence of local evolutions, e.g.

$$e^{-it\sum_\alpha H_\alpha} \approx \prod_\alpha e^{-itH_\alpha}. \tag{1.17}$$

This does not hold exactly because the Hamiltonian terms generally do not all commute, and hence neither do their exponentials, but because $H$ is local we at least know that each exponential commutes with many others. For simplicity, let us consider the simple case of the Heisenberg model (1.2) on a chain, in which case $H$ can be subdivided into two disjoint sets $H = H_a + H_b$, with each element commuting only with other elements in its set and non-commuting with the rest. Then we can implement the global time evolution with a non-commuting sequence of local evolutions $e^{-itH_a}$ and $e^{-itH_b}$.

To this end we first factor the time-evolution into $n$ segments, i.e. $e^{-itH} = (e^{-itH/n})^n$, and then take the limit to an infinite amount of segments. Then we can invoke the Lie product formula [44]

$$e^{-it(H_a+H_b)} = \lim_{n\to\infty}\left(e^{-itH_a/n}e^{-itH_b/n}\right)^n, \tag{1.18}$$

indicating that (1.17) holds exactly in the limit $t \to 0$. Since $H_a$ and $H_b$ are each a sum of commuting operators, their exponentials can be decomposed exactly into a product of two-qubit evolutions, e.g.

$$e^{-itH_a} = \prod_{j=1}^{N/2} e^{-itH_{2j,2j+1}}, \tag{1.19}$$

with $H_{j,j+1}$ the Heisenberg coupling between sites $j$ and $j+1$. The same decompositions hold for the imaginary time-evolution operator, and for complex $t$ in general.

In the following chapter we will use (1.18) to construct a powerful algorithm for the simulation of equilibrium physics. Afterwards, in Ch. 3 we will extensively discuss generalizations of (1.18).

CHAPTER 2

# Quantum Monte Carlo

## 2.1 Worldline Quantum Monte Carlo

With the decomposition (1.18) we can map the time-evolution operator of a $d$-dimensional quantum spin system onto that of a $(d+1)$-dimensional classical Ising model [45, 46]. Concretely, we split the time-evolution operator into $n$ segments according to (1.18), and insert a resolution of identity $\sum_{\vec{\sigma}} |\vec{\sigma}\rangle \langle\vec{\sigma}| = I$ between each pair of exponentials, where $\vec{\sigma} = (\sigma_1, \sigma_2, \ldots)$ is the computational basis. For the Heisenberg model this yields

$$e^{-itH} = \lim_{n\to\infty} \sum_{\vec{\sigma}_1,\vec{\sigma}_2,\ldots} e^{-itH_A/n} |\vec{\sigma}_1\rangle \langle\vec{\sigma}_1| e^{-itH_B/n} |\vec{\sigma}_2\rangle \langle\vec{\sigma}_2| e^{-itH_A/n} \ldots . \tag{2.1}$$

The segmentation induces an extra time dimension, which encodes the quantum nature of the Hamiltonian, appearing only when the evolution operators of neighboring segments non-commute. The time-evolution operator is now represented as a sum over paths through this time. If we apply it to a state, we can propagate it along the different paths, and by summing the results we get the time-evolved state. If we can determine bases $\vec{\sigma}_j$ such that all the amplitudes in (2.1) are simultaneously real and non-negative, then we can sample them from a $(d+1)$-dimensional Ising model, for which we can use traditional Monte Carlo techniques such as the Metropolis algorithm [47].

For real-time evolution this program is spoiled by the factor $i$ in each exponential. On account of this, we cannot determine the sought-after bases, such that we are forced to deal with complex amplitudes. Consequently, the accurate estimation of (2.1) will rely on the cancellation of many complex numbers, such that the number of samples required to reach a fixed precision grows exponentially with the system size. This is called the *sign problem*, and it brings us back to the exponential asymptotic cost of constructing the time-evolution operator in its entirety [48]. In contrast, if all amplitudes would be simultaneously real and non-negative, the cost would be only linear in system size.

Now we consider the imaginary-time evolution operator, for which (2.1) holds with $t \to -i\tau$. If we want to calculate the expectation value of an observable $O$ with (1.16), we have to apply $O$ to (2.1) and take the trace. By representing the trace as a sum over computational basis states, it is immediately clear that $\mathrm{Tr}e^{-\tau H}$ is now a sum over periodic paths. Consequently, $\langle O\rangle$ is simply an average over these paths. Because we no longer have the factor $i$, there is now a possibility that the sampling scheme can be carried out efficiently, depending on the particular model at hand. For instance, the Heisenberg model (1.2) on a bipartite lattice enjoys this property, and the resulting linear complexity allows for the large-scale simulation of its equilibrium properties [49]. This method is called *worldline quantum Monte Carlo* (QMC) formulated in discrete time [50–54]. See [55] for an extensive introduction to this method and a summary of early works that use it. At zero temperature we can perform the equilibrium simulations only indirectly, by performing simulations at various finite temperatures and extrapolating its expectation values to zero.

In practice, we have to use a finite segmentation $n$ in (2.1), which introduces a discretization error that has to be extrapolated in $n$ to obtain statistically exact results. This could be computationally costly, and it is therefore interesting to consider the particular error structure of a truncated version of (1.18). Perhaps we can improve it by modifying the evolution operator of each segment. We will discuss this in great detail in Ch. 3. Nowadays, this method has been formulated directly in the continuous-time limit $n \longrightarrow \infty$, eliminating the discretization error altogether. It is called continuous-time worldine QMC [56–59] and has proved itself in many large-scale numerical studies [60–66]. As mentioned above, when there is no sign problem, the cost is linear in system size, for systems of any dimensionality. In that case it is arguably the most powerful method for equilibrium simulation known today.

## 2.2 Stochastic Series Expansion

When time-reversal symmetry is broken, e.g. by a magnetic background field, we can construct a QMC algorithm that is even more powerful than worldline QMC. We will use this for the equilibrium simulation of experimental 3D quantum magnets in Article 1 [42], which we will discuss in Sec. 2.4.

To uncover it we have to take a different approach to the imaginary time-evolution operator. We immediately restrict ourselves to equilibrium averages and consider the canonical partition function $Z$, which we expand as a Taylor series

$$Z = \sum_{\sigma} \sum_{n=0}^{\infty} \frac{(-\tau)^n}{n!} \langle\sigma| H^n |\sigma\rangle \,. \tag{2.2}$$

We utilize the locality of $H = \sum_b H_b$ to write $H^n$ as a sum of products, i.e.

$$H^n = \sum_{S_n \in C_n} \prod_{b \in S_n} H_b, \tag{2.3}$$

where $C_n$ contains all operator products $S_n$ with $n$ bond operators $H_b$, e.g. $S_4 = H_6 H_3 H_2 H_1$.

Inspired by the success of Monte Carlo in simulating the Ising model, Handscomb formulated the initial proposal for QMC based on this expansion, where a Markov Chain over $C_n$ is used to sample operator strings with classical Monte Carlo provided there is no sign problem [67]. For the ferromagnetic Heisenberg model, where all terms in (2.2) become positive, he derived analytical expressions for the expectation value of an arbitrary operator string, and with this he introduced the first QMC algorithm for a quantum many-body system [68]. It gave access to previously intractable system sizes, for special Hamiltonians [69–71].

This was subsequently generalized to arbitrary sign-problem-free Hamiltonians, where the operator string expectation values are also sampled [72–74]. To see how this can be done, we first note that each product in (2.3) propagates a state $|\sigma\rangle$ in (2.2), i.e.

$$|\sigma(k)\rangle = \prod_{b \in S_n(k)} H_b \, |\sigma\rangle \,, \tag{2.4}$$

where $S_n(k)$ contains the first $k$ operators in $S_n$, e.g. for the same example as before we get $S_4(2) = H_2 H_1$. Importantly, the initial state $|\sigma\rangle$ is always a classical state, and propagating it with bond Hamiltonians at most flips spins, never creating a superposition. Hence we have again induced a classical representation of the quantum partition function by creating an extra time dimension, along which the states are propagated by the $H_b$. This can be made explicit by substituting (2.4) and (2.3) into (2.2), yielding

$$Z = \sum_{\sigma} \sum_{n=0}^{\infty} \sum_{S_n \in C_n} \frac{(-\tau)^n}{n!} \prod_{k=1}^{n} \langle \sigma(k)_{b_k} | \, H_{b_k} \, | \sigma(k-1)_{b_k} \rangle \,. \tag{2.5}$$

Here $b_k$ is the bond of the $k$th operator in $S_n$ and $|\sigma_b(k)\rangle$ represents the $z$-spins of the bond $b$, e.g. $|\sigma_0(3)\rangle = |\uparrow\downarrow\rangle$.

The distribution of terms in the Taylor expansion (2.2) is concentrated in a small window of $n$. In particular, we have $\langle n \rangle = \beta N_b |E_b|$ with width $\sqrt{\langle n \rangle}$, where $N_b$ is the amount of terms in the Hamiltonian and $E_b = \langle H_b \rangle$ [74]. So we only need to consider a small amount of terms in order to get a manifestly unbiased simulation algorithm with computational complexity $O(N)$. Many observables $\langle O \rangle$ can be calculated directly from the information contained in $Z$, e.g. by averaging over the paths (2.4) [73].

This algorithm is called *stochastic series expansion* (SSE) QMC, and coupled with cluster sampling algorithms [56, 57, 74, 75] it became a powerful method for simulating the equilibrium properties of sign-problem-free systems [76–81]. SSE with directed loops [74] outperforms continuous worldline QMC with worms [82] for anisotropic Heisenberg models in a background magnetic field. This is rooted in their inherently different ways of sampling loops [74].

## 2.3 Topological phase transition in the XY model

As a testament to the power of QMC for sign-problem-free Hamiltonians, we were able to simulate a 3D anisotropic spin-1/2 Heisenberg model in a magnetic field at low temperature for up to a million spins. This was done in the context of Article 1 of this thesis, where an experimental material is studied that realizes this Heisenberg model. We will discuss this in detail in the commentary in Sec. 2.4. The material is of interest because it realizes a weakly-interacting stack of strongly-interacting square lattice magnets, providing some level of experimental access to exotic phenomena such as topological phase transitions. Specifically, in a small temperature range the material behaves as the antiferromagnetic XY model on a square lattice

$$H = \sum_{\langle i,j \rangle} S_i^x S_j^x + S_i^y S_j^y \,. \tag{2.6}$$

This model exhibits a phase transition that is characterized by a proliferation of topological defects and hence global instead of local ordering. In fact, thermal fluctuations prevent any local ordering at finite temperature in local two-dimensional systems that have a continuous symmetry [83].

Topological defects already occur in the ferromagnetic classical XY model on a square lattice, which we write as

$$H = -\sum_{\langle i,j} \cos(\phi_i - \phi_j). \tag{2.7}$$

Here each spin is represented as a normalized vector with angle $\phi$ relative to a fixed axis, i.e. $S_i^\alpha = (\cos\phi_i, \sin\phi_i)$. For illustration purposes we have chosen a ferromagnetic coupling, which is classically equivalent to the antiferromagnetic case if we neglect boundary conditions. With quantum mechanics this is no longer true in general. While the ferromagnet still has a classical ground state $|0\rangle = |\uparrow\uparrow\uparrow \ldots\rangle$, the antiferromagnet no longer has $|\mathbb{Z}_2\rangle = |\uparrow\downarrow\uparrow \ldots\rangle$ as its ground state. Instead, the ground state is now a superposition called the Néel state, which is antiferromagnetically ordered in the sense that its local magnetization $\langle\sigma_i^z\rangle$ is still equal to that of the classical state $|\mathbb{Z}_2\rangle$. At finite temperature the order is destroyed for both the ferromagnetic and antiferromagnetic XY model, due to the gapless spin-waves that accompany the $U(1)$ symmetry of (2.6) [83]. In 1D and 2D these proliferate because the entropy associated to the spin waves already outweighs their energy cost at any finite temperature, to the extent of preventing any local ordering. For 3D systems this is no longer the case, such that magnetic ordering can occur at finite temperature. This will become relevant in the experimental setting of Article 1.[1]

Because we are concerned with the low-energy properties of (2.7), we can expand the angles around the uniform ground state that is shown in Fig. 2.1a and keep only the quadratic term, i.e.

$$H - E_0 \approx \frac{1}{2} \sum_{\langle i,j\rangle} (\phi_i - \phi_j)^2. \tag{2.8}$$

This term clearly favors smooth configurations, so a natural first guess would be to consider only spin waves. In that case we have for separated spins [35]

$$\langle S_i^x S_j^x + S_i^y S_j^y \rangle \propto |r_i - r_j|^{-\frac{k_B T}{4\pi}}, \tag{2.9}$$

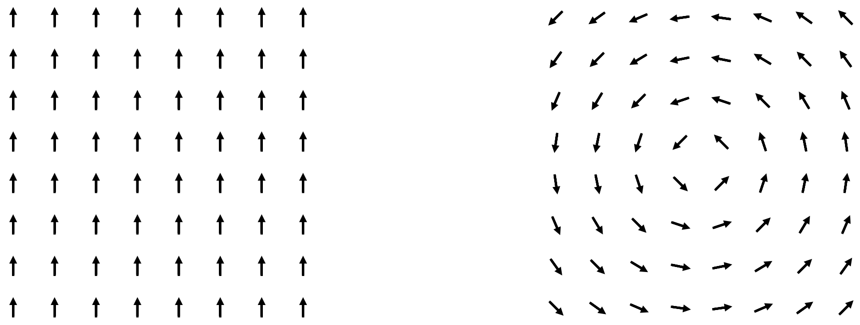
which vanishes for $|r_i - r_j| \to \infty$ such that thermally-excited spin waves indeed prevent long-range magnetic ordering at any non-zero temperature [34, 83]. Nonetheless, the correlations decay as a power-law, implying an infinite correlation length.

There also exist smooth *vortex* configurations, such as the one shown in Fig. 2.1b. When moving in a counter-clockwise closed path around the center of the vortex, such that we end up at the same spin, we know that the $\phi_i$ along the path sum to $2\pi$. In contrast, for the ground state from Fig. 2.1a they sum to zero along any closed path. If the vortex was reversed, they would instead sum to $-2\pi$. Clearly this is independent of the specific path that we choose, as long as it encloses the center and winds around it once. This indicates that the vortex is a topological defect in the ground state, where the direction of the vortex poses as a topological charge, i.e. we can assign $q = +1$ to the vortex and $q = -1$ to the *antivortex*. When a path encloses $N_v$ vortices and $N_{av}$ antivortices its angles generally sum to $2\pi q = N_v + N_{av}$ where $q \in \mathbb{Z}$.

Within the low-energy theory (2.8) the vortex is gapped from the ground state by $E_v - E_0 \approx 2\pi \ln(R)$, where $R$ is the radius of the vortex. If we consider large systems, then it is immediately clear that we can treat separately the vortex excitations with gaps $\Delta E \sim O(\ln(R))$ from the spin wave excitations with

---

[1] It should be noted that on finite lattices there is no symmetry breaking, and as such the statistically-exact QMC algorithms inherit the symmetries of the Hamiltonian. Therefore they cannot directly access the order parameter that is associated to the thermodynamic symmetry breaking, and instead we have to resort to more intricate probes such as the decay of spin correlations.



(a)

(b)

Figure 2.1: In panel (a) we show the uniform configuration that is the ground state of the 2D XY Hamiltonian. In panel (b) we show an isolated topological vortex, which is a smooth configuration of spins that has excitation energy

$\Delta E \sim O(1/R^2)$. Essentially, we can consider the vortices as equilibrium configurations and define the spin waves with respect to each. This makes the topological defects *metastable*, since the spin waves and vortices are effectively decoupled due to their different energy scales. However, at finite temperature $T$ we also need to take into account the configurational entropy of a vortex $S_v$, since the equilibrium configuration then coincides with the minimum of the free energy $F = E_v - TS_v$. Taking into account that the vortex center can roughly be placed on $R^2$ different sites, we estimate $S_v \approx 2k_B \ln(R)$. Hence the free energy is $F = 2(\pi - k_B T)\ln(R)$, predicting that above $k_B T_{\text{BKT}} = \pi$ the increase in energy is outweighted by the increase in entropy, seemingly enabling the proliferation of vortices. This suggests that (2.9) does not hold above $T_{\text{BKT}}$ and that we instead have to take into account the vortices. To uncover this in more detail we need to considering interacting vortices [35].

Far below $T_{\text{BKT}}$ there is a small chance that thermal fluctuations excite a vortex, because this takes $O(\ln(R))$ energy, but a bound vortex-antivortex pair only requires $O(1)$. If these are created already at $T < T_{\text{BKT}}$, then by increasing $T$ we can imagine that increasingly strong thermal fluctuations cause a vortex and its antivortex to drift apart, until they become unbound and a vortex can in principle exist on its own. By decomposing the $\phi_i$ into a spin-wave part that does not wind and a vortex contribution that encodes the winding, the low-energy theory (2.8) naturally decouples the spin-waves and vortices [35, 36]. The vortex system then becomes a neutral gas of vortices and anti-vortices, with equally-charged defects repelling eachother and oppositely-charged defects attracting eachother. The coupling decays logarithmically with separation. This effective theory can be used to show that there is indeed a phase transition associated to the unbinding of vortex-antivortex pairs.

Specifically, at $T \leq T_{\text{BKT}}$ the average separation between each vortex and its antivortex is finite (in an infinite system) [35]. In that case the vortex component of $\phi$ is long-range ordered, but the spin-wave component has the correlations (2.9) as usual and therefore prevents any long-range ordering of $\phi$ as a whole. At $T > T_{\text{BKT}}$ the average separation becomes infinite, at which point the vortex component of $\phi$ is no longer long-range ordered. As a result, the spin correlations now decay exponentially, such that the correlation length is finite. Accordingly, at $T_{\text{BKT}}$ the magnetic susceptibility jumps down from infinity to a finite value, indicating a sudden change in the response to a magnetic field. This can be understood in light of the vortex-antivortex unbinding, since at $T > T_{\mathrm{BKT}}$ even a tiny in-plane magnetic field can separate vortices and antivortices.

The vortex gas theory can be used to show that the spin-spin correlation length $\xi$ satisfies [35]

$$\xi \sim e^{1.5\sqrt{\frac{T_{\mathrm{BKT}}}{T-T_{\mathrm{BKT}}}}}, \tag{2.10}$$

in the limit $T \longrightarrow T_{\mathrm{BKT}}$ from above. This is qualitatively different than the critical scaling of a symmetry-breaking transition that is associated to long-range order, where $\xi$ instead diverges as a power of $T_{\mathrm{BKT}}/(T - T_{\mathrm{BKT}})$. The divergence in (2.10) is more severe than any such power. It can be shown that at $T = T_{\mathrm{BKT}}$ the spin-spin correlation function decays as $|r_i - r_j|^{-1/4}$ and that consequently the in-plane magnetic susceptibility $\chi$ diverges as $\chi \sim \xi^{2-\frac{1}{4}}$ [36]. This is in stark contrast with any known local symmetry-breaking transition.

So in the seemingly simple classical system (2.6) we have a finite-temperature phase transition between a topologically trivial and a topologically non-trivial phase, marked by the presence or absence of free topological charges. We have seen that it is accompanied by a qualitative change in spin correlations, but without local ordering or symmetry breaking in general. It is called the BKT transition, and it is a type of *topological phase transition*. Such transitions are not captured by the perturbative framework of Landau and Ginzburg, and are unveiled only when considering non-perturbative effects. The BKT transition also occurs in superfluids and the melting theory of solids [35], with the main requirement being the existence of metastable configurations that are gapped with $\Delta E \sim O(\ln R)$. For example, the 2D Heisenberg model (1.2) does not have metastable vortices. This is because there are now three equivalent spin directions in the low-energy theory, such that there would appear two angles in (2.8). These need to be summed simultaneously in order to define topologically charged vortices, but it turns out that the increased dimensionality reduces the gapping to $E_v - E_0 \sim O(1)$ [35].

Knowing that the main difference between the two phases is the behavior of the spin-spin correlations, which decay as a power law at $T \leq T_{\mathrm{BKT}}$ and as an exponential at $T > T_{\mathrm{BKT}}$, we can construct a global order parameter that captures this. As it turns out, in the absence of long-range order but in the presence of power-law correlations, applying a uniformly-increasing twist $\delta\phi_j = j\theta/L$ to the spins $\phi_i$ along one of the two spatial directions increases the free energy significantly, in the sense that the *spin stiffness* [84, 85]

$$\rho_s = \frac{1}{L^2}\frac{\partial^2 F}{\partial\theta^2} \tag{2.11}$$

is nonzero. As the name indicates, it provides a notion of rigidity of the spin system. Intuitively, the stronger a spin is correlated to the others, the more energy it will cost to rotate it on its own. When the spin correlations decay exponentially, we instead have $\rho_s = 0$, such that $\rho_s$ is the global parameter that we are looking for. It can be shown that at the critical point $\rho_s = 2T_{\mathrm{BKT}}/\pi$ [85]. In the limit of infinite system size $\rho_s$ steps discontinuously from this value to zero, but the behavior of $\rho_s$ as a function of both $T$ and $L$ can be used to extract $T_{\mathrm{BKT}}$ and various critical exponents. We utilize such a procedure in Article 1. Crucially, the spin stiffness is readily accesible with the QMC algorithms from this chapter. See Sec. 3.5 of Ref. [86] for an extensive introduction to the spin stiffness and the BKT transition in the context of QMC.

The transition is also accompanied by a maximum in the specific heat, which occurs when vortex-antivortex unbinding is happening at the maximal rate. Because the average vortex-antivortex separation is still finite at $T_{\mathrm{BKT}}$, as this temperature marks the point after which the first pairs can be imagined to start separating, it is only natural that the peak in the specific heat occurs slightly above $T_{\mathrm{BKT}}$, when the separation rate peaks [87].

The BKT mechanism for topological phase transitions carries over into the quantum realm, and the alignment induced by the local coupling appears to be irrelevant. For example, the antiferromagnetic quantum 2D XY model with spin-$1/2$ constituents is estimated to have a BKT transition with the same critical scaling and exponents as the ferromagnetic classical XY model that we considered above [39]. But now an analytical treatment is difficult compared to the classical scenario. Luckily, the XY model on a bipartite lattice is sign-problem-free, such that the QMC algorithms can be employed to perform efficient simulations. This allows for highly-accurate estimates of the critical scalings. The vortices generalize in a straightforward manner by encoding them into plaquette operators [88]. These have been used in QMC to illustrate that the specific heat peak indeed appears at the maximum rate of vortex-antivortex unbinding [89].

## 2.4 Commentary on Article 1

**Field-Tunable Berezinskii-Kosterlitz-Thouless Correlations in a Heisenberg Magnet**

Phys. Rev. Lett. 130, 086704 (2023)

*D. Opherden, M. S. J. Tepaske, F. Bärtl, M. Weber, M. M. Turnbull, T. Lancaster, S. J. Blundell, M. Baenitz, J. Wosnitza, C. P. Landee, R. Moessner, D. J. Luitz, H. Kühne*

Experimentally, a square lattice of spins has only been realized as an embedding in a 3D material [90–93]. In particular, there are various materials that realize a stack of weakly-coupled anisotropic Heisenberg lattices, each described internally by the 2D XXZ Hamiltonian

$$H = J \sum_{\langle i,j \rangle} S_i^x S_j^x + S_i^y S_j^y + (1-\Delta) S_i^z S_j^z . \tag{2.12}$$

Here $J > 0$ ranges wildly, from a few Kelvin to a few thousand. The exchange anisotropy $\Delta > 0$ is often on the order of $D \sim O(10^{-2})$ and is intrinsic to the material. Because the low-energy theory of the XXZ model (2.12) is that of the XY model (2.6), naively one would expect a BKT transition [40, 41]. However, in reality the interlayer antiferromagnetic coupling $J' \ll J$ spoils this. A review of candidate materials is given in Article 2 [43] that we will discuss below in Sec. 2.5.

Because the XY model is predicted to have an exponentially divergent correlation length as $T \to T_{\mathrm{BKT}}$, there appear large correlated regions of spins close to $T_{\mathrm{BKT}}$. When these become sufficiently large, the tiny $J'$ will be sufficient to trigger 3D antiferromagnetic ordering [90]. Since this occurs before the correlation length gets to diverge, the BKT transition is in reality foregone by a Néel transition at $T_{\mathrm{LRO}} > T_{\mathrm{BKT}}$. Nonetheless, if $J'/J$ is relatively small then the correlated XY regions are relatively large before they start interacting with other layers. By probing this regime we can investigate genuine BKT physics. The extent of this regime is controlled by $\Delta$ and $J'/J$, which can in turn be controlled by a uniform magnetic $z$-field and hydrostatic pressure, respectively. $\Delta$ can be increased in this way because the uniform $z$-field competes with the antiferromagnetic $S_i^z S_j^z$ coupling, whereas $J'/J$ can be increased because the pressure pushes the square lattices closer together.

In Article 1 of this cumulative thesis [42], shown in App. A, the author has used SSE with directed loops to simulate the low-temperature equillibrium properties of a 3D cubic anisotropic spin-1/2 Heisenberg magnet in an external magnetic field. This system is sign-problem-free, allowing us to reach system sizes with up to a million spins. We have investigated the specific Hamiltonian that models the molecular material CuPOF of Ref. [93], to provide theoretical backing for their experimental findings of BKT correlations. By tuning the magnetic field from zero to 17T, the system is interpolated between the layered XXZ system and a trivial ferromagnet. At small field strengths the effect of the magnetic field can be described as a field-induced exchange anisotropy [94], which we investigate with QMC in Fig. S4 of the Supplemental Material. At large field strengths, $\Delta$ effectively becomes negative, which induces a canting of the in-plane XY ordering into the $z$-direction.

To characterize the effects of a tiny $J'$ at different field strengths, we determined $T_{\mathrm{BKT}}$ at $J' = 0$ and $T_{\mathrm{LRO}}$ at $J' > 0$ for various magnetic fields. To this end we calculated the spin stiffness (2.11) with QMC and applied finite-size scaling theory. This procedure is illustrated in Fig. 3 of Article 1. Then, by fitting the experimentally determined spin-lattice relaxation rate to the correlation length (2.10), in the temperature regime where the XY model is realized, we can extract an experimental estimate of $T_{\mathrm{BKT}}$. The experimentalists determined $T_{\mathrm{LRO}}$ from the onset of the magnetization, which we have reproduced to some extent with QMC in Fig. S3 of the Supplemental Material. There we also determined the in-plane correlation length that triggers the Néel transition when it is on the order of $\infty\prime^{\in}$ lattice spacings, along with the uniform $z$-magnetization that captures the canting and the staggered $z$-magnetization which vanishes as expected. The theoretical and experimental $T_{\mathrm{BKT}}$ and $T_{\mathrm{LRO}}$ were determined across the full range of non-trivial field strengths, shown in Fig. 1 of Article 1. They display a non-linear dependency of the criticality on field strength, with the theoretical predictions behaving qualitatively the same as the experimental findings. So a simple sign-problem-free Heisenberg model is able to qualitatively capture the low-temperature equilibrium properties of CuPOF, showing a rich interplay of Néel and BKT correlations that can be tuned through an external field.

An interesting direction for further study would be to consider the vortex operators from Ref. [88]. This could provide a more detailed picture of the extent to which vortices are significant in the regime that we have studied. It would also give a glimpse of the fate of the vortices at the moment that the interlayer coupling becomes noticable, i.e. whether they instantly vanish or if there is perhaps some remnant.

The author of this thesis has performed all the numerical simulations to obtain the theoretical results of the article, and has performed the finite-size scaling to obtain the theoretical critical temperatures. The author has also performed the fitting of the universal scaling laws to the experimental data in Fig. 2 to obtain the experimental $T_{\mathrm{BKT}}$. Lastly, the author has written the article in conjunction with the other authors.

## 2.5 Commentary on Article 2

### Berezinskii—Kosterlitz—Thouless correlations in copper-based quasi-2D spin systems (Review Article)

Low Temp. Phys. 49, 819–826 (2023)

*D. Opherden*, *F. Bärtl*, *M. S. J. Tepaske*, *C. P. Landee*, *J. Wosnitza*, *H. Kühne*

A crucial aspect to the study in Article 1 is the experimental feasibility of accessing the XY regime of interest across a wide range of $\Delta$. The material CuPOF has several advantageous properties that make it the best known material for this task. This is because of its exceptional isolation of layers, with $J'/J \sim \mathcal{O}(10^{-4})$, while having a modest $J = 6.8$K such that a ferromagnet can be experimentally realized with a magnetic field [42]. Moreover, with an intrinsic anisotropy of $\Delta \approx 0.01$, the XY model is realized in a resolvable temperature range of a few hundred millikelvin above $T_{\mathrm{LRO}}$, which is itself around 2K for intermediate magnetic fields. Above the crossover temperature $T_{\mathrm{co}} > T_{\mathrm{LRO}}$ the exchange anisotropy washes out due to the thermal fluctuations, separating an isotropic from a planar regime [40, 41, 94]. Hence there are three relevant temperatures: the $T_{\mathrm{BKT}}$ that occurs for $J' = 0$, the actual $T_{\mathrm{LRO}}$, and $T_{\mathrm{co}}$. These essentially dictate the extent to which BKT correlations can be probed.

Article 2 of this cumulative thesis [43], shown in App. B, is a review article on the XY regime of CuPOF and other modern candidate materials. We consider the qualities that make for a suitable magnet in which field-tunable BKT correlations can be observed, as well as the methods of determining this, and we compare previous QMC studies with the experimental results. In Fig. 1 we compare the experimental $T_{\mathrm{LRO}}$ of various quasi-2D materials with different intrinsic anisotropies to the theoretical $T_{\mathrm{BKT}}$ from the QMC studies in Refs. [40, 41]. This is done at zero field strength. It gives the first indication that CuPOF is a good candidate for observing BKT correlations, since it is one of two materials for which $T_{\mathrm{LRO}}$ falls below the predicted $T_{\mathrm{BKT}}$.

We provide further support for this by comparing in Fig. 2 the experimental $T_{\mathrm{co}}$ to the theoretical $T_{\mathrm{co}}$ from Ref. [95], again at zero external field. These are both obtained from the minimum of the out-of-plane magnetic susceptibility, which was introduced in the QMC study of Ref. [95]. The experimental and theoretical results agree very well. For comparison we also show again the theoretical $T_{\mathrm{BKT}}$ from Ref. [41] along with the experimental $T_{\mathrm{LRO}}$. The section between the $T_{\mathrm{co}}$ and $T_{\mathrm{BKT}}$ curves can be considered as the region in which XY physics dominates.

If we look for a field-tunable magnet then only CuPOF remains as a suitable candidate. This can be seen in Table 1, where the relevant couplings are summarized. Only CuPOF and its analogs have $J \sim \mathcal{O}(1)$, such that they can be experimentally tuned through the entire range of non-trivial field strengths. In Fig. 3 we show the experimental magnetic susceptibility minima at various field strengths, illustrating how $T_{\mathrm{co}}$ can be increased by increasing the field strength. This supports the concept of field-induced anisotropy. To investigate this more closely from the theoretical side, we plot in Fig. 4 the $T_{\mathrm{co}}$ and $T_{\mathrm{BKT}}$ as determined by QMC in Ref. [94] for $\Delta = 0$ with variable field strength, together with the $T_{\mathrm{co}}$ and $T_{\mathrm{BKT}}$ for zero field and variable $\Delta$ from Refs. [40, 41]. The excellent agreement indicates that at low field strength the models are equivalent in many respects, which we already got a glimpse of in the supplemental material of Article 1.

It is clear from Articles 1 and 2 that the simple sign-problem-free anisotropic XXZ model captures the physics of a wide range of layered materials. This enabled QMC to be a driving force, as it can simulate the finite-temperature equilibrium properties of the Hamiltonian for system sizes that are inaccesible with any other known theoretical method. It provides the foundation for an extensive finite-size scaling analysis that gives highly accurate predictions of critical temperatures, to achieve the best possible comparison between theory and experiment.

The author of this thesis is responsible for the theoretical analysis and comparison with QMC. The author has furthermore assisted in writing the article.

## CHAPTER 3

# Iterative algorithms

The QMC algorithms from the previous chapter are only efficient in simulating the equillibrium properties of sign-problem free Hamiltonians. This excludes real-time evolution and the exotic equilibrium properties of e.g. frustrated magnets that frequently appear in experiments [96–99]. Outside of this, we have to consider the decomposition in (1.18) directly, without any mapping to a classical model. For example, we can alternatively calculate the trace in (1.16) as the average return amplitude from an imaginary-time-evolved state $P |\psi\rangle$ to its initial state $O |\psi\rangle$, for all bit strings $|\psi\rangle$. At zero temperature we only need to evolve a single state, analogous to the real-time evolution of a state.

The particular form of the time-evolution operator decomposition now becomes important, as for the discrete worldline QMC from Ch. 2, because in practice we have to use a finite $n$ truncation of (1.18). Furthermore, in Ch. 4 and Ch. 5 we will see that approximate decompositions with a small amount of exponentials are often crucial to modern time-evolution methods. Each exponential has to be applied to the state, incurring a computational cost and potentially decohering the state. Then the exact decomposition (1.18) leads to a completely mixed state.

In this chapter we will develop various alternatives to (1.18), where the time-evolution operator of a single segment is made more complex to improve the approximation accuracy for a fixed segmentation $n$. In Ch. 4 on tensor networks and Ch. 5 we will see that this is crucial for modern time evolution algorithms on both classical and quantum computers.

## 3.1 Approximate time evolution

Before we start looking for approximate decompositions, we have to define a measure of accuracy that is physically sensible. We do this with a *distance* between the targeted time-evolution operator $U$ and its approximation $C$. A natural option is the normalized completely bounded trace distance [100, 101],

$$d_{\mathrm{tbt}}(C, U) = \max_{\rho} \frac{1}{2} ||(C \otimes I)\rho(C^{\dagger} \otimes I) - (U \otimes I)\rho(U \otimes I)^{\dagger}||_{\mathrm{t}}. \tag{3.1}$$

The maximum is taken over arbitrary density matrices $\rho$ (satisfying $||\rho||_{\mathrm{t}} \leq 1$ by definition) that live in a tensor product of the physical Hilbert space with an auxillary system. The distance (3.1) is induced by the trace norm $||\rho||_{\mathrm{t}} = \mathrm{Tr}\sqrt{\rho^{\dagger}\rho}$ on density matrices, and the accompanying normalized trace distance $d_{\mathrm{t}}(\rho, \sigma) = ||\rho - \sigma||/2$ encodes the distinguishability of states $\rho$ and $\sigma$ [102]. The auxillary space is required to efficiently discriminate between $C$ and $U$. The identity in (3.1) acts on this auxillary space and is there solely to account for its presence during a discrimination task. When $C$ and $U$ are affected by decoherence, the distance (3.1) is a natural choice even when not considering the distinguishibility test [100], and analogously for time-evolution algorithms that rely on auxillary systems. The operational meaning of precision as defined by (3.1) is clear: It is the maximum extent to which the difference between $C$ and $U$ can induce distinguishible states that should be indistinguishable when $C$ and $U$ are equivalent. The errors of the decompositions that we will introduce below are defined with respect to this measure.

There is no closed form for calculating (3.1) [101], calling for a bound on (3.1) that is relatively easy to evaluate. This is essential for applications where we often need to quantify the distance between a time-evolution operator and its approximation, e.g. when attempting to minimize this distance in a variational algorithm. It can be shown that $d_{\text{tbt}}(C, U) < d_{\text{spec}}(C, U)$, where $d_{\text{spec}}(C, U)$ is the *spectral distance* between $C$ and $U$, which is the operator distance induced by the Euclidean norm on the Hilbert space of states [101]

$$d_{\text{spec}}(C, U) = ||C - U||_{\text{spec}} = \max_{|\psi\rangle} \sqrt{\langle\psi| (C - U)^{\dagger}(C - U) |\psi\rangle}. \tag{3.2}$$

Here $|\psi\rangle$ can be an arbitrary state, such that $d_{\text{spec}}(C, U) \approx 0$ only when $C|\psi\rangle$ and $U|\psi\rangle$ are close even in the worst case. Now, $d_{\text{spec}}(C, U)$ has a closed form: it is equal to the square root of the largest eigenvalue of $(C - U)^{\dagger}(C - U)$. This norm is also often used to quantify the error of analytical decompositions [103]. However, its calculation requires (partial) matrix diagonalization, which is expensive compared to matrix multiplication. Hence for the purpose of e.g. numerical optimization it is worthwile to look for a distance that involves matrix multiplication at most.

The Frobenius distance generalizes the Euclidean norm to operators, i.e.

$$d_{\text{F}}(C, U) = ||C - U||_{\text{F}} = \sqrt{\text{Tr}((C - U)^{\dagger}(C - U))} = \sqrt{\sum_{ij} |C_{ij} - U_{ij}|^2}, \tag{3.3}$$

and it is closely related to the spectral and trace distance, all being based on the eigenvalues of $(C - U)^{\dagger}(C - U)$. In particular, the trace distance of an operator is the sum of the square root of its eigenvalues, i.e. the sum of its singular values, the Frobenius distance is the square root of the sum of its eigenvalues, and the spectral norm is the square root of the largest eigenvalues. This indicates how the different distances prioritize states differently, with the trace and Frobenius distances averaging over the entire Hilbert space while the spectral distance only considers the worst case.

To calculate $d_{\text{F}}(C, U)$ we only need to perform matrix multiplication followed by a trace, and since clearly $d_{\text{spec}}(C, U) \leq d_{\text{F}}(C, U)$ we find the efficiently calculable bound $d_{\text{tbt}}(C, U) \leq d_{\text{F}}(C, U)$. Hence if we determine $d_{\text{F}}(C, U) = \epsilon$, we know that $d_{\text{tbt}}(C, U) \leq \epsilon$, giving us an intuitive measure of how close $C$ and $U$ are. We can normalize the Frobenius norm by averaging over the eigenvalues instead of summing them, to which end we divide the trace in (3.3) by the Hilbert space dimension $2^N$. This normalized measure introduces a notion of error, which we use as a cost function for the optimization procedure in Articles 4 and 5, as well as in the preprint [104].

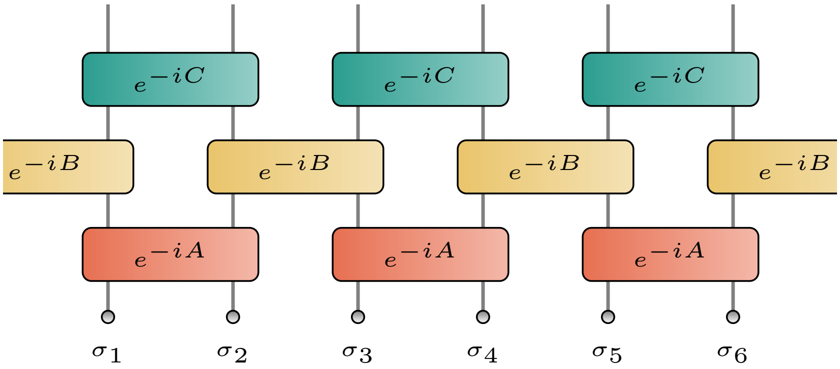


Figure 3.1: A translationally-invariant circuit $C$ consisting of general two-qubit gates $G$ laid out in a brickwall pattern.

## 3.2 First-order decomposition

Now we are ready to consider approximate decompositions, starting with (1.18). Trotter showed that (1.18) generalizes to the case where the operators are unbounded (e.g. for Hamiltonians of infinite lattice systems) [105], in which case we cannot write the closed form (1.18). Suzuki further extended it to the case of more than two operators and formulated it in the practical language of approximants [106], i.e. the $n$th approximant of (1.18) is

$$e^{-it\sum_j H_j} = \left(\prod_j e^{-itH_j/n}\right)^n + O(t^2/n). \tag{3.4}$$

The error depends solely on the norm of all nested commutators of the Hamiltonian terms [103]. These vanish when all the terms commute, in which case the error is zero. For instance, for the Heisenberg model we can easily use the triangle inequality to show that the error in (3.4) is equal to $||[H_A, H_B]||_{\text{spec}} t^2/(2n) + O(t^3/n^2)$. This explicit dependence of the error bound on the locality of the system under consideration, as quantified through the commutators of the Hamiltonian terms, can be used to prove rigorously that the amount of local exponentials required to reach a fixed precision $\epsilon$ scales linearly with system size [107]. In [103] a rigorous theory of Trotter commutator error bounds was introduced.

## 3.3 Variational compiling

More generally, we can build $C$ from an arbitrary product of exponentials, that we can design whichever way we want. Each *gate* $G$ in the product can be chosen to arbitrarily couple any amount of qubits, and the terms can be executed in any order. This is called a *circuit* or *product formula*. Then we can try to modify the couplings in each $G$ such that their product approximates the desired time evolution. One possible choice is the first-order decomposition (3.4), which consists of uniform gates. For practical reasons, we will only consider circuits $C$ consisting of gates $G$ that act at most on nn qubits, laid out in a *brickwall architecture*. We call these brickwall circuits, which arise naturally for Hamiltonians that satisfy (1.19), but in general we can use it for more complex Hamiltonians.

An example of such a circuit is depicted in Fig. 3.1, for a qubit chain with periodic boundary conditions (PBC). Every gate is in principle a general unitary, e.g. $G_1 = e^{-iA}$ with $A$ an arbitrary $16 \times 16$ complex matrix that can be varied to affect the approximation accuracy of $C$. Such *parameterized circuits* are essential to variational approaches, where the parameters are varied to find an optimal approximation accuracy for a given layout and amount of gates. This process is called *variational compiling*. Here it is essential that an efficiently calculable distance measure such as the Frobenius distance (3.3) is used. For manifest translational symmetry we choose all gates in a half-layer to be equal, serving to approximate a translationally-invariant time-evolution operator. Naturally, a translationally-invariant circuit can be executed for any amount of qubits. In Article 3 [108], Article 4 [109], and the preprint [104], we use this approach to numerically determine high-accuracy approximations of various time-evolution operators.

The gates in every half brickwall layer can be applied simultaneously, and more generally the brickwall architecture has the smallest circuit *depth* for a fixed number of gates. The depth is the maximum amount of gates encountered by a forward-moving path from the input to the output qubits. If we think of the vertical direction in Fig. 3.1 as time, then a brickwall circuit has the minimum qubit idling time for a fixed number of gates. Depth also bounds the extent to which gate errors can propagate, such that it is especially relevant in the presence of non-correctable errors. In Ch. 4 we will see how this applies to classical tensor network algorithms that undergo algorithmic decoherence on account of keeping calculations tractable, and in Ch. 5 we will see how it similarly applies to quantum computer algorithms that undergo physical decoherence. There it is desirable to use circuits with as small a depth and as little elementary two-qubit gates as possible.

When looking for a *shallow* circuit approximation $C$ to a time-evolution operator $U(t)$, i.e. containing a relatively small amount of layers of gates, a fundamental property that $C$ should reproduce is the distance $W(t)$ that information can travel in time $t$. One way to quantify this is through the out-of-time-ordered correlators (OTOCs) [110–112]

$$C^{\alpha\beta}_{i,j}(t) = ||[\sigma^{\alpha}_i(t), \sigma^{\beta}_j]||^2_F, \tag{3.5}$$

where $\sigma^{\alpha}_j(t) = U\sigma^{\alpha}_j U^{\dagger}$. The OTOCs vanish when information cannot travel from the spin at site $i$ to the spin at site $j$. By varying $i$ and $j$ at a time $t$ of interest, we get an idea of $W(t)$. The Lieb-Robinson velocity implies $W(t) \propto te^{-\lambda t}$ for local Hamiltonians [113], where $\lambda > 0$ signifies a leaking of the perfect lightcone.

Analogously, circuits have a maximum-velocity width $R(M)$, which can be determined by starting at an input qubit and taking the forward-moving path that yields the largest displacement in sites after $M$ layers [114]. A sensible approximation should have $R(M) \geq W(t)$. It is easily inferred from Fig. 3.1 that a brickwall circuit has $R(M) = 2M + 1$. In contrast, a sequential staircase architecture has $R(M) = N$, since by design it allows for transport from one side to the other with only a single layer [115].

In the preprint [104] we consider in detail the potential benefits of reducing the parameter count of fixed-depth shallow circuits. We achieve this by incorporating the local constraints and symmetries of various models directly into the variational circuit. This greatly reduces the amount of circuit parameters and hence makes the optimization much faster, allowing for the optimization of larger circuits. However, we found that it can reduce the expressibility of fixed-depth shallow circuits. In the case of the PXP model we were able to construct restricted circuits that have the brickwall architecture, and there it immediately follows that on account of the restriction we have the more than eightfold reduction to $R(M) = M/4$. As a result, the time $t$ where the lightcone of the targeted time-evolution operator is larger than $R(M)$ occurs relatively early, after which the approximation breaks down. This is illustrated in Fig. 8 of [104], where we compare the OTOCs for fixed circuit depth. However, at this $t$ the $R(M)$ of an

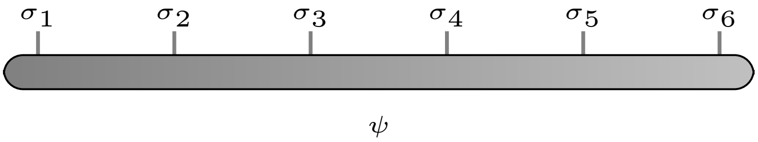


Figure 3.2: A quantum state of six qubits in the tensor product basis: $\psi_{\sigma_1\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6}$.

equally-sized unrestricted (i.e. only translationally-invariant) brickwall circuit is still much larger than the lightcone, and we found that it can encode high-accuracy approximations up to $2t$. This illustrates how circuit design is especially relevant when searching for accurate fixed-depth shallow circuits.

## 3.4 Computational complexity

To simulate the targeted dynamics we have to apply the circuit to a state, for which we need to sequentially apply its gates. In Fig. 3.2 we show a quantum state for six qubits, represented as a $(2, 2, \ldots)$ tensor $\psi_{\sigma_1\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6}$ to make the tensor-product structure of the Hilbert space explicit. Pictorally, it is already clear how the gates from Fig. 3.1 are applied to the state from Fig. 3.2.

The gate count $N_g$ of a circuit is a natural measure of its computational complexity. This is the amount of nn two-qubit gates required for $C$ to approximate the targeted unitary with precision $\epsilon$. Long-ranged gates or multi-qubit gates have to be decomposed into nn two-qubit gates before counting. This measure of complexity allows for comparison between widely different algorithms, and it is an especially relevant measure for noisy classical and quantum calculations. Importantly, this measure of complexity pertains only to the circuit, and not to the mechanism with which it is applied to a state. In particular, if we want to simulate arbitrary circuits on a classical computer, e.g. by considering their effect on arbitrary states, then applying a gate to a state is exponentially costly in $N$. In contrast, as we will discuss in Ch. 5, this is in principle only polynomially costly on a quantum computer. In Ch. 4 we will consider a formalism that allows for approximate Hilbert space calculations based on entanglement truncation. This lets us apply a circuit to a state on a classical computer with polynomial complexity, as long as the initial and propagated states are lowly entangled.

## 3.5 Higher order circuits

Before considering the numerical approach from Sec. 3.3 we will first explore analytical decompositions such as (1.18), to see what we are up against. To determine circuits for which the error scales as a higher power than (3.4), we have to consider more complicated approximants, such that the accuracy can be improved without having to use a finer segmentation $n$. To get a sense of how this will go, we note the celebrated Baker-Campbell-Hausdorff (BCH) formula, which states that for the case of two non-commuting exponentials

$$e^{-itA}e^{-itB} = e^{\sum_{n=1}^{\infty}(-it)^n Q_n}. \tag{3.6}$$

The coefficients $Q_n$ are the averages of the nested commutators available at order $n$ [116, 117], e.g. $Q_1 = A + B$, $Q_2 = [A, B]/2$, and $Q_3 = ([A, [A, B]] - [B, [A, B]])\,/12$. When $[A, B] = 0$ we get the scalar multiplication law $e^{-it(A+B)} = e^{-itA}e^{-itB}$.

With the BCH formula we can determine the coefficients $C_n$ of the dual Zassenhaus formula [117–119]

$$e^{-it(A+B)} = e^{-itA}e^{-itB}\prod_{k=2}^{\infty} e^{(-it)^k C_k}. \tag{3.7}$$

For example, $C_2 = [B, A]/2$ and $C_3 = ([[B, A], A] + 2[[B, A], B])/6$. These can be determined recursively [117], and it immediately provides a $m$th order approximation to $e^{-it(A+B)}$ by truncating the product [106], i.e.

$$e^{-it(A+B)} = \left(e^{-itA/n}e^{-itB/n}\prod_{k=2}^{m} e^{(-it)^k C_k/n^k}\right)^n + O\left(\frac{t^{m+1}}{n^m(m+1)!}\right). \tag{3.8}$$

Here we again introduced the segmentation $n$ to induce a perturbative parameter. By fixing $n$ such that $t/n \ll 1$ we then get the exact exponential for $m \longrightarrow \infty$. It generalizes naturally to the case of multiple non-commuting sets of Hamiltonian terms [106]. Because the amount of BCH and Zassenhaus coefficients increases exponentially with the order $m$, the amount of gates in (3.8) also grows exponentially with $m$. We can therefore expect that to increase $m$ for a general circuit approximation will require us to add an amount of gates that scales exponentially with $m$.

We find a similar structure for the approximation of a single exponential $e^{-itH}$. The definition (1.7) can be regarded as a $m \longrightarrow \infty$ approximation with $n = 1$, whereas the other well known definition

$$e^{-itH} = \lim_{n\longrightarrow\infty}(I - itH/n)^n, \tag{3.9}$$

instead has $m = 1$ and $n \longrightarrow \infty$. We can combine both definitions into the approximant [106]

$$e^{-itH} = \sum_{k=0}^{m}\left[\frac{1}{k!}\left(\frac{-itH}{n}\right)^k\right]^n + O(t^{m+1}/n^m), \tag{3.10}$$

which is the analog of (3.8). However, increasing $m$ by one now only requires one extra term in the expansion. This illustrates how the matching of orders is exponentially less complex when using an approximant that is a sum instead of a product as for (3.8).

## 3.6 Symmetric circuits

Nonetheless, the direct truncation approximation (3.8), first introduced by Suzuki [106], can be improved. In particular, shortly after Ref. [106], multiple reports surfaced of a simpler second-order circuit for the case of two non-commuting operators [120–122]

$$e^{-it(H_a+H_b)} = e^{-itH_a/(2n)}e^{-itH_b/n}e^{-itH_a/(2n)} + O(t^3/n^2). \tag{3.11}$$

Because of its symmetry, the circuit (3.11) is odd in $t$, i.e. its Zassenhaus formula is

$$e^{-itH_a/(2n)}e^{-itH_b/n}e^{-itH_a/(2n)} = e^{-itA/n}e^{-itB/n}\prod_{k=1}^{\infty} e^{(-it)^{2k+1}C_{2k+1}/n^{2k+1}}, \tag{3.12}$$

This implies the Taylor expansion

$$\left(e^{-itH_a/(2n)}e^{-itH_b/n}e^{-itH_a/(2n)}\right)^n = e^{-itH} + E_3(t)\frac{t^3}{n^2} + E_5(t)\frac{t^5}{n^4} + \ldots, \tag{3.13}$$

where the coefficient $E_k$ depends on $C_k$ and hence consists of nested commutators between $H_a$ and $H_b$.

This exposes the error structure of a symmetric circuit, and it illustrates how higher-order approximations can be obtained relatively easily by summing second-order circuits with different $n$. Such sums of powers of symmetric circuits are called *multiproduct formulas* [123, 124]. While a single circuit grows exponentially with $m$, with a multi-product the amount of gates grows polynomially with $m$, because it is straightforward to eliminate terms in the Taylor expansion. Essentially, the terms with different $n$ have different accuracies $\epsilon$ when executed on their own, and the multiproduct formula is an extrapolation to increase the accuracy, analogous to Richardson extrapolation [125].

For real-time evolution, generalizing from circuits to sums of circuits breaks the manifest unitarity of the approximation, since a linear combination of unitaries is generally non-unitary. Nonetheless, they have found wide application in classical simulation methods [123, 124].

## 3.7 Designing efficient algorithms

If we include the order of approximation $m$ as an algorithmic parameter, then product formulas may be considered inefficient, as they have an inherintly exponential scaling of the gate count $N_g$ with $m$. This is also reflected in its scaling with $1/\epsilon$, which is polynomial [126]. For multiproduct formulas it is instead polylogarithmic, which is optimal [126].

From the previous section it is clear how to design efficient algorithms, that scale optimally with $m$ and hence $1/\epsilon$. For instance, by first determining a circuit approximation to the targeted unitary, and decomposing its remainder into a linear combination of unitaries, the optimal scaling is achieved [127]. We could also implement the Taylor series directly, but implementing a sum of Hamiltonian powers with complex coefficients is a daunting task that appears unappealing from an algorithmic point of view. In principle, we can implement the Taylor series directly, by performing a classical simulation of the quantum algorithm of Ref. [128], but it leads to an inferior algorithm.

Of course, these asymptotic considerations are only relevant when gates can be executed perfectly, which is not the case in many relevant settings. An example is noisy quantum computing without error correction, which we will consider in much detail in Ch. 5. Here, the particular amount of gates in a circuit is crucial for the potential of a quantum algorithm, with too many gates leading to severe mixing with the enviroment.

## 3.8 Improved first-order bound

In the case of two non-commuting operators, the symmetrical second-order circuit (3.11) has the favorable property that it is structurally close to the first-order circuit (3.4). This can be utilized to determine some special properties of the first-order circuit. For example, tracing over (3.11) and utilizing the cyclicity of the trace, we find that the trace of the first-order circuit is second order, i.e.

$$\mathrm{Tr}\left[e^{-it(H_a+H_b)}\right] = \mathrm{Tr}\left[e^{-itH_a}e^{-itH_b}\right] + O(t^2/n^2). \tag{3.14}$$

It moreover induces an error scaling of the full operator that is significantly better than the $O(t^2/n)$ from (3.4), which was already noticed for particular local Hamiltonians in Ref. [129]. This was explained in Ref. [130] as self-inference of subsequent Trotter errors, but it can be derived easily by utilizing that a sequence of second-order circuits becomes a slightly modified first-order circuit with an additional half-layer. In Ref. [131] this was used to derive the improved scaling

$$e^{-it(H_a+H_b)} = e^{-itH_a}e^{-itH_b} + \min(C_2t^2/n, C_1t/n + C_3t^3/n^2), \tag{3.15}$$

with the coefficients $C_1 = \min(||H_a||, ||H_b||)$, $C_2 = ||[H_a, H_b]||/2$, and $C_3 = (\min(S) + \max(S)/2)/12$ with $S = \{||[H_1, [H_1, H_2]]||, ||[H_2, [H_2, H_1]]||\}$. Together, the results in Refs. [130, 131] imply that the first- and second-order Lie-Trotter-Suzuki circuits have the same asymptotic number of gates, namely $N_g \sim O\left(\sqrt{n^3t^3/\epsilon}\right)$.

Interference effects like this can always occur, and the rigorous Trotter bounds do not take them into account. As a result, emperical bounds are frequently observed [129, 132–135].

## 3.9 Recursive construction of higher-order circuits

Suzuki generalized the symmetric second-order circuit to the case of $r$ general Hamiltonian terms $H_j$ [50], i.e.

$$S_2(t) = \prod_{j=1}^{r} e^{-itH_j/2} \prod_{j=r}^{1} e^{-itH_j/2}. \tag{3.16}$$

He also realized that it can be used as a building block for a recursive construction of higher-order circuits that are simpler than (3.8) [51, 52]. In particular, we can generate a Lie-Trotter-Suzuki circuit of arbitrary order $2m$

$$e^{-it\sum_j H_j} = (S_{2m}(-it/n))^n + O(t^{2m+1}/n^{2m}), \tag{3.17}$$

by using a recursively defined symmetric approximant $S_{2m}$ such as

$$S_{2m} = \left[S_{2m-2}(p_m t)\right]^2 S_{2m-2}\left((1-4p_m)t\right)\left[S_{2m-2}(p_m t)\right]^2, \tag{3.18}$$

with coefficient

$$p_m = \frac{1}{4 - 4^{\frac{1}{2m-1}}}. \tag{3.19}$$

Due to the symmetrical structure of (3.18), odd powers in $t$ vanish in its Taylor expansion, such that the $(2m-1)$th order approximation is automatically also the $2m$th order approximation [51]. Therefore we only consider even powers in the definition (3.17).

The choice (3.18) is not unique, and there is in fact a smaller option

$$S_{2m} = S_{2m-2}(c_m t)S_{2m-2}\left((1-2c_m)t\right)S_{2m-2}(c_m t). \tag{3.20}$$

However, there is no simple analytical formula for $c_m$, which instead have to be determined by explicitly matching orders [51]. They are also in principle unbounded and generally blow up for $m \to \infty$, whereas $p_m < 1$. For concreteness we use (3.18), such that the gate count scales as $N_g \sim O(5^k)$. In any case, the circuits with $m > 2$ contain both forward and backward evolutions.

Some circuits cannot be derived via (3.18) and its generalizations to non-symmetric or complex-time circuits [53, 136, 137]. For example, the Zassenhaus truncation (3.8) and the Ruth formula with its generalizations to higher orders [138–140], which were developed for the simulation of classical mechanics. However, Suzuki introduced a more general mathematical framework [141] in which these can all be derived systematically [142]. Besides these analytical approaches, more recently there have also been numerical works that find better product formulas [143–146]. Furthermore, while we have been concerned with time-independent Hamiltonians, analogous circuits have been found for the time-dependent case, in which case we have to use time-ordered exponentials [147–153].

From a practical point of view, the symmetrical circuits from this section are usually the best. These only involve exponentials of the local Hamiltonian terms, whereas e.g. the Zassenhaus truncation (3.8) contains exponentials of polynomials of the local terms. Moreover, as we have seen in (3.13), the symmetry simplifies the error structure.

## 3.10 Computational complexity and no fast-forwarding

Let us determine the overall computational complexity of the $2m$th order Lie-Trotter-Suzuki circuit (3.17) for a Hamiltonian with $r$ terms that generally do not commute. The amount of exponentials $N_{\text{exp}}$ in (3.17) for segmentation $n$ and order $m$ is equal to $N_{\text{exp}} = 2r5^{k-1}n$. To fix the precision $\epsilon$, we can take the segmentation $n = \lceil 2^{\frac{1}{2m}}(5^{m-1}2rt)^{1+\frac{1}{2m}}/\epsilon^{\frac{1}{2m}}\rceil$ [154]. With this we get the complexity

$$N_{\text{exp}} \sim O\left(5^{2m}r(rt)^{1+\frac{1}{2m}}/\epsilon^{\frac{1}{2m}}\right). \tag{3.21}$$

To obtain the corresponding $N_g$ scaling we need to decompose every gate that acts on more than two nn qubits into nn gates. If the $H_j$ are $l$-qubit operators, then the bound in (3.21) gets at least a factor $l$.

The bound (3.21) is for the Lie-Trotter-Suzuki circuits of Hamiltonians with arbitrary $H_j$, and as such it might overestimate the complexity of particular cases. One such example was given in Sec. 3.8, with the first- and second-order circuits having equal gate complexity for a Hamiltonian with only nn couplings. For such a Hamiltonian the bound in (3.21) can be tightened. This is because the Trotter errors depend on the commutators between the Hamiltonian terms, such that a strong locality has an explicit effect on the bound. Specifically, it allows us to eliminate a factor $r$ from the bound [107]. Since now all gates are already nn, and $r \sim O(N)$, we get complexity

$$N_g \sim O\left(5^{2m}(Nt)^{1+\frac{1}{2m}}/\epsilon^{\frac{1}{2m}}\right). \tag{3.22}$$

For fixed order $m$, this bound shows that when we have a Lie-Trotter-Suzuki circuit that approximates the time-evolution operator at time $t$ with accuracy $\epsilon$, then to reach the same accuracy at time $2t$ we need to more than double $N_g$. In terms of the segmentation $n$, it means that when $t$ is doubled we need to at least double $n$. Strict doubling only occurs in the limit of infinite order $m \to \infty$, when every segment is evolved exactly. However, because $N_g$ scales exponentially with $m$, the bound (3.22) has a minimum in $m$, which is derived in Ref. [154]. Hence the scaling is always superlinear in $Nt$.

A linear scaling $N_g \sim \mathcal{O}(t)$ is optimal, which is called the *no-fast-forwarding* theorem. It lies at the heart of iterative algorithms, where time is closely connected to the specifics of the algorithm. It has been rigourously shown to hold for the iterative approximation of sparse Hamiltonians in general [154]. There are particular simple Hamiltonians where asymptotic fast-forwarding is attainable with an iterative algorithm [155–157]. Analogously, the linear scaling $N_g \sim \mathcal{O}(N)$ is optimal for local Hamiltonians such as the Heisenberg model (1.2), which can also be shown rigorously [158]. It is intuitively clear that the combined scaling $N_g \sim \mathcal{O}(Nt)$ is optimal, simply because we are simulating one physical system with another. As discussed in Sec. 3.5, the bound (3.22) scales suboptimally in $1/\epsilon$ due to the exponential complexity of eliminating orders with a product formula. The overall optimal scaling for local Hamiltonians $N_g \sim \mathcal{O}(Nt\text{poly}(\log(1/\epsilon)))$ can be achieved by utilizing the Lieb-Robinson bounds [113, 159] together with e.g. sums of circuits [158].

To avoid the fundamental no-fast-forwarding bound $\mathcal{O}(t)$ we need to use non-iterative methods, where time is simply a parameter of the targeted unitary. Concretely, we fix the gate count in our circuit approximation and increase the precision by modifying its gates [160–163]. This can be done with the variational compiling approach from Sec. 3.3. It seems especially promising for circuits that have a relatively small gate count, since there it seems feasible that significant improvements can be made relative to iterative methods. This usually comes at the cost of numerical optimization, where the gates are parameterized and then taylored to particular Hamiltonian for a particular $N$ and $t$.

Since numerical optimization limits the non-iterative approach to small $N_g$, we cannot probe its asymptotic behavior in $t$ for constant $\epsilon$. In practice, we perform optimization up to some thresholds $N_g$ and $t$, and to go beyond these we use the optimized circuits iteratively. As such, it obeys the asymptotic non-fast-forwarding bound. Nonetheless, it could improve the scaling of the Lie-Trotter-Suzuki circuits, as well as its offsets and prefactors. This is important for practical use, as we will discuss in Ch. 4 and Ch. 5.

## 3.11 Commentary on Article 3

### Optimal compression of quantum many-body time evolution operators into brickwall circuits

SciPost Phys. 14, 073 (2023)
*M. S. J. Tepaske*, *D. Hahn*, *D. J. Luitz*

In Article 3 of this cumulative thesis [108], which is shown in App. C, we consider the low-$N_g$ performance of the Lie-Trotter-Suzuki circuits with various orders of approximation, and compare this with variationally-compiled parameterized brickwall circuits. We did this for the nn Heisenberg model (1.2) and the nn + nnn Heisenberg model, where the nnn coupling is equally strong to the nn coupling. The nn Heisenberg model is naturally simulated on a quantum computer such as IBM's guadalupe [164], where a loop of 12 qubits can be formed, on which we can act with nn CNOT gates, arbitrary single-qubit $z$-rotations, $\sigma_x$ and $\sqrt{\sigma_x}$.

For the nnn Heisenberg model we need to insert additional SWAP gates to implement the nnn gates that appear in the Lie-Trotter-Suzuki circuits, so that they no longer have the brickwall architecture. We compared these with parameterized brickwall circuits of nn gates, that were numerically optimized to have minimal Frobenius distance (3.3) to the time-evolution operator. The calculations were performed with the tensor network machinery that we will describe in Ch. 4. We chose every gate to be a general unitary, to probe the maximum expressibility of the brickwall architecture, outside of any hardware-specific constraints. Moreover, we used open boundary conditions (OBC) since it is more flexible than PBC as far as real quantum computers go, and it allows us to use efficient tensor network methods.

We find that, for both Hamiltonians, the shallow optimized circuits are significantly better than the Lie-Trotter-Suzuki circuits, for a fixed number of nn two-qubit gates. Hence, at the cost of numerical optimization, we can effectively eliminate the SWAP gates from the Lie-Trotter-Suzuki circuits, and instead we can use the simplest possible architecture of nn gates, while improving the accuracy for a fixed number of gates. Moreover, by stacking the optimized circuits we find the scaling from Ref. [131], i.e. iteratively they behave like second-order Lie-Trotter-Suzuki circuits. Nonetheless, the offsets are significantly smaller for the parameterized circuits, requiring multiple times as little resources to achieve a precision threshold.

We also find that the variational circuits can break the no-fast-forwarding bound. Specifically, for the small circuits and the particular models and times that we have considered, when we have a brickwall circuit of $N_g$ two-qubit gates that simulates up to time $t$ with precision $\epsilon$, then we can often variationally determine a circuit with $2N_g$ two-qubit gates that simulates up to time $2t$ with a precision better than $\epsilon$. This can only be seen implicitly from Fig. 1 in Article 3. However, recently we also considered the potential benefits of incorporating local constraints and symmetries of the targeted Hamiltonian directly into the variational circuits [104]. In that preprint we explicitly consider the compilation performance as a function of time, with different curves corresponding to different circuit depths. Then it can be directly read off from the plot if a doubling of $M$ can more than double $t$ at fixed $\epsilon$. For example, in Fig. 3 of Ref. [104] we optimize circuits for the PXP model, and there it can be seen that going from $M = 4$ to $M = 8$ allows us to go from $t = 1$ to $t = 2.3$ for the error threshold $\epsilon = 10^{-2}$. In contrast, the Lie-Trotter-Suzuki circuits are observed to adhere to the fast-forwarding bound.

The author of this thesis has developed the framework of variational compiling, with the goal of finding shallow circuit approximations of quantum many-body time-evolution operators, together with the other authors. He has performed all the numerical simulations, carried out the theoretical analyses, and wrote the article together with the other authors.

CHAPTER 4

# Tensor networks

To simulate quantum time evolution with one of the circuit approximations from Ch. 3 on a classical computer generally requires us to perform matrix-vector multiplication in a Hilbert space that grows exponentially with system size. When the sign problem prevents us from using the QMC algorithms from Ch. 5, and there is no special structure such as integrability that allows for analytic calculation, we have to resort to clever approximation algorithms that work within a subspace of the full Hilbert space which grows merely polynomially with system size.

One class of dynamics that can be simulated efficiently in this way is that of locally-correlated states. This can be uncovered by formulating quantum states and operators in an explicitly local manner, namely as a *tensor network*, which gives control over the range of correlations that can be captured. In this chapter we consider this in some detail, to motivate Article 4 [165], shown in App. D, which we will discuss in Sec. 4.10. There we introduced a new class of 3D tensor networks along with an associated time-evolution algorithm. As we will discuss in this chapter, imaginary time-evolution of tensor networks allows for the efficient simulation of the low-lying eigenstates of local Hamiltonians, which are predominantly locally correlated. Real-time evolution is generally restricted to small times, since for a generic local Hamiltonian the correlations spread through the entire system until it is correlated equally on all length scales.

Nonetheless, by utilizing tensor networks for circuit optimization at the bleeding edge of what is possible on a modern classical supercomputer, we can determine circuits that can in principle be applied a few times on a quantum computer to generate classically intractable states. This was part of the motivation behind Article 3 [108] and Article 5 [109], which are shown in App. C and App. E, respectively, and discussed in Sec. 3.11 and Sec. 5.10. We will discuss this in more detail in Ch. 5 on quantum computing.

## 4.1 Manifest locality

The brickwall architecture of Fig. 3.1 consists of local gates, each acting on two neighboring qubits. Intuitively, such a gate should only affect the state locally, but this is not immediately clear when performing calculations in the representation of Fig. 3.2. There the tensor-product structure of the many-body Hilbert space is used to write the six-qubit state $|\psi\rangle$ as a $(2,2,2,2,2,2)$-dimensional *tensor* $\psi_{\sigma_1\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6}$. Then it is already clear that we do not need to construct a $2^N \times 2^N$ unitary for the circuit,

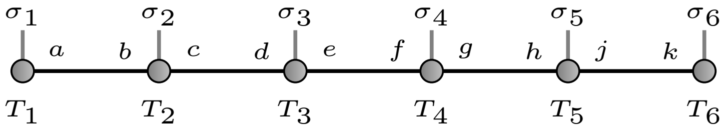


Figure 4.1: A tensor network state for six qubits.

in order to apply it to $|\psi\rangle$ through a matrix-vector product. Instead, we only need to sequentially apply a $4 \times 4$ matrix with the state reshaped as a $4 \times 2^{N-2}$ matrix. This reduces the computational complexity from $O(2^{2N})$ to $O(2^{N+2})$, which is however still exponentially complex. One such gate application immediately modifies all elements of $\psi_{\sigma_1\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6}$, obfuscating its locality.

This motivates us to look for a representation where $\psi$ can be modified in a manifestly local manner, which will naturally lead us to an approximation scheme based on the truncation of entanglement. In the same spirit as for the circuits from Ch. 3, we consider a 1D *tensor network state* (TNS) $|\psi\rangle$ that is a simultaneous *contraction* of local tensors $T^{\sigma_j}_{j,ab}$, i.e.

$$\psi_{\sigma_1\sigma_2\sigma_3\sigma_4\sigma_5\sigma_6} = \sum_{abcdef} T^{\sigma_1}_{1,a} T^{\sigma_2}_{2,ab} T^{\sigma_3}_{3,bc} T^{\sigma_4}_{4,cd} T^{\sigma_5}_{5,de} T^{\sigma_6}_{6,ef}. \tag{4.1}$$

A *virtual* index $a$ corresponds to a $X$-dimensional artificial Hilbert space with orthonormal basis $|a\rangle$, where $X$ is commonly called the *bond dimension*, and each lattice site now hosts a direct product of one qubit space with one or two virtual spaces, e.g. at site 2 we have the basis $|\sigma_2\rangle\,|a\rangle\,|b\rangle$. In higher-dimensional tensor networks there are usually more virtual spaces per site, depending on the lattice connectivity and the type of tensor network. The tensor $T^{\sigma_2}_{2,ab}$ maps the virtual spaces $a$ and $b$ onto the qubit space $\sigma_2$, allowing us to generate qubit states from virtual states [166, 167].[1]

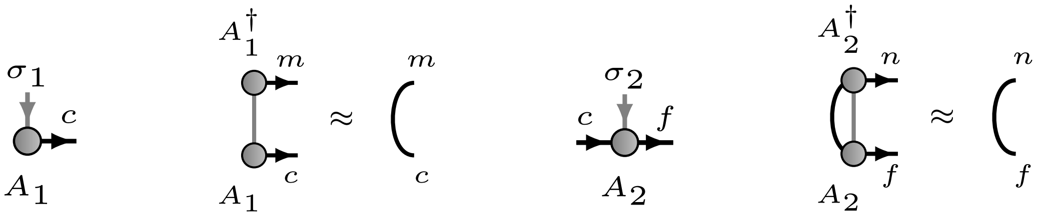
A contraction such as $\sum_b T^{\sigma_2}_{2,ab} T^{\sigma_3}_{3,bc}$ creates a maximally-entangled pair of virtual particles $a$ and $c$ between neighboring sites. Hence the state in (4.1) has a maximally-entangled virtual pair for every bond of the lattice, which are mapped to a $N$-qubit state by the $T_j$. This is illustrated in Fig. 4.1, where connected tensors imply a contraction. With this construction, it is immediately clear that the generated spin states are entangled hierarchically, according to the lattice geometry. For example, sites 1 and 3 are only entangled because they are both entangled with site 2. Depending on the dimension $X$ of the virtual spaces, this can severely restrict the range across which qubits can be entangled.

## 4.2 Isometric tensor networks

To uncover the connection with entanglement, let us consider the center bond in Fig. 4.1, which splits the chain into subsystems $A$ and $B$. We will convert the TNS into a special representation with respect to this bond, such that the tensor networks on $A$ and $B$ form orthonormal bases. For this we only need to use the gauge freedom that exists on each bond on account of the ability to insert $\sum_c U^\dagger_{ac} U_{cb} = I_{ab}$ on every virtual bond $ab$. Starting with the leftmost two tensors in Fig. 4.1, we insert this and determine

---

[1] A circuit like that in Fig. 3.1 is clearly a tensor network by design, consisting of gate tensors with only physical spaces.



(a) (b) (c) (d)

Figure 4.2: The isometric tensors used to construct the orthonormal basis on subsystem $A$. In panel (a) we show the corner tensor, which satisfies the condition shown in panel (b), i.e. it reduces to a virtual identity upon contraction with the conjugate qubit space. In panel (c) we show a bulk tensor, which satisfies the condition shown in panel (d), now also requiring the contraction of one virtual space with its conjugate. The isometric tensors are analogously defined for subsystem $B$, for which we only have to reverse the direction of the virtual spaces. The arrows signify the direction of collapse of an isometry: by contracting over the incoming legs of a tensor we obtain an identity on every outgoing leg.

$U_{ac}$ such that $\sum_a T^{\sigma_1}_{1,a} U^\dagger_{ac} = A^{\sigma_1}_{1,c}$, where $A^{\sigma_1}_{1,c}$ satisfies the orthonormality constraint

$$\sum_{\sigma_1} A^{\sigma_1}_{1,c} A^{\dagger\sigma_1}_{1,m} = I_{cm}. \tag{4.2}$$

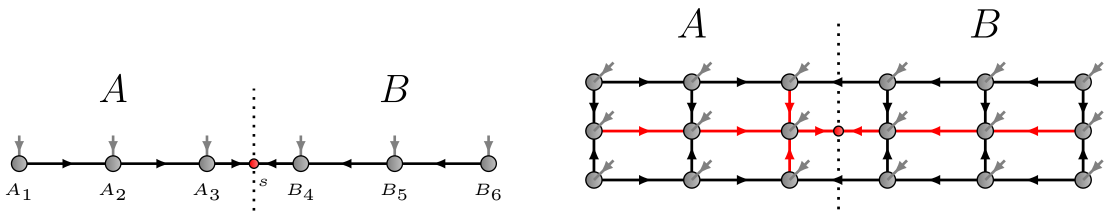
Here $I_{cl}$ is the identity operator of the first virtual space. In Fig. 4.2a we show a diagram for $A_1$, and in Fig. 4.2b we show the constraint (4.2). Here the arrows indicate the specifics of the isometry: By contracting over the incoming arrows we get identities for the outcoming arrows.

The appropriate $U$ follows from a QR decomposition of $T_1$, for which fast numerical linear algebra routines exists, and it immediately yields $A_1$ and $U_{cb}$. Then we absorb $U_{cb}$ into $T^{\sigma_2}_{2,bd}$, and we invoke the gauge freedom of the next bond. This time we choose $U^\dagger$ such that it transforms $T_2$ into an *isometric* tensor $A_2$, which satisfies

$$\sum_{\sigma_2 c} A^{\sigma_2}_{2,cf} A^{\dagger\sigma_2}_{2,cn} = I_{fn}. \tag{4.3}$$

In Fig. 4.2c we show $A_2$ and in Fig. 4.2d we show the isometry constraint (4.3). We continue this process until all of the tensors in $A$ are isometric, and for the last $T$ we keep the final gauge tensor $U$. This yields the tensors inside area $A$ in Fig. 4.3a.

Now we carry out the same procedure for $B$, starting from the rightmost tensor, so that finally we get the *isometric tensor network* state (isoTNS) shown in Fig. 4.3a. Upon contraction with its conjugate, it becomes clear that the isometric tensors serve to implement $X$-dimensional orthonormal bases $|A_n\rangle$ and $|B_m\rangle$. Concretely, starting with the outermost tensors, we invoke the orthonormality condition from Fig. 4.2b. Afterwards, we invoke the isometry condition from 4.2d until we reach the center, leaving us with identity matrices of the virtual spaces of the middle bond. The virtual tensor $s$ on the middle bond, shown in red, is the product of the two final gauge tensors from the orthonormalization of $A$ and $B$. It is



(a)

(b)

Figure 4.3: In panel (a) we show an isoTNS in 1D, with the singular values located at the middle bond. The isometric tensors $A_j$ and $B_j$ implement the orthonormality of the bases on subsystem $A$ and $B$, respectively. Upon contraction with the conjugate isoTNS, the isometric tensors sequentially contract to identities, starting from the corners. They eventually reduce to a pair of identities that act on the virtual spaces of the middle bond. Here the singular values are located, and on account of the orthonormality on $A$ and $B$ they directly encode the entanglement between $A$ and $B$. In panel (b) we show a 2D isoTNS, with the singular values on the middle bond. The isometric tensors are now slightly more complex than in panel (a), but they implement orthonormal bases for the middle bond in the same fashion. Unlike in panel (a), we now truncate the entanglement across a system-wide boundary by truncating only one of many bonds that crosses it.

a $X \times X$ matrix, which we diagonalize to find the *singular value decomposition* (SVD)

$$|\psi\rangle = \sum_n \lambda_n |A_n\rangle |B_n\rangle, \tag{4.4}$$

where the $X$ eigenvalues $\lambda_n$ are the *singular values*. They satisfy $\sum_n |\lambda_n| = 1$ because $|\psi\rangle$ is normalized.

To see their significance we construct the density matrix $\rho = |\psi\rangle\langle\psi|$ and trace out the $B$ subsystem

$$\rho_A = \mathrm{Tr}_B\left[|\psi\rangle\langle\psi|\right] = \sum_n |\lambda_n|^2 |A_n\rangle\langle A_n|. \tag{4.5}$$

This shows that $|\lambda_n|^2$ are the eigenvalues of $\rho_A$, already indicating that the singular values control the entanglement between subsystems $A$ and $B$. This is quantified by the *entanglement entropy* [168]

$$S = -\mathrm{Tr}\left[\rho_A \ln \rho_A\right] = -\sum_n |\lambda_n|^2 \ln\left(|\lambda_n|^2\right). \tag{4.6}$$

So in the isoTNS representation it is immediately clear that the virtual particles of the middle bond control the entanglement between $A$ and $B$ in Fig. 4.3a, on account of the isometric tensors constituting orthonormal bases for the bond. Their Hilbert space dimension $X$ determines the maximum entanglement that the TNS can encode between $A$ and $B$, which occurs for the maximal superposition $\lambda_n = 1/\sqrt{X}$, yielding $S = \ln X$. A product state instead has $\lambda_1 = 1$ and hence $S = 0$.

The bond gauge freedom can be used to move the singular values to every bond, indicating that the 1D isoTNS in Fig. 4.3a is simply a particular gauge of the arbitrary TNS in Fig. 4.1. Now it is clear that the TNS provides an entanglement-controlled subspace of the full many-body qubit Hilbert space, which grows upon increasing the bond dimensions.

## 4.3 Circuit application

Applying a nn gate $G$ to a TNS is manifestly local, since we can ignore all tensors except for those involved. First, we perform a SVD of $G$, which yields

$$G = \sum_{l=1}^{4} \kappa_l \left|i_l\right\rangle \left\langle i_l\right| \otimes \left|j_l\right\rangle \left\langle j_l\right| . \tag{4.7}$$

Here $\kappa_l$ are operator singular values, now satisfying $\sum_n |\kappa_n|^2 = 16$ on account of unitarity. The states $|i_l\rangle , |j_l\rangle$ are 2-dimensional orthonormal bases for neighboring sites $i$ and $j$, considered as a two-qubit system. Because $G$ acts only on two qubits, the virtual space is at most 4-dimensional. The representation (4.7) is called a *tensor network operator* (TNO), and it can be defined for many-body operators analogously to how did for states. Now, a virtual space controls the *operator entanglement entropy* (opEE) between two subsystems [169]. If we rescale $\kappa_n$ such that that they are normalized, yielding $\tilde{\kappa_n}$, then we can define it as in (4.6), i.e.

$$\text{opEE} = -\sum_n |\tilde{\kappa}_n|^2 \ln\left(|\tilde{\kappa}_n|^2\right) . \tag{4.8}$$

As such, a TNO with large $X$ can encode a large amount of opEE, which we can truncate optimally by first bringing it into the isoTNO gauge.

To see the significance of the opEE, we apply $G$ (4.7) to the middle bond of the isoTNS in Fig. 4.3a, yielding

$$GA_3\lambda B_4 = |f\rangle |h\rangle \sum_{\sigma_3\sigma_4} \sum_{l=1}^{4} \sum_{g=1}^{X} \kappa_l \lambda_g \left|i_l\right\rangle \left\langle i_l|\sigma_3\right\rangle A^{\sigma_3}_{3,fg} \left|j_l\right\rangle \left\langle j_l|\sigma_4\right\rangle B^{\sigma_4}_{4,gh}. \tag{4.9}$$

We now have the $(4, 2, X, \tilde{X})$ tensors $\sum_{\sigma_3} \langle i_l|\sigma_3\rangle A^{\sigma_3}_{3,fg}$ and $\sum_{\sigma_4} \langle j_l|\sigma_4\rangle B^{\sigma_4}_{4,gh}$, where $X$ is the bond dimension of the contracted space $g$ and $\tilde{X}$ the bond dimension of the uncontracted spaces $f, h$. These tensors are no longer isometric, but we can use the gauge freedom of the direct product of the virtual spaces $l$ and $g$ to recover the isometric representation. Concretely, we insert $\sum_{k=0}^{\mathcal{X}} U_{(mo)k} U^{\dagger}_{k(np)} = I_{(mo)(np)}$ twice, each with a different gauge transformation, $U$ and $U'$, which are $4X \times \mathcal{X}$ isometries. This implies $\mathcal{X} \leq 4X$, with the fourfold increase occuring for unitary gauge transformations. We choose $U$ and $U'$ such that the tensors again become isometric, which imposes the additional condition $\mathcal{X} \leq 2\tilde{X}$. Consequently, for uniform bond dimension $X = \tilde{X}$ we cannot saturate the bound $\mathcal{X} = 4X$, instead at most $\mathcal{X} = 2X$. In this case, the maximum $S$ increases with ln(2). In contrast, for $X \geq 2\tilde{X}$ we can choose $\mathcal{X} = 4X$, increasing the maximum $S$ with ln(4), which is the fundamental maximum increase in entanglement due to a nn qubit gate.

The non-isometric remainder $\sum_{lg} U_{k(lg)} \kappa_l \lambda_g U'^{\dagger}_{(lg)d}$ is diagonalized to determine the $\mathcal{X}$ singular values $\mu_k$. We have then recovered the isometric representation

$$GA_3\lambda B_4 = |f\rangle |h\rangle \sum_{k=1}^{\mathcal{X}} \mu_k \left|\sigma_3'\right\rangle A'^{\sigma_3'}_{3,fk} \left|\sigma_4'\right\rangle B'^{\sigma_4'}_{4,kh} = A_3' \mu B_4'. \tag{4.10}$$

This gives a glimpse of how opEE relates to the generation of entanglement in a state, and how this is naturally captured by a TNS. For example, if we apply a gate with $\tilde{\kappa}_0 = 1$, i.e. a direct product of two one-qubit unitaries and hence with opEE = 0, to a state with $X = \tilde{X} = 1$, then the result clearly has $\mathcal{X} = 1$ as well. We can apply such a gate at every bond with only a tiny $N$-independent cost, which would instead be exponential in $N$ if we would work directly in the many-body Hilbert space.

## 4.4 States and operators of local Hamiltonians

The maximum opEE of a $N$-body unitary is $\ln(2^{N/2})$, which can be encoded by an isoTNO when $X = 2^{N/2}$. Therefore, exactly representing an arbitrary many-qubit operator as an isoTNO is exponentially complex in the amount of qubits that it acts on. Hence it is clear that the opEE quantifies the non-locality of a unitary, which is explored in Ref. [169]. Analogously, it quantifies how much entanglement a unitary generates on average for all its input states. Real-time evolution is known to generate linear growth of $S$ for generic Hamiltonians and initial product states [170]. For low $t$, it entangles states only locally, but for large $t$ it yields states that are entangled across the entire system. In this case, an exact representation as a TNS requires a bond dimension that grows linearly with time, until it reaches the maximum entropy. In some cases this process is relatively slow, e.g. logarithmically for many-body-localization [171] and many-body scarring [172, 173], but e.g. for the Heisenberg model (1.2) it is linear [174].

States that are predominantly locally entangled can be encoded by a TNS with $X \sim \mathcal{O}(\text{poly}(N))$ [175]. One group of such states are the low-lying eigenstates of local Hamiltonians, in any dimension, for which $S$ scales only with the boundary size between $A$ and $B$ [176, 177]. This is essentially because only the qubits close to the boundary contribute to the entanglement across it. Unlike with the QMC algorithms from Ch. 2, there is no classical sampling involved and hence no sign problem, such that in principle any Hamiltonian can be simulated. This makes it seem plausible that imaginary-time evolution can be performed efficiently for local Hamiltonians by utilizing TNS.

When approximating a quantum many-body time-evolution operator $U$ with a brickwall circuit $C$, such as that in Fig. 3.1, the gates must be able to generate sufficient entanglement across each bond. We now know that an entanglement entropy of $S$ across the middle bond requires us to apply at least $S/2$ gates to the corresponding qubits. Moreover, with the brickwall architecture of Fig. 3.1, non-neighboring qubits are only entangled indirectly, as it encodes a lightcone of correlation spreading. It is therefore clear that we need at least enough brickwall layers to reach the non-locality of the targeted $U$. These properties give a lower bound on the minimum circuit size that is required for a suitable approximation. For local Hamiltonian, correlations also spread in a lightcone, which is known as the Lieb-Robinson bound [113, 159]. The spreading of a particular Hamiltonian can be quantified by calculating out-of-time-ordered correlators (OTOCs) [110–112]. We will consider this in more detail in Ch. 5.

These considerations are mostly relevant when we are concerned with small circuits, where we want to pack as much information into every gate. For a Lie-Trotter-Suzuki circuit with $t/n \sim \mathcal{O}(10^{-2})$ the gates are close to the identity and therefore have small opEE, which is a consequence of the fact that an accurate Lie-Trotter-Suzuki circuit relies on the iteration over many tiny segments. This incremental approach becomes unfavorable when each gate is accompanied by decoherence. Then, if the incurred errors are not corrected, the application of many near-identity gates would simply lead to a completely mixed state. This highlights the objective of *compressing* accurate Lie-Trotter-Suzuki circuits into smaller circuits, as discussed in Article 3 of Sec. 3.11.

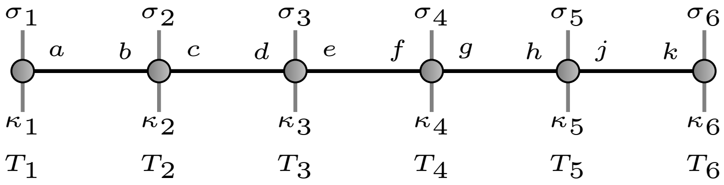


Figure 4.4: A tensor network operator for six qubits.

## 4.5 Efficient time evolution

To keep calculations tractable, we have to curb the growth of the bond dimension in (4.10) when $\mathcal{X}$ surpasses some threshold $X$. The numerical linear algebra algorithms that implement (4.10) always saturate the bound $\mathcal{X} = 2X$ or $\mathcal{X} = 4X$ on the size of the new virtual space, but for small times many of the new singular values will be zero. As such, we can omit them without losing any information, such that the actual SVD is obtained. If none of the new singular values are negligible, we have to omit physical information when truncating back to the threshold $X$, thereby inevitably degrading the TNS approximation. However, with (4.6) in mind we can omit the virtual states with the smallest singular values, such that the loss of entanglement is minimal.

By maintaining a fixed threshold $X$ we can use the procedure outlined in Sec. 4.3 to approximately execute any circuit with complexity $O(\mathrm{poly}(N))$. We only need to move around the singular values, which is free on account of the bond gauge freedom, and apply every gate with optimal truncation. This algorithm is called time-evolving-block-decimation (TEBD) [168], and it lets us do real- or imaginary-time evolution with one of the Lie-Trotter-Suzuki circuits from Sec. 3.5. We can initialize the TNS as a product state and then carry out TEBD, e.g. to simulate quench dynamics for real-time evolution or to determine a ground-state approximation at large imaginary time. Importantly, the form (4.4) also allows us to calculate an expectation value for the middle two qubits, e.g. $\langle\psi|\,\sigma_3^z\sigma_4^z\,|\psi\rangle$, using only $A_3$, $\lambda$, and $B_4$. As such, we never need to perform the full contraction in (4.1), thereby avoiding the $O(2^N)$ complexity as long as $X \ll 2^N$.

As mentioned earlier, for real-time evolution a constant $X < 2^{N/2}$ at some evolution time becomes insufficient to reach fixed accuracy as a function of $N$, after which it has complexity $X \sim O(2^{N/2})$ to accomodate for additional long-range entanglement. The alternative is to let the accuracy degrade, which effectively decoheres the TNS [178, 179]. As a result, real-time TEBD is generally limited to short times for generic local Hamiltonians, unlike imaginary-time evolution where the infinite-time evolution yields a locally-entangled state. When we are instead interested in simulating time evolution under decoherence, the artificial decoherence due to truncation can mimic the physical decoherence of the system under investigation. Then to simulate e.g. a noisy quantum computation from Ch. 5 we might only need a relatively small bond dimension.

We can also use TEBD to determine a TNO approximation to a circuit. To obtain a TNO we simply add an extra qubit space to each tensor in Fig. 4.1, yielding the tensor network in Fig. 4.4. As we found earlier, in the isoTNO representation the singular values of each bond encode the opEE across that bond, which we bound to keep calculations tractable. We initialize the TNO as the identity by setting $X = 1$ and $T_{j,00}^{\sigma\kappa} = I^{\sigma\kappa}$ for all sites $j$. Then we use TEBD to apply the circuit to the $\sigma_j$, truncating the opEE whenever necessary, yielding a TNO approximation of the circuit.

Tensor networks are taylored to circuit approximations like those from Ch. 3, where entanglement is explicitly generated only locally, and as such they are naturally simulated with TEBD. In contrast, with multi-product formulas entanglement is additionally generated globally, due to the non-unitarity of summing states.[2] Adding two tensor networks doubles $X$ on all bonds [175], making it twice as costly as TEBD. Truncating back to $X$ induces an *algorithmic decoherence*, since there are now truncations occuring outside of the gate application in (4.10). In principle there can then exist a maximum in the accuracy as a function of segmentation, in contrast to vanilla 1D TEBD. Because of these unfavorable properties it yields an inferior tensor network algorithm. This summation of tensor networks can also be realized stochastically, if we consider it as a sum of circuits instead of states. With a convex sum we can sample the circuits in a straightforward manner, and by applying each sample to a state of interest we can estimate expectation values.

Finally, an isometric tensor network can also be updated variationally, by restricting the tensor parameters to e.g. a Stiefel manifold and minimizing an objective function [181–185]. With this in mind, the differential programming machinery for general tensor networks [186] was extended to isoTNS [187], opening the door to efficient implementation with modern automatic differentiation frameworks such as PyTorch [188].

## 4.6 Higher-dimensional isometric tensor networks

Let us see if we can generalize the isoTNS construction from Sec. 4.2 to higher-dimensional tensor networks. If we take a tensor network that is arranged and connected as a square lattice with OBC, such as in Fig. 4.3b, then it is immediately clear that the local truncation of entanglement will be less straightforward. In particular, dividing the lattice in two equal halves $A$ and $B$ now yields a system-wide boundary. In the spirit of 1D isoTNS, we would like to find a representation such that we can truncate the entanglement between $A$ and $B$ by truncating only a single bond that crosses the boundary. This is in contrast to 1D, where the boundary contains a single bond. Alternatively, we could use a *tree* TNS, where the connectivity of the tensor network is chosen such that the boundary of every possible subdivision into $A$ and $B$ is crossed by a single virtual bond. In that case, the TNS is essentially 1D, and we can perform TEBD as in 1D. But such an approach cannot directly encode the correct entanglement distribution for the locally-entangled states of the underlying qubit lattice.

For concreteness we focus on the center bond again. To determine a SVD (4.4) across this bond, we now know that this boils down to determining an appropriate set of isometric tensors. In this case, we also need isometric tensors $A^{\sigma_j}_{j,abcd}$ that collapse in both spatial directions upon contraction with the conjugate isoTNS, e.g.

$$\sum_{\sigma_j ab} A^{\sigma_j}_{2,abcd} A^{\dagger\sigma_j}_{j,abef} = I_{ce} I_{df}. \tag{4.11}$$

As an example, in Fig. 4.3b we show a 2D isoTNS that encodes orthonormal bases for the middle bond. Now we can apply a gate to this bond and afterwards optimally truncate the entanglement between $A$ and $B$. We can do this for all the bonds shown in red, for which we only need to use the bond gauge freedom

---

[2] In principle, we can implement a non-unitary multi-product formula with a unitary circuit, if we extend the system with auxillary qubits. This extended system can then be represented as a tensor network, and e.g. the quantum algorithm from Ref. [180] could be implemented with TEBD.

to move the singular values. Hence we can truncate the entanglement across all possible choices of rectangles for $A$ and $B$. However, in practice we want to evolve all pairs of nn qubits, such that we have to apply gates across all bonds. Since we want to optimally truncate the entanglement after each gate application, this means that we also have to move the singular values to the black bonds in Fig. 4.3b, which is currently situated in an orthonormal basis. For example, to evolve the top row of bonds, we have to move the red row upwards, flipping the vertical bonds that connect the rows. This can be done with the gauge freedom that exists between entire columns, inserting isometries that act on many virtual spaces. Essentially, we want to perform a QR decomposition for an entire row.

To maintain the locality of the tensor network, we want to implement this global operation with a sequence of local operations. Unlike in 1D, the bond gauge freedom no longer suffices. To see this, we can use it to move the singular values in Fig. 4.3b to one of the black bonds that connects with the red string. This creates a loop of isometries, spoiling the orthonormal basis as it can no longer be reduced to identities upon contraction with the conjugate. Hence, we have to move the row while taking care not to create loops, which highlights a fundamenental difference between 1D and higher-D isoTNS. To achieve this we can utilize the gauge freedom between a tensor and two neighboring tensors that are spatially orthogonal. In 2D the required gauge transformation is achieved by splitting a tensor into a triangle [189], while in 3D we can do it with the tetrahedron splitting introduced in Article 4 of this thesis [165]. This lets us simultaneously flip two orthogonal arrows in Fig. 4.3b. Now we can move the singular values to every bond, giving rise to TEBD$^2$, which is the generalization of TEBD to 2D isoTNS.

## 4.7 Algorithmic decoherence

It appears that we are forced to utilize a gauge freedom that is not native to the tensor network layout, coupling three tensors simultaneously, thereby potentially generating multipartite entanglement. As a result, the bond dimensions can grow even without applying a gate, analogous to the situation for summing tensor networks that we encountered in Sec. 4.5. If we maintain a constant threshold $X$ then we can lose information just by moving the singular values across all bonds, such that applying a circuit is accompanied by algorithmic decoherence.

This emphasizes the need for clever algorithms, where circuits are applied with minimal decoherence. It highlights a potential application of non-iterative approximations, as discussed in Sec. 3.10, where we try to pack as much information into each gate as possible. This indirectly minimizes the overall decoherence by minimizing the amount of costly isometry manipulations. If the circuit approximation error is less than the overall error due to decoherence, we do not need to look for a high-accuracy approximation.

With a tree TNS, e.g. the 1D TNS from Fig. 4.1, it does not matter if the time evolution is spread out over many gates that are close to the identity, as is the case for accurate Lie-Trotter-Suzuki circuits. This is because truncation only occurs after applying a gate. Because of this, for fixed $X$ we can always increase the approximation accuracy by increasing the segmentation, whereas with higher-dimensional isoTNS there is a maximum in the accuracy at a finite segmentation [165, 189]. In the limit of infinite segementation the error becomes maximal, since near-identity gates effectively only supply decoherence.

## 4.8 Long-range couplings

It should be noted that some loops appear impossible to resolve. For example, already in 1D with PBC we can have a loop across the entire system. In fact, here the isometric construction does not work at all, since a system-wide loop is inevitable if we want to be able to put the singular values on every bond. A solution is to remove one bond and couple the corresponding pair of tensors with a system-wide coupling. In particular, we can use SWAP gates to move the two qubit spaces to a pair of tensors that share a bond, then apply a gate, and then move the spaces back to their original tensors [190]. But this in some sense spoils the locality of the tensor network, which can no longer be constructed such that it simultaneously encodes the natural pattern of correlations and allows for the singular values to be moved across every bond.

A similar situation occurs when we want to simulate a Hamiltonian with couplings that are not native to the tensor network, i.e. which couple tensors that are not directly connected. Here we also have to use SWAP gates, and if this coupling is equally strong to the native couplings, then again the TNS does not encode the proper hierarchy of correlations. As a result, in the extreme case, to encode entanglement on all length scales the bond dimensions need to be of order $X \sim O(2^{N/2})$. In a practical setting, it is not a priori clear to which extent we should reproduce the Hamiltonian couplings with the tensor network connectivity.

For instance, the states and operators of a qubic 3D qubit system can be represented either as a 3D tensor network, with each tensor having at most six virtual spaces, or e.g. as a snake following a 3D path, with each tensor having only two virtual spaces. TEBD is cheaper for the snake at fixed $X$, but the local couplings of the underlying qubit system appear as long-range couplings of the snake, such that short-range entanglement of the qubit system becomes long-range entanglement for the snake. As such, the required bond dimension $X$ is larger than for a cubic tensor network. It is generally unclear which approach allows one to reach the highest accuracy, which depends on the specific hardware and software used to perform the calculations. Therefore a particular system is often simulated with different tensor networks [191].

## 4.9 Non-isometric tensor networks

Alternatively, we can use non-orthonormal bases, in which case we have to contract the entire TN to perform entanglement truncation, which is generally inefficient [192]. To combat this, an approximate *environment* is constructed for the bond of interest [167, 193]. For example, if we consider the 2D TNS from Sec. 4.6, then after contracting it with its conjugate we can interpret the bulk columns of the tensor network as TNOs, and the first and final columns as TNSs. Then the contraction of the environment boils down to applying 1D isoTNOs to a 1D isoTNS, truncating the entanglement whenever necessary to maintain a constant threshold $X$ [167]. It can be naturally extended to infinite TNS [194].[3] It has significantly higher cost than TEBD$^2$ at fixed $X$, and also has a worse asymptotic cost [193].

Analogous to the algorithmic decoherence of the 2D isoTNS, there is an additional error due to the entanglement truncation of the environment. It is unclear how these compare in general, since their sources are so different. With TEBD$^2$, the error instead stems from the entanglement truncations that are necessary to manipulate manifestly orthonormal bases with constant $X$. Similarly, when we want to calculate the expectation value of an observable at a particular tensor, then in the isoTNS representation

---

[3] This is in contrast with QMC, where thermodynamic estimates instead have to be obtained through finite-size scaling.

we can do this exactly with small constant cost. However, if afterwards we want to calculate it for a different tensor, then we incur an error if it lies outside of the red lines in Fig. 4.3b. In contrast, for a general 2D TNS, the expectation value is always calculated approximately, because the environment always has to be calculated.

As of yet it is unclear how the isometric and general approaches compare for e.g. the real- or imaginary-time evolution with various types of local Hamiltonians. One important aspect would be to consider the maximum performance that can be achieved with a feasible amount of computational resources. Then it would also be interesting to consider how the performance scales with the amount of resources, which would determine the approach that is likely to be more useful in the future. Of course, such a comparison is only legitimate if the employed algorithms are equally optimal in some sense, but this is difficult to estimate. Since TEBD for higher-dimensional isoTNS has acquired popularity only recently, and since its machinery seems to be in an early stage of developement, it seems plausible that it will improve. In contrast, higher-dimensional TNS have been around for longer, and could therefore be expected to undergo less algorithmic advances.

## 4.10 Commentary on Article 4

**Three-dimensional isometric tensor networks**
Phys. Rev. Research 3, 023236 (2021)
*M. S. J. Tepaske*, *D. J. Luitz*

In Article 4 of this cumulative thesis [165], shown in App. D, the author has introduced isoTNS in 3D and generalized the TEBD algorithm. This was benchmarked with the imaginary time-evolution of the 3D transverse-field Ising model for a ground state search. Here a second-order Lie-Trotter-Suzuki circuit was utilized for the state propagation, using a segmentation that roughly minimizes the overall decoherence. We determined an approximate ground state for a range of field strengths and calculated some equilibrium properties of the state in order to benchmark it against numerically exact results. The computational cost of our time-evolution algorithm has the best scaling in bond dimension of any known tensor network algorithm.

It is likely that the higher-dimensional versions of TEBD could benefit from the circuit compression of Article 3 [108], discussed in Sec. 3.11, to reduce the effect of the artificial decoherence. By packing more opEE into each gate than the Lie-Trotter-Suzuki circuits, we can achieve higher accuracy with less gates, which is likely to reduce the overall decoherence. With less information being lost due to the specifics of the algorithm, this could bring us closer to the inherent expressibility of isoTNS and provide access to states that were previously intractable. The potential gain from circuit compression could be even bigger if we can determine a compressed circuit whose error is at least equal to the decoherence error. If this error is not too large, say $\epsilon \sim \mathcal{O}(10^{-2})$, then we can often determine shallow approximations that outperform the Lie-Trotter-Suzuki circuits by orders of magnitude. Of course, for this we have to tackle the circuit optimization problem from Article 3.

IsoTNS also have a natural interpretation in terms of a quantum Markov chain, which can be considered as the initial representation of 1D isoTNS [195]. This representation was already used to formulate *holographic* quantum algorithms, where the isometry structure of a 1D isoTNS is interpreted as a time direction along which qubits travel [196–200]. The virtual spaces are then represented with $\mathcal{O}(\log_2(X))$ ancilla qubits, and the isometric tensors are converted into unitary gates, yielding a quantum circuit. In this way the isoTNS can be sequentially generated, which can be used to perform $D + 1$-dimensional time evolution on a $D$-dimensional quantum computer. When $X > 2$, the gates act on multiple qubits at once, so they first need to be decomposed into nn two-qubit gates to be straightforwardly executed on a quantum computer like that of IBM [164]. Here the compression machinery of Article 3 could again be of great use. The sequential construction can be straightforwardly generalized to higher-dimensional isoTNS [201–203]. Unlike the classical representation of isoTNS, the orthogonality center is now fixed, with its only remaining function to signal the end of the circuit.

The isoTNS representation was only recently brought to attention in 2D, where a TEBD$^2$ algorithm was introduced that has the same computational complexity as TEBD in 1D [189]. In Article 4 we generalize this to TEBD$^3$ for 3D, which is significantly more expensive on account of the tetrahedron splitting [165]. In Ref. [204] a DMRG algorithm was introduced for 2D isoTNS, and in the same article they used 2D isoTNS for real-time evolution. Moreover, in Ref. [205] it was shown that string-net liquids, which are exotic qubit states, can be exactly represented as isoTNS. All these studies focused on finite and isolated qubit systems, but they were subsequently generalized to fermions [206], thermal systems [207], and systems that are infinite in one spatial dimension [208].

The author has conceptualized the contents of the paper, performed the numerical simulations, and wrote the article. David J. Luitz has provided valuable input during all stages.

CHAPTER 5

# Quantum computing

A fundamental solution to the classical intractablity of quantum time evolution is to not construct the Hilbert space on a classical computer, where it has to be done in terms of classical bits and hence with exponential complexity, but to instead do it on a *quantum computer*. To this end we have to replace the classical bits with small quantum systems, in the simplest case qubits. By coherently coupling $N$ qubits we then get a many-body Hilbert space with dimension $2^N$. If we can create and control such a system with perfect accuracy and coherence, we have access to the full Hilbert space with only $N$ constituents, compared to $\mathcal{O}(2^N)$ classical bits. By performing unitary operations on the qubits, e.g. coupling a pair of qubits with coherent laser photons, we can in principle perform arbitrary quantum dynamics [209, 210].

## 5.1 Early days of quantum computing

Many flavors of quantum information processing were already present in the work of Benioff, who built a *universal classical computer* (i.e. a reversible Turing machine [211]) from quantum components [212–214]. Feynman provided an intuitive picture of how quantum coherence could be further harnessed to achieve a *universal quantum computer* [209]. To him it was already clear that, for the simulation of a general $N$-qubit state, a universal classical computer would require $\mathcal{O}(2^N)$ memory and runtime, and that this cannot be avoided due to the sign problem. The natural solution would be to use a computer that behaves entirely according to quantum mechanics, consisting of qubits that can only be coupled in specific ways. This would in principle allow for an exponential improvement in computational complexity over the universal classical computer.

In order to simulate systems that are not already such a qubit system, they first need to be mapped onto one. In that case we require an additional pre-processing step, where the targeted time-evolution algorithm is decomposed into the elementary qubit operations. This shifts the burden onto the quantum algorithms, and allows for hardware to be specifically designed and optimized for a particular type of qubit system. In principle, for it to be universal it should be able to implement arbitrary many-qubit unitaries and the time evolution of arbitrary time-dependent Hamiltonians.

Deutsch formalized the universal quantum computer, essentially generalizing the reversible Turing machine to be entirely quantum mechanical, additionally allowing for logical computations to be carried out in superposition [210]. Because the universal classical computer is a subset of the universal quantum computer, it allows for precise comparison between their computational power. Providing a natural framework to study quantum complexity [215], it was used to determine many instances of quantum superiority, where exponential improvements on asymptotic scalings were rigorously proven [210, 216, 217].

## 5.2 Quantum circuits

Quantum computing became more accessible when it was realized that the circuit language of Fig. 3.1 provides an alternate framework for the universal quantum computer [218–220]. This is analogous to how any universal classical computer can be represented as a Boolean circuit acting on a collection of bits [221]. The diagrammatics of such *quantum circuits* allows for easy design of quantum algorithms, providing an intuitive picture of the quantum simulation on multiple levels. It abstracts the physical implementation of the qubit couplings into gates, just like in Fig. 3.1. Then algorithms can be formulated in terms of general multi-qubit unitaries, or directly in terms of the elementary operations that can be implemented on specific hardware. Here we are again concerned with constructing a global unitary from local evolutions, providing a direct connection with the circuits from Sec. 3.5.

Moreover, quantum circuits provide an intuitive framework for quantum complexity theory [220, 222]. They were also extended to the arena of open quantum systems [100], but here they lose practicality because actual quantum devices can only realize an ensemble by amassing pure state trajectories. Nonetheless, it is a useful tool for classical pre- and post-processing, as evidenced by Article 5 on quantum error mitigation [109], which we will discuss in Sec. 5.10.

An immediate question arises about the universality of circuits with a restricted set of gates: What should this set contain in order to be able to generate arbitrary $N$-qubit unitaries? For reversible Boolean circuits, it has to contain at least a three-bit logic gate in order to be able to construct arbitrary Boolean functions [221]. For quantum circuits we need two-qubit gates at most [223], with reversibility being a natural consequence of quantum mechanics since a time-evolution operator $e^{-itH}$ always has inverse $e^{itH}$. In fact, a simple class of two-qubit gates is already universal [224], as well as any programmable multi-qubit gate [225, 226]. A powerful result is that a single entangling two-qubit gate together with general one-qubit unitaries forms a universal gateset [227]. An example of such a two-qubit gate is the *quantum logical* controlled NOT (CNOT) operation, which flips the left qubit if the right qubit is spin down and otherwise does nothing.

More generally, quantum circuits built from simple universal gatesets can be used to approximate any unitary up to a desired precision [228]. As with the circuits from Ch. 3, allowing for approximation is an important aspect, since imperfect control and decoherence inevitably lead to errors. Therefore, we can maintain an approximation accuracy on the order of the decoherence error, to maximize the use of the noisy computational resources. In the same spirit, we can efficiently decompose general one-qubit gates into a discrete gateset, according to the Solovay-Kitaev theorem. It states that a general one-qubit gate can be decomposed up to precision $\epsilon$ into a sequence of $\mathcal{O}(\mathrm{poly}(\log(1/\epsilon)))$ discrete one-qubit gates that are pulled from a finite gateset [229]. In practice, it implies high accuracy with short sequences, while allowing for a decomposition error. Altogether, these results imply that we can in principle approximate an arbitrary unitary to arbitrary precision with a circuit that is built from a discrete gateset that contains the CNOT and a finite amount of one-qubit unitaries.

## 5.3 Experimental implementation

A quantum circuit implements a sequence of local time evolutions, which have to be physically implemented in a quantum computer. This is often done by coherently coupling qubits through externally-controlled photons and fields, which can realize various time-dependent Hamiltonians for the qubits. These are then tuned to ensure that the integrated time-evolution operator implements the required unitary. Currently, there are multiple quantum computing architectures that are already feasible and which appear as potential candidates for the scalable platforms of the future.

So far we have discussed the *digital* quantum computer, where universality is key. Alternatively, there are experimental setups where the time-evolution operator of particular classes of local quantum many-body Hamiltonians can be directly realized. The implementable unitaries of such an *analog* quantum computer in principle form a subset of the implementable unitaries of an equally-sized digital quantum computer. Nonetheless, their relative simplicity allows for more straightforward scalability.

For example, optical lattices of ultracold atoms can be used to simulate common quantum many-body Hamiltonians [230–235]. Qubits are often encoded in neutral Rydberg atoms, and on account of the Rydberg blockade effect these interact according to a simple spin Hamiltonian [236–239]. Analog quantum computers consisting of trapped Rydberg atoms coupled by external fields and photons have been used to perform the largest quantum time evolution experiments to date [240–247]. Similarly, trapped ions can be used to encode qubits, and the Coulomb interaction can be engineered via external control to realize tuneable spin Hamiltonians [248]. This has also allowed for many large-scale quantum time evolution studies [245, 249–254].

These atom or ion-based systems can moreover be used for the purpose of digital quantum computing. For example, the Rydberg blockade mechanism can be used to construct a two-qubit CNOT [255, 256]. With this a digital quantum computer can be constructed [257–259]. Analogously, logic gates can be implemented on trapped ion platforms [260, 261]. This has also been used to construct digital quantum computers [262, 263] that have already been employed for various simulation studies [200, 264–272]. Less explored options include platforms based on nuclear spins that can be controlled through nuclear magnetic resonance [273], silicon spins that can be manipulated coherently at relatively high temperatures [274, 275], and photonic systems where polarization is used to encode qubits [276, 277].

Perhaps the furthest developed and most popular digital computing platform consists of locally-coupled superconducting qubits [278–280]. One such qubit is essentially a non-harmonic oscillator for which the two lowest energy levels can be (imperfectly) isolated. They are experimentally realized as electrical circuits that contain superconducting Josephson junctions [281], for which the energy spacing lies in the microwave regime. Neighboring qubits are coupled through microwave resonators [282, 283]. This allows for an off-resonant driving that evolves the pair with an effective XZ Hamiltonian, i.e. the corresponding time-evolution operator is

$$U_{XZ}(t) = e^{-it\sigma_0^x \otimes \sigma_1^z}, \tag{5.1}$$

where $\sigma_j^{I,x,y,z}$ is the Pauli algebra of the $j$th qubit. We can then use (5.1) as the elementary two-qubit gate in the variational circuits from Sec. 3.3, which can then be translated to a schedule of microwave pulses with as little overhead as possible. This digital quantum computing platform has arguably been used most extensively, with many studies on quantum many-body simulation [162, 284–296]. Recently, such machines have become large and accurate to the extent that it is no longer clear if their computations can be simulated classically with a feasible amount of computing resources [178, 191, 296, 297].

By focusing on a restricted gateset, much effort can go toward reducing the decoherence of a specific process, while it also simplifies quantum error mitigation. For $t = \pi/4$ the gate (5.1) becomes the CNOT gate, which due to decoherence and imperfect control currently has $\mathcal{O}(1\%)$ median error on a standard superconducting device such as the IBM_ALGIERS computer with 27 qubits [164]. One-qubit gate errors are an order of magnitude lower, making the two-qubit gates the main source of decoherence during a quantum algorithm. There is also a transient error due to the coupling of the qubits with the environment, making the brickwall circuit architecture from Sec. 3.3 favorable because it minimizes the idling time. Finally, there is also an error in the measurement process, which is of the same order as the CNOT error. All these factors can be used to estimate an effective error per two-qubit gate for a corresponding brickwall circuit, which can then be used to determine the *quantum volume* of a device [298]. This is a figure of merit for the amount of qubits times the amount of brickwall layers that can be executed before reaching an error threshold. In practice it serves as a quality factor for digital quantum computers.

Trapped atom or ion platforms often have a two-qubit gate error that is an order of magnitude lower than for superconducting platforms [299]. Even more, they can be directly applied to qubits that are far apart, which could be beneficial when simulating long-range interacting systems. Recently, it was shown that these trapped systems can be configured into any desired 3D configuration with relative ease [246]. This could eventually give rise to e.g. analog quantum computers that can directly simulate the intricate 3D geometries that often appear in experiments [42, 300–302].

## 5.4 Efficient circuits

While the results from Sec. 5.2 display the universality of quantum circuits, for practical purposes we look for efficient circuits, which implement a particular unitary up to precision $\epsilon$ using $\mathcal{O}(\mathrm{poly}(N, t, \log(1/\epsilon)))$ nn one- and two-qubit gates $N_g$. Here $N$ denotes the amount of qubits that the circuit acts on, and $t$ is the evolution time of the targeted unitary (or more generally a parameter that controls its complexity). This is the same situation as we encountered in Sec. 3.5, where we also looked for circuits with favorable cost scaling. Again, the polylogarithmic dependency on $1/\epsilon$ is optimal, representing a simulation overhead that arises even when e.g. a Turing machine simulates another [126].

It turns out that an arbitrary unitary cannot be implemented efficiently [228]. If we can determine a unitary that can be implemented efficiently on a quantum computer but not on a classical computer, we can in principle use it to perform classically intractable calculations. However, analogously to the higher-dimensional isometric tensor networks from Sec. 4.6, each gate application is in reality accompanied by decoherence, which is now physical instead of algorithmic. If this can be overcome, then the circuits from Sec. 3.5 provide a path to high-accuracy time evolution with $N_g \sim \mathcal{O}(N)$ for all times. In Ref. [226] the first-order Lie-Trotter-Suzuki circuit (3.4) was used to formulate the first efficient quantum simulation algorithm for local Hamiltonians.

For the time-evolution operator of local Hamiltonians, a decomposition into CNOTs can be obtained by starting from a Lie-Trotter-Suzuki circuit, and then decomposing every gate into nn CNOTs and one-qubit unitaries [132]. Similarly, in the universality studies [223–225, 227] a hierarchy of decompositions is used to decompose an arbitrary unitary. First, the unitary is analytically decomposed into large multi-controlled gates, which are then decomposed into Toffoli gates, and finally into CNOTs. This sequential scheme was afterwards improved, reducing overheads and improving bounds [303–308]. Clearly, such a hierarchical procedure is prone to global redundancies, with different parts of the circuit being synthesized independently. One approach to resolving redundancies and inefficiencies, especially with a particular hardware implementation in mind, is to utilize circuit identities [309–314]. This can drastically reduce the gate count in some instances [315].

There are also many Hamiltonians of interest that do not have a tensor product structure, but which are sparse. These feature for example in adiabatic quantum computation [316, 317] and various quantum algorithms [318–321]. In that case, one option is to decompose $H$ into a sum of simple terms [154, 317, 322], and use the same techniques as for local systems.

## 5.5 Effects of decoherence

Even if a quantum algorithm can in principle yield an exponential speedup over the best classical algorithm, this situation might change drastically in the presence of decoherence. In particular, in the absence of error correction, the state will mix increasingly with the enviroment as noisy gates are applied to it, until it ends up in a completely mixed state. In any case, to minimize the detrimental effects of decoherence we can try to pack a quantum algorithm into as little gates as possible. This was the prime motivation behind Article 3 of this thesis [108] and the preprint [104], which we discussed in Sec. 3.11.

To achieve quantum supremacy in this age of noisy quantum computing requires us to find classically intractable problems that can be solved with high accuracy using a relatively small amount of gates $N_g$. Currently, IBM devices have quantum volumes that are on the order of the qubit count, e.g. IBM_ALGIERS has a quantum volume of 128 at 27 qubits [164], such that only a few brickwall layers of CNOTs can be executed before complete mixing is expected to occur. At this point, the asymptotic scaling of gates is not important, since it pertains to the behavior of the large and clean devices of the future. Instead we are currently restricted to the regime of small and noisy computers without error correction. With this in mind, the most straightforward way of generating classically intractable states is with random circuits, to generate as much entanglement as possible for a fixed number of gates. A more meaningful goal is to perform a classically intractable calculation that approximates a meaningful unitary, such as a time evolution operator [135]. In any case, the classical verification of potential showcasings of quantum supremacy is crucial [191, 296, 297].

For locally-connected quantum computers such as the superconducting platforms, where the elementary gates act on qubits that are physically neighboring, the simplest circuit architecture is that of a brickwall as in Fig. 3.1. In that case, gates do not need to be supplied with SWAP gates before being executed on the quantum computer. We discussed this in detail in Article 3 [108], which we covered in Sec. 3.11. Then the gate count of a circuit is a direct indication of the potential influence of decoherence.

If the total error incurred per gate is below some threshold, we can in principle make the circuit execution fault-tolerant through error correction [323–325]. Here, redundancy is used to detect and correct errors, such that we could in principle execute arbitrarily large circuits with perfect precision. Then large circuits such as those for the quantum Fourier transform can be coherently executed, opening the door to e.g. Shor's algorithm for integer factoring [326]. This would allow for unmistakable demonstrations of quantum supremacy. It can be considered as an ultimate goal of quantum computing, but as of yet the experimental aspects appear out of reach [327]. For now, we instead rely on classical pre- and post-processing to make the most out of what we have, prompting research such as that of Article 3 [108], Article 5 [109], and the preprint [104].

## 5.6 Alternatives to product formulas

The simplest approach to time evolution on a quantum computer is conceptually analogous to the TEBD algorithm for tensor networks from Sec. 4.2. Namely, we store the gates in a classical memory and execute them sequentially on the quantum computer. However, we can now also utilize ancilla qubits to coherently implement non-unitary operations. This allows for many alternatives to sequential circuits, often with improved asymptotic scaling. For example, we can execute the whole circuit coherently instead of sequentially [154, 218, 317, 322].

For this we need to introduce ancilla qubits, enough to encode the gates into a state. For example, to select gate $G_j$ from a circuit $C$, and apply it to a state $|\psi\rangle$, we can use

$$\text{select}(C)\,|j\rangle\,|\psi\rangle = |\phi_j\rangle\,U_j\,|\psi\rangle\,, \tag{5.2}$$

where $|\phi_j\rangle$ is a garbage state that is there solely to make the select operation unitary (i.e. reversible). The state $|j\rangle$ picks a gate at a particular bond and should therefore have at least $N_g$ different basis states, i.e. we need $\mathcal{O}(\log_2(N_g))$ ancilla qubits. In the current age of quantum computing, with experimental devices containing $\mathcal{O}(10^2)$ noisy qubits, this ancilla overhead is intolerable, especially since these algorithms also come with a large constant gate overhead [128, 135, 317]. Essentially, they are tailored to the quantum computers of a future generation, where such overheads become insignificant relative to the asymptotic cost.

The select($C$) operation also allows for the coherent implementation of a linear combination of unitaries [328]. As we mentioned in Sec. 3.10, this can be used to formulate time-evolution algorithms with a gate count that scales optimally as $\mathcal{O}(\text{poly}(\log(1/\epsilon)))$, i.e. exponentially better than a single circuit. Many such quantum algorithms were developed, for example using a discrete fractional query model [126], a quantum walk (which saturates the no-fast-forwarding bound even for non-local Hamiltonians) [319, 329, 330], multiproduct formulas [180, 331], a direct truncation of the Taylor series of $e^{-itH}$ [128], and representing the Lie-Trotter-Suzuki remainder as a linear combination of unitaries [127]. The block encoding that a linear combination of unitaries realizes can also be implemented differently, namely by encoding the eigenvalues of the Hamiltonian into ancilla qubits, which are then rotated to induce time evolution [332, 333]. For local Hamiltonians we can use one of these algorithms with optimal scaling in $1/\epsilon$ to construct a quantum algorithm that saturates all optimality bounds, by relying on the Lieb-Robinson bound [158]. All mentioned algorithms require at least $\mathcal{O}(N)$ ancilla qubits. In Ref. [135] the gate counts were estimated for a few of the mentioned algorithms, in the context of reaching a particular evolution time for a simple spin model with a particular precision. All the given estimates contain millions of gates, still out of range for modern devices, and even then their emperical bounds predict that the product formulas outperform the rest while requiring no ancilla qubits.

Of course, all these algorithms can be implemented classically as well, e.g. using a single tensor network for the system and ancilla qubits. This can be useful in order to verify and analyze various aspects of an algorithm.

## 5.7 Randomization

If we only care about the expectation values that a circuit produces when acting on a state, we can improve its accuracy by introducing randomness, so that we work with an ensemble of circuits instead of a single circuit [334, 335]. For example, if we take a product formula, instead of always applying one specific gate at a particular point in the algorithm, we now sample the gate from a probability distribution [336–338]. This is done by replacing each gate with a sum of gates, with the coefficients corresponding to probabilities [109]. Randomization can double the approximation order of Lie-Trotter-Suzuki circuits [339]. It also allows for the stochastic implementation of a linear combination of unitaries, requiring only a single ancilla and shallow circuits compared to a coherent implementation [127, 340]. This could benefit the mentioned time-evolution algorithms that rely on this subroutine, by providing a more hardware-friendly alternative that could be viable in the current era of quantum computing.

The use of such *randomization* is natural when looking for circuit approximations in the presence of decoherence. This is because the decoherence already introduces a classical probability distribution, effectively doubling the Hilbert space to accomodate for density matrices. Such an ensemble of states can then be directly propagated by an ensemble of circuits, called a *supercircuit* [100], and we can extend the variational compilation framework of Sec. 3.3 to determine the optimal supercircuit approximation of e.g. a noisy time-evolution operator. Here each constituent *supergate*, i.e. an ensemble of gates, has to be parameterized. An obvious choice is to make it into a probability distribution over various gates, with the probabilities also functioning as variational parameters. We utilized this in Article 5 [109], which we discuss in Sec. 5.10.

## 5.8 Variational quantum compiling

When performed in the context of quantum computing, the variational compiling approach from Sec. 3.3 is called *variational quantum compiling*, and it has been at the core of many studies on quantum time evolution [104, 108, 160–163, 341–353]. This approach has also been used to represent states of interest as a circuit applied to some initial state [115, 354–362]. In most of these variational compiling studies the computational load of the circuit optimization is intended to be placed onto a digital quantum computer [160–163, 341–349]. With the current level of noise this program still suffers from many ailments that are absent with a classical optimization algorithm [342, 363]. Hence there are also variational compiling approaches where the optimization is geared to a classical computer [104, 108, 350–353]. Then the quantum computer is only needed to execute the final optimized circuit, in order to simulate the targeted dynamics. As we will discuss below, this allows for the generation of classically intractable states, even though the optimization is classically tractable. Naturally, here a restricted gateset is often used to taylor the variational compiler to a specific quantum computing platform.

The classical optimization problem at the core of variational compiling has only recently become feasible for more than three qubits, with the rise of supercomputers and efficient optimization algorithms based on gradient descent [186]. In the first place, this is because the circuit parameters form a high-dimensional space, and determining the minimimum distance (3.3) boils down to minizing a nonlinear function over this space. Then, the distance needs to be calculated many times, and this can be costly when done with high accuracy for more than a few qubits.

Determining the minimum distance through a simultaneous global scan of all circuit parameters becomes infeasible already for a modest amount of parameters and qubits, so we resort to a local scheme. Here the circuit is initialized with some parameters, and these are then continuously updated based on local information about the distance function. To first order only the gradient is used, from which the *gradient descent* algorithm derives its name. Now we can get stuck in local minima of the distance function, which is a well-known problem of variational compiling and the optimization of complicated non-linear functions in general [354].

To classically optimize a circuit that acts on many qubits, we can represent the circuit and the targeted unitary as truncated isoTNOs, as explained in Sec. 4.2. But this will inevitably fail for large times with a generic Hamiltonian, in which case the optimization is limited to a few qubits. Crucially, we can perform the optimization only for a few qubits, or only at small times, and then "extrapolate" the circuit. For example, we can use tensor networks to perform the distance calculation at relatively small times on a large lattice, with the bond dimension chosen such that we are just able to perform the calculations without significant information loss. Then we can simply stack the circuit twice to reach double the time, which we can no longer verify on a classical computer. Alternatively, if we have a quantum computer with PBC, we can use translationally-invariant circuits, which are straightforwardly used at any system size. Then we can perform the optimization at e.g. $N = 8$ and use the optimized circuits on a larger quantum computer. In Ref. [104] we have found that this extrapolation works extremely well for the models considered. Either way, we can determine highly-accurate circuits with relatively small gate counts via a classically-tractable optimization procedure. This is at the core of the variational quantum compiling approach that we have taken in Article 3, Article 5, and the preprint [104].

## 5.9 Mitigating errors

After we have determined the optimal circuit parameters to approximate the targeted unitary, we can now use them as instructions for the quantum computer. By performing many shots for a couple of initial states, and measuring the output states, we can check how close the experiment is to the theoretical prediction. Naturally, the decoherence affects this correspondence, with deep circuits yielding a complete loss of information. Interestingly, for some cases it is known that the optimal circuit parameters remain unchanged when including noise [161, 343].

To combat the information loss, we can perform *quantum error mitigation*, which is a classical post-processing step. With *probabilistic error cancellation* (PEC) we can effectively cancel the errors due to the decoherence [364, 365]. For this we first create many new realizations of the circuit, by extending it with random gates that are distributed according to a particular *mitigating ensemble*. Then we can determine a mitigated observable by calculating the observable for each of the random circuits, performing many shots for every circuit, and then averaging over the results. If done exactly, the amount of required samples grows exponentially with $N$, which is why in Article 5 we aim to reduce the amount of gates in the mitigating ensemble through variational compiling.

Alternatively, we can use *zero-noise extrapolation*, where the decoherence is varied to obtain data at various noise levels [125, 191, 366]. This is then extrapolated to zero noise in an attempt to mitigate noise. For a brickwall circuit of CNOTs, the overall decoherence due to two-qubit gates can be varied by inserting products of CNOTs into the circuit, with a product of two CNOTs being equal to the identity.

## 5.10 Commentary on Article 5

**Compressed quantum error mitigation**
Phys. Rev. B 107, L201114 (2023)
*M. S. J. Tepaske*, *D. J. Luitz*

In Article 5 of this cumulative thesis [109], shown in App. E, we use variational quantum compiling to determine shallow mitigating ensembles for PEC on the level of an entire circuit. Concretely, we look for a supercircuit that (partially) reverts the decoherence of a circuit, where the decoherence stems from the accumulation of two-qubit decoherence. In practice, after such a *denoiser* has been determined, the circuits that make up its ensemble have to be sampled and applied after every sample of the noisy circuit. By averaging over many samples, the errors are mitigated up to the capacity of the denoiser. Because we also have to average over many quantum trajectories, the cost of PEC is high. This emphasizes the utility of a PEC approach where the size of the mitigating ensemble can be tuned, in contrast to the traditional approach from Refs. [364, 365]. By keeping the ensemble small we can enforce practicality, since it requires less sampling. This comes at the cost of imperfect mitigation, i.e. it has an approximation error relative to a perfectly mitigating ensemble, but this can be a practical advantage when we are mitigating the errors of circuits that already have an associated approximation error. For example, in the Article we chose to mitigate second-order Lie-Trotter-Suzuki circuits of the 1D Heisenberg model, so at most we need to look for a denoiser that is equally accurate to the Lie-Trotter-Suzuki circuit. Alternatively, we can look for denoisers that only consist of a couple gates, in order to construct fast partial mitigation procedures.

The denoiser is unphysical, and even if we had access to a quantum computer that realizes density matrices and allows for arbitrary physical maps, it would not be executable. It is the analog of a non-unitary circuit in the supercircuit formalism. It makes little sense to look for a coherent implementation, because we cannot physically realize a supercircuit, so a natural approach is to utilize the randomization from Sec. 5.7 to realize it stochastically. To facilitate this, we parameterize the supergates in a denoiser, which is a supercircuit with a brickwall architecture, as a convex sum of elementary gates and measurements. Every supergate can then be sampled based on its coefficients, and the resulting collection of unitary circuits interspersed with measurements realizes the ensemble of the denoiser. This is similar to many of the methods that we mentioned in Sec. 5.6, where a sum of unitaries has to be realized on a quantum computer.

In contrast to Article 3, here we utilized translationally-invariant circuits. For this we have to restrict the noisy time-evolution operator to be translationally invariant. This is crucial in order to simulate the performance of the denoiser on larger quantum computers. In particular, since we are now working with supercircuits, exact calculations have cost $O(4^N)$. Consequently, the distance minimization is already expensive with $N = 8$ qubits. We can also calculate the distance stochastically, by utilizing typicality, i.e. interpreting the superoperator trace as an average over random density matrices, and replacing the ensemble calculations with averages over pure state trajectories. This way, we only need to perform pure-state calculations, at the cost of needing many samples. Besides allowing us to check the performance of the denoiser at system sizes that are larger than the size for which it was optimized, it also opens the door to the extrapolation procedure of Sec. 5.8.

There are many interesting variations on the problem that we have studied in this Article. For example, instead of compiling a noisy circuit into the product of the noiseless version and a denoiser, we could directly compile it into an ensemble of pure-state operations. In other words, we could directly incorporate the error mitigation into the time evolution, following the approach from Article 3. It could also be worthwhile to consider different denoiser architectures. For example, it could be advantageous to denoise after every couple layers, instead of waiting until the entire circuit has been completed.

The author of this thesis has developed the theory, performed the numerical simulations, and wrote the article. All with input from David J. Luitz.

# Bibliography

[1] I. Newton, *Philosophiae Naturalis Principia Mathematica*, Edmond Halley, Cambridge, 1687.

[2] I. Newton, *Opticks*, William Innys, London, 1704.

[3] C. Huygens, *Treatise on Light*, Pierre van der Aa, Leyden, 1690.

[4] P. E. B. Jourdain, *Maupertuis and the principle of least action*, Monist **22** (1912) 414.

[5] W. Hamilton, *On a General Method in Dynamics*, Philos. Trans. Royal Soc. (1834).

[6] E. Schrödinger, *Quantisierung als Eigenwertproblem*, Ann. Phys. **384** (1926) 361.

[7] E. Schrödinger, *Quantisierung als Eigenwertproblem*, Ann. Phys. **384** (1926) 489.

[8] E. Schrödinger, *Quantisierung als Eigenwertproblem*, Ann. Phys. **385** (1926) 437.

[9] E. Schrödinger, *Quantisierung als Eigenwertproblem*, Ann. Phys. **386** (1926) 109.

[10] E. Schrödinger, *Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinem*, Ann. Phys. **384** (1926) 734.

[11] E. Schrödinger, *Der stetige Übergang von der Mikro- zur Makromechanik*, Sci. Nat. **14** (1926) 664.

[12] L. de Broglie, *Recherches sur la Théorie des Quanta*, Ann. Phys. **10** (1925) 22.

[13] W. Pauli, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, Z. Phys. **33** (1925) 879.

[14] M. Born and J. P., *Zur Quantenmechanik*, Z. Phys. **34** (1925) 858.

[15] M. Born, W. Heisenberg and J. P., *Zur Quantenmechanik. II.*, Z. Phys. **35** (1926) 557.

[16] W. Heisenberg, *Zur Theorie des Ferromagnetismus*, Z. Phys. **49** (1928) 619.

[17] P. Weiss, *L'hypothèse du champ moléculaire et la propriété ferromagnétique*, J. Phys. Theor. Appl. **6** (1907) 661.

[18] W. Lenz, *Beitrag zum Verständnis der magnetischen Erscheinungen in festen Körpern*, Z. Phys. **21** (1920) 613.

[19] E. Ising, *Beitrag zur Theorie des Ferromagnetismus*, Z. Phys. **31** (1925) 253.

[20] L. Onsager, *Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition*, Phys. Rev. **65** (1944) 117.

[21] L. D. Fosdick, *Calculation of Order Parameters in a Binary Alloy by the Monte Carlo Method*, Phys. Rev. **116** (1959) 565.

[22] J. R. Ehrman, L. D. Fosdick and D. C. Handscomb, *Computation of Order Parameters in an Ising Lattice by the Monte Carlo Method*, J. Math. Phys. **1** (1960) 547.

[23] T. Ruijgrok and H. Van der Vlist, *On the Hamiltonian and Lagrangian formulation of classical dynamics for particles with spin*, Phys. A **101** (1980) 571.

[24] G. E. Uhlenbeck and S. Goudsmit, *Ersetzung der Hypothese vom unmechanischen Zwang durch eine Forderung bezüglich des inneren Verhaltens jedes einzelnen Elektrons*, Sci. Nat. **13** (1925) 953.

[25] W. Pauli, *Über den Einfluß der Geschwindigkeitsabhängigkeit der Elektronenmasse auf den Zeemaneffekt*, Z. Phys. **31** (1925) 373.

[26] W. Gerlach and O. Stern, *Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld*, Z. Phys. **9** (1922) 349.

[27] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, 2020.

[28] P. Coleman, *Introduction to Many-Body Physics*, 2015.

[29] H. A. Brown and J. M. Luttinger, *Ferromagnetic and Antiferromagnetic Curie Temperatures*, Phys. Rev. **100** (1955) 685.

[30] H. E. Stanley and T. A. Kaplan, *High-Temperature Expansions-the Classical Heisenberg Model*, Phys. Rev. Lett. **16** (1966) 981.

[31] F. Wegner, *Spin-ordering in a planar classical Heisenberg model*, Z. Phys. **206** (1967) 465.

[32] G. S. Joyce, *Classical Heisenberg Model*, Phys. Rev. **155** (1967) 478.

[33] H. E. Stanley, *Dependence of Critical Properties on Dimensionality of Spins*, Phys. Rev. Lett. **20** (1968) 589.

[34] V. L. Berezinskiĭ, *Destruction of Long-range Order in One-dimensional and Two-dimensional Systems having a Continuous Symmetry Group I. Classical Systems*, J. Exp. Theor. Phys. **32** (1971) 493.

[35] J. M. Kosterlitz and D. J. Thouless, *Ordering, metastability and phase transitions in two-dimensional systems*, J. Phys. C **6** (1973) 1181.

[36] J. M. Kosterlitz, *The critical properties of the two-dimensional xy model*, J. Phys. C **7** (1974) 1046.

[37] S. Miyashita, H. Nishimori, A. Kuroda and M. Suzuki, *Monte Carlo Simulation and Static and Dynamic Critical Behavior of the Plane Rotator Model*, Prog. Theor. Phys. **60** (1978) 1669.

[38] M. Suzuki, S. Miyashita and A. Kuroda, *Monte Carlo Simulation of Quantum Spin Systems. I*, Prog. Theor. Phys. **58** (1977) 1377.

[39] H.-Q. Ding and M. S. Makivi ć, *Kosterlitz-Thouless transition in the two-dimensional quantum XY model*, Phys. Rev. B **42** (1990) 6827.

[40] H.-Q. Ding, *Could in-plane exchange anisotropy induce the observed antiferromagnetic transitions in the undoped high-$T_c$ materials?*, Phys. Rev. Lett. **68** (1992) 1927.

[41] A. Cuccoli, T. Roscilde, V. Tognetti, R. Vaia and P. Verrucchi, *Quantum Monte Carlo study of $S = \frac{1}{2}$ weakly anisotropic antiferromagnets on the square lattice*, Phys. Rev. B **67** (2003) 104414.

[42] D. Opherden et al., *Field-Tunable Berezinskii-Kosterlitz-Thouless Correlations in a Heisenberg Magnet*, Phys. Rev. Lett. **130** (2023) 086704.

[43] D. Opherden et al., *Berezinskii—Kosterlitz—Thouless correlations in copper-based quasi-2D spin systems (Review Article)*, Low Temp. Phys. **49** (2023) 819.

[44] S. Lie, *Theorie der Transformationsgruppen I*, Math. Ann. **16** (1880) 441.

[45] M. Suzuki, *Relationship among Exactly Soluble Models of Critical Phenomena. I*): 2D Ising Model, Dimer Problem and the Generalized XY-Model*, Prog. Theor. Phys. **46** (1971) 1337.

[46] M. Suzuki, *Relationship between d-Dimensional Quantal Spin Systems and (d+1)-Dimensional Ising Systems: Equivalence, Critical Exponents and Systematic Approximants of the Partition Function and Spin Correlations*, Prog. Theor. Phys. **56** (1976) 1454.

[47] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *Equation of State Calculations by Fast Computing Machines*, J. Chem. Phys. **21** (1953) 1087.

[48] G. Pan and Z. Y. Meng, “The sign problem in quantum Monte Carlo simulations”, *Encyclopedia of Condensed Matter Physics (Second Edition)*, ed. by T. Chakraborty, Second Edition, Academic Press, 2024 879.

[49] M. Troyer, M. Imada and K. Ueda, *Critical Exponents of the Quantum Phase Transition in a Planar Antiferromagnet*, J. Phys. Soc. Jpn. **66** (1997) 2957.

[50] M. Suzuki, *Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics*, J. Math. Phys. **26** (1985) 601.

[51] M. Suzuki, *Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations*, Phys. Lett. A **146** (1990) 319.

[52] M. Suzuki, *General theory of fractal path integrals with applications to many-body theories and statistical physics*, J. Math. Phys. **32** (1991) 400.

[53] M. Suzuki, *Hybrid exponential product formulas for unbounded operators with possible applications to Monte Carlo simulations*, Phys. Lett. A **201** (1995) 425.

[54] H. A. Forbert and S. A. Chin, *Fourth-order diffusion Monte Carlo algorithms for solving quantum many-body problems*, Phys. Rev. B **63** (2001) 144518.

[55] M. Suzuki, *Quantum statistical monte carlo methods and applications to spin systems*, J. Stat. Phys. **43** (1986) 883.

[56] B. B. Beard and U.-J. Wiese,
*Simulations of Discrete Quantum Systems in Continuous Euclidean Time*,
Phys. Rev. Lett. **77** (1996) 5130.

[57] N. V. Prokof'ev, B. V. Svistunov and I. S. Tupitsyn, *Exact, complete, and universal continuous-time worldline Monte Carlo approach to the statistics of discrete quantum systems*,
J. Exp. Theor. Phys. **87** (1998) 310.

[58] L. Pollet, K. V. Houcke and S. M. Rombouts,
*Engineering local optimality in quantum Monte Carlo algorithms*,
Journal of Computational Physics **225** (2007) 2249.

[59] E. Gull et al., *Continuous-time Monte Carlo methods for quantum impurity models*,
Rev. Mod. Phys. **83** (2011) 349.

[60] B. Frischmuth, B. Ammon and M. Troyer,
*Susceptibility and low-temperature thermodynamics of spin-½ Heisenberg ladders*,
Phys. Rev. B **54** (1996) R3714.

[61] M. Troyer, H. Kontani and K. Ueda, *Phase Diagram of Depleted Heisenberg Model for Ca*$\mathrm{V}_4\mathrm{O}_9$,
Phys. Rev. Lett. **76** (1996) 3822.

[62] M. Greven, R. J. Birgeneau and U. .-.-J. Wiese,
*Monte Carlo Study of Correlations in Quantum Spin Ladders*, Phys. Rev. Lett. **77** (1996) 1865.

[63] S. Trotzky et al.,
*Suppression of the critical temperature for superfluidity near the Mott transition*,
Nat. Phys. **6** (2010) 998.

[64] E. Burovski, N. Prokof'ev, B. Svistunov and M. Troyer,
*Critical Temperature and Thermodynamics of Attractive Fermions at Unitarity*,
Phys. Rev. Lett. **96** (2006) 160402.

[65] S. Wessel and M. Troyer, *Supersolid Hard-Core Bosons on the Triangular Lattice*,
Phys. Rev. Lett. **95** (2005) 127205.

[66] C. Zhang, B. Capogrosso-Sansone, M. Boninsegni, N. V. Prokof'ev and B. V. Svistunov,
*Superconducting Transition Temperature of the Bose One-Component Plasma*,
Phys. Rev. Lett. **130** (2023) 236001.

[67] D. C. Handscomb, *The Monte Carlo method in quantum statistical mechanics*,
Math. Proc. Cam. Phil. Soc. **58** (1962) 594.

[68] D. C. Handscomb, *A Monte Carlo method applied to the Heisenberg ferromagnet*,
Math. Proc. Cam. Phil. Soc. **60** (1964) 115.

[69] J. W. Lyklema, *Quantum-Statistical Monte Carlo Method for Heisenberg Spins*,
Phys. Rev. Lett. **49** (1982) 88.

[70] S. Chakravarty and D. B. Stein, *Monte Carlo Simulation of Quantum Spin Systems*,
Phys. Rev. Lett. **49** (1982) 582.

[71] D. H. Lee, J. D. Joannopoulos and J. W. Negele,
*Monte Carlo solution of antiferromagnetic quantum Heisenberg spin systems*,
Phys. Rev. B **30** (1984) 1599.

[72] A. W. Sandvik and J. Kurkijärvi, *Quantum Monte Carlo simulation method for spin systems*, Phys. Rev. B **43** (1991) 5950.

[73] A. W. Sandvik, *A generalization of Handscomb's quantum Monte Carlo scheme-application to the 1D Hubbard model*, J. Phys. A **25** (1992) 3667.

[74] O. F. Syljuåsen and A. W. Sandvik, *Quantum Monte Carlo with directed loops*, Phys. Rev. E **66** (2002) 046701.

[75] U. J. Wiese and H. P. Ying, *A determination of the low energy parameters of the 2-d Heisenberg antiferromagnet*, Z. Phys. B **93** (1994) 147.

[76] A. W. Sandvik and D. J. Scalapino, *Order-disorder transition in a two-layer quantum antiferromagnet*, Phys. Rev. Lett. **72** (1994) 2777.

[77] A. W. Sandvik, *Finite-size scaling of the ground-state parameters of the two-dimensional Heisenberg model*, Phys. Rev. B **56** (1997) 11678.

[78] P. Sengupta, A. W. Sandvik and D. K. Campbell, *Bond-order-wave phase and quantum phase transitions in the one-dimensional extended Hubbard model*, Phys. Rev. B **65** (2002) 155113.

[79] A. W. Sandvik, *Evidence for Deconfined Quantum Criticality in a Two-Dimensional Heisenberg Model with Four-Spin Interactions*, Phys. Rev. Lett. **98** (2007) 227202.

[80] J. Lou, A. W. Sandvik and N. Kawashima, *Antiferromagnetic to valence-bond-solid transitions in two-dimensional* $SU(N)$ *Heisenberg models with multispin interactions*, Phys. Rev. B **80** (2009) 180414.

[81] Y. Cui et al., *Proximate deconfined quantum critical point in SrCu2(BO3)2*, Science **380** (2023) 1179.

[82] N. Prokof'ev, B. Svistunov and I. Tupitsyn, *"Worm" algorithm in quantum Monte Carlo simulations*, Phys. Lett. A **238** (1998) 253.

[83] N. D. Mermin and H. Wagner, *Absence of Ferromagnetism or Antiferromagnetism in One- or Two-Dimensional Isotropic Heisenberg Models*, Phys. Rev. Lett. **17** (1966) 1133.

[84] M. E. Fisher, M. N. Barber and D. Jasnow, *Helicity Modulus, Superfluidity, and Scaling in Isotropic Systems*, Phys. Rev. A **8** (1973) 1111.

[85] D. R. Nelson and J. M. Kosterlitz, *Universal Jump in the Superfluid Density of Two-Dimensional Superfluids*, Phys. Rev. Lett. **39** (1977) 1201.

[86] A. W. Sandvik, A. Avella and F. Mancini, "Computational Studies of Quantum Spin Systems", *AIP Conference Proceedings*, AIP, 2010.

[87] R. Gupta et al., *Phase Transition in the* 2DXY *Model*, Phys. Rev. Lett. **61** (1988) 1996.

[88] D. D. Betts, F. C. Salevsky and J. Rogiers, *Vortices in the two-dimensional s=1/2 XY model*, J. Phys. A **14** (1981) 531.

[89] E. Loh, D. J. Scalapino and P. M. Grant, *Monte Carlo studies of the quantum* XY *model in two dimensions*, Phys. Rev. B **31** (1985) 4712.

[90] J. Als-Nielsen, S. T. Bramwell, M. T. Hutchings, G. J. McIntyre and D. Visser, *Neutron scattering investigation of the static critical properties of Rb2CrCl4*, J. Phys. **5** (1993) 7871.

[91] M. Greven et al., *Spin correlations in the 2D Heisenberg antiferromagnet Sr2CuO2Cl2: Neutron scattering, Monte Carlo simulation, and theory*, Phys. Rev. Lett. **72** (1994) 1096.

[92] O. Mustonen et al., *Spin-liquid-like state in a spin-1/2 square-lattice antiferromagnet perovskite induced by d10–d0 cation mixing*, Nat. Commun. **9** (2018) 1085.

[93] D. Opherden et al., *Extremely well isolated two-dimensional spin-$\frac{1}{2}$ antiferromagnetic Heisenberg layers with a small exchange coupling in the molecular-based magnet CuPOF*, Phys. Rev. B **102** (2020) 064431.

[94] A. Cuccoli, T. Roscilde, R. Vaia and P. Verrucchi, *Field-induced* XY *behavior in the* $S = \frac{1}{2}$ *antiferromagnet on the square lattice*, Phys. Rev. B **68** (2003) 060402.

[95] A. Cuccoli, T. Roscilde, R. Vaia and P. Verrucchi, *Detection of XY Behavior in Weakly Anisotropic Quantum Antiferromagnets on the Square Lattice*, Phys. Rev. Lett. **90** (2003) 167205.

[96] J. S. Helton et al., *Spin Dynamics of the Spin-*$1/2$ *Kagome Lattice Antiferromagnet* $\mathrm{ZnCu_3(OH)_6Cl_2}$, Phys. Rev. Lett. **98** (2007) 107204.

[97] Y. Okamoto, M. Nohara, H. Aruga-Katori and H. Takagi, *Spin-Liquid State in the* $S = 1/2$ *Hyperkagome Antiferromagnet* $\mathrm{Na_4Ir_3O_8}$, Phys. Rev. Lett. **99** (2007) 137207.

[98] S. A. Zvyagin et al., *Pressure-tuning the quantum spin Hamiltonian of the triangular lattice antiferromagnet Cs2CuCl4*, Nat. Commun. **10** (2019).

[99] B. Sana et al., *Spin-liquid-like state in a square lattice antiferromagnet*, arXiv:2304.13116 (2023).

[100] D. Aharonov, A. Kitaev and N. Nisan, “Quantum Circuits with Mixed States”, *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC ’98, Dallas, Texas, USA: Association for Computing Machinery, 1998 20, ISBN: 0897919629.

[101] J. Watrous, *The Theory of Quantum Information*, Cambridge University Press, 2018.

[102] J. Eisert, K. Audenaert and M. B. Plenio, *Remarks on entanglement measures and non-local state distinguishability*, J. Phys. A **36** (2003) 5605.

[103] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe and S. Zhu, *Theory of Trotter Error with Commutator Scaling*, Phys. Rev. X **11** (2021) 011020.

[104] M. S. J. Tepaske, D. J. Luitz and D. Hahn, *Optimal compression of constrained quantum time evolution*, arXiv:2311.06347 (2023).

[105] H. F. Trotter, *On the product of semi-groups of operators*, Proc. Am. Math. Soc. **10** (1959) 545.

[106] M. Suzuki, *Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems*, Commun. Math. Phys. **51** (1976) 183.

[107] A. M. Childs and Y. Su, *Nearly Optimal Lattice Simulation by Product Formulas*, Phys. Rev. Lett. **123** (2019).

[108] M. S. J. Tepaske, D. Hahn and D. J. Luitz, *Optimal compression of quantum many-body time evolution operators into brickwall circuits*, SciPost Phys. **14** (2023) 073.

[109] M. S. J. Tepaske and D. J. Luitz, *Compressed quantum error mitigation*, Phys. Rev. B **107** (2023).

[110] D. J. Luitz and Y. Bar Lev, *Information propagation in isolated quantum systems*, Phys. Rev. B **96** (2017) 020406.

[111] J. Maldacena, S. H. Shenker and D. Stanford, *A bound on chaos*, JHEP **2016** (2016) 106.

[112] A. I. Larkin and Y. N. Ovchinnikov, *Quasiclassical Method in the Theory of Superconductivity*, J. Exp. Theor. Phys. **28** (1969) 1200.

[113] E. H. Lieb and D. W. Robinson, *The finite group velocity of quantum spin systems*, Commun. Math. Phys. **28** (1972) 251.

[114] P. W. Claeys and A. Lamacraft, *Maximum velocity quantum circuits*, Phys. Rev. Res. **2** (2020).

[115] S.-H. Lin, R. Dilip, A. G. Green, A. Smith and F. Pollmann, *Real- and Imaginary-Time Evolution with Compressed Quantum Circuits*, PRX Quantum **2** (2021) 010342.

[116] J. E. Campbell, *On a Law of Combination of Operators bearing on the Theory of Continuous Transformation Groups*, Proceedings of the London Mathematical Society **s1-28** (1896) 381.

[117] M. Suzuki, *On the convergence of exponential operators—the Zassenhaus formula, BCH formula and systematic approximants*, Commun. Math. Phys. **57** (1977) 193.

[118] W. Magnus, *On the exponential solution of differential equations for a linear operator*, Commun. Pure. Appl. Math. **7** (1954) 649.

[119] R. M. Wilcox, *Exponential Operators and Parameter Differentiation in Quantum Physics*, J. Math. Phys. **8** (1967) 962.

[120] J. E. Hirsch, R. L. Sugar, D. J. Scalapino and R. Blankenbecler, *Monte Carlo simulations of one-dimensional fermion systems*, Phys. Rev. B **26** (1982) 5033.

[121] H. De Raedt and B. De Raedt, *Applications of the generalized Trotter formula*, Phys. Rev. A **28** (1983) 3575.

[122] M. Suzuki, *Transfer-matrix method and Monte Carlo simulation in quantum spin systems*, Phys. Rev. B **31** (1985) 2957.

[123] S. Blanes, F. Casas and J. Ros, *Extrapolation of Symplectic Integrators*, Celest. Mech. Dyn. Astron. **75** (1999) 149.

[124] S. A. Chin, *Multi-product splitting and Runge-Kutta-Nyström integrators*, Celest. Mech. Dyn. Astron. **106** (2010) 391.

[125] S. Endo, Q. Zhao, Y. Li, S. Benjamin and X. Yuan, *Mitigating algorithmic errors in a Hamiltonian simulation*, Phys. Rev. A **99** (2019) 012334.

[126] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari and R. D. Somma, "Exponential improvement in precision for simulating sparse Hamiltonians", *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, ACM, 2014.

[127] P. Zeng, J. Sun, L. Jiang and Q. Zhao, *Simple and high-precision Hamiltonian simulation by compensating Trotter error with linear combination of unitary operations*, arXiv:2212.04566 (2022).

[128] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari and R. D. Somma, *Simulating Hamiltonian Dynamics with a Truncated Taylor Series*, Phys. Rev. Lett. **114** (2015) 090502.

[129] M. Heyl, P. Hauke and P. Zoller, *Quantum localization bounds Trotter errors in digital quantum simulation*, Sci. Adv. **5** (2019) eaau8342.

[130] M. C. Tran, S.-K. Chu, Y. Su, A. M. Childs and A. V. Gorshkov, *Destructive Error Interference in Product-Formula Lattice Simulation*, Phys. Rev. Lett. **124** (2020).

[131] D. Layden, *First-Order Trotter Error from a Second-Order Perspective*, Phys. Rev. Lett. **128** (2022).

[132] S. Raeisi, N. Wiebe and B. C. Sanders, *Quantum-circuit design for efficient simulations of many-body quantum dynamics*, New J. Phys. **14** (2012) 103017.

[133] R. Babbush, J. McClean, D. Wecker, A. Aspuru-Guzik and N. Wiebe, *Chemical basis of Trotter-Suzuki errors in quantum chemistry simulation*, Phys. Rev. A **91** (2015).

[134] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker and M. Troyer, *Elucidating reaction mechanisms on quantum computers*, Proc. Natl. Acad. Sci. **114** (2017) 7555.

[135] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross and Y. Su, *Toward the first quantum simulation with quantum speedup*, Proc. Natl. Academy Sci. **115** (2018) 9456.

[136] M. Suzuki, *Fractal path integrals with applications to quantum many-body systems*, Physica A **191** (1992) 501.

[137] M. Suzuki, *Convergence of general decompositions of exponential operators*, Commun. Math. Phys. **163** (1994) 491.

[138] R. D. Ruth, *A Canonical Integration Technique*, IEEE Trans. Nucl. Sci. **30** (1983) 2669.

[139] E. Forest and R. D. Ruth, *Fourth-order symplectic integration*, Phys. D **43** (1990) 105.

[140] H. Yoshida, *Construction of higher order symplectic integrators*, Phys. Lett. A **150** (1990) 262.

[141] M. Suzuki, *Quantum analysis—non-commutative differential and integral calculi*, Commun. Math. Phys. **183** (1997) 339.

[142] N. Hatano and M. Suzuki, “Finding Exponential Product Formulas of Higher Orders”, *Quantum Annealing and Other Optimization Methods*, Springer Berlin Heidelberg, 2005 37.

[143] W. Kahan and R.-C. Li, *Composition constants for raising the orders of unconventional schemes for ordinary differential equations*, Math. Comput. **66** (1997) 1089.

[144] I. Omelyan, I. Mryglod and R. Folk, *Optimized Forest–Ruth- and Suzuki-like algorithms for integration of motion in many-body systems*, Comput. Phys. Comm. **146** (2002) 188.

[145] T. Barthel and Y. Zhang, *Optimized Lie-Trotter-Suzuki decompositions for two and three non-commuting terms*, Ann. Phys. **418**, 168165 (2020) 168165.

[146] J. Ostmeyer, *Optimised Trotter decompositions for classical and quantum computing*, J. Phys. A **56** (2023) 285303.

[147] H. De Raedt, *Product formula algorithms for solving the time dependent Schrödinger equation*, Comp. Phys. Rep. **7** (1987) 1.

[148] J. Huyghebaert and H. D. Raedt, *Product formula methods for time-dependent Schrodinger problems*, J. Phys. A **23** (1990) 5777.

[149] M. Suzuki, *General Decomposition Theory of Ordered Exponentials*, Proc. Jpn. Acad. **69** (1993) 161.

[150] D. Baye, G. Goldstein and P. Capel, *Fourth-order factorization of the evolution operator for time-dependent potentials*, Phys. Lett. A **317** (2003) 337.

[151] N. Wiebe, D. Berry, P. Høyer and B. C. Sanders, *Higher order decompositions of ordered operator exponentials*, J. Phys. A **43** (2010) 065203.

[152] D. Poulin, A. Qarry, R. Somma and F. Verstraete, *Quantum Simulation of Time-Dependent Hamiltonians and the Convenient Illusion of Hilbert Space*, Phys. Rev. Lett. **106** (2011) 170501.

[153] M. Vogl, P. Laurell, A. D. Barr and G. A. Fiete, *Analog of Hamilton-Jacobi theory for the time-evolution operator*, Phys. Rev. A **100** (2019) 012132.

[154] D. W. Berry, G. Ahokas, R. Cleve and B. C. Sanders, *Efficient Quantum Algorithms for Simulating Sparse Hamiltonians*, Commun. in Math. Phys. **270** (2007) 359.

[155] F. Verstraete, J. I. Cirac and J. I. Latorre, *Quantum circuits for strongly correlated quantum systems*, Phys. Rev. A **79** (2009) 032316.

[156] Y. Atia and D. Aharonov, *Fast-forwarding of Hamiltonians and exponentially precise measurements*, Nat. Commun. **8** (2017).

[157] B. Peng, S. Gulania, Y. Alexeev and N. Govind, *Quantum time dynamics employing the Yang-Baxter equation for circuit compression*, Phys. Rev. A **106** (2022).

[158] J. Haah, M. B. Hastings, R. Kothari and G. H. Low,
*Quantum Algorithm for Simulating Real Time Evolution of Lattice Hamiltonians*,
SIAM J. Comput. (2021) FOCS18–250–FOCS18.

[159] S. Bravyi, M. B. Hastings and F. Verstraete,
*Lieb-Robinson Bounds and the Generation of Correlations and Topological Quantum Order*,
Phys. Rev. Lett. **97** (2006).

[160] C. Cirstoiu et al., *Variational fast forwarding for quantum simulation beyond the coherence time*,
NPJ Quantum Info. **6**, 82 (2020) 82.

[161] B. Commeau et al.,
*Variational Hamiltonian Diagonalization for Dynamical Quantum Simulation*,
arXiv:2009.02559 (2020).

[162] J. Gibbs et al., *Long-time simulations for fixed input states on quantum hardware*,
NPJ Quantum Info. **8**, 135 (2022) 135.

[163] K. Heya et al., *Subspace variational quantum simulator*, Phys. Rev. Res. **5** (2023) 023078.

[164] *IBM quantum services*,
`https://quantum-computing.ibm.com/services/resources?tab=systems`,
Accessed: 09-11-2023.

[165] M. S. J. Tepaske and D. J. Luitz, *Three-dimensional isometric tensor networks*,
Phys. Rev. Res. **3** (2021).

[166] F. Verstraete, D. Porras and J. I. Cirac, *Density Matrix Renormalization Group and Periodic Boundary Conditions: A Quantum Information Perspective*, Phys. Rev. Lett. **93** (2004) 227205.

[167] F. Verstraete and J. I. Cirac,
*Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions*,
arXiv:cond-mat/0407066 (2004).

[168] G. Vidal, *Efficient Classical Simulation of Slightly Entangled Quantum Computations*,
Phys. Rev. Lett. **91** (2003) 147902.

[169] T. Zhou and D. J. Luitz,
*Operator entanglement entropy of the time evolution operator in chaotic systems*,
Phys. Rev. B **95** (2017) 094206.

[170] H. Kim and D. A. Huse,
*Ballistic Spreading of Entanglement in a Diffusive Nonintegrable System*,
Phys. Rev. Lett. **111** (2013).

[171] M. Serbyn, Z. Papić and D. A. Abanin,
*Universal Slow Growth of Entanglement in Interacting Strongly Disordered Systems*,
Phys. Rev. Lett. **110** (2013).

[172] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn and Z. Papi ć,
*Quantum scarred eigenstates in a Rydberg atom chain: Entanglement, breakdown of thermalization, and stability to perturbations*, Phys. Rev. B **98** (2018) 155134.

[173] F. M. Surace et al.,
*Lattice Gauge Theories and String Dynamics in Rydberg Atom Quantum Simulators*,
Phys. Rev. X **10** (2020) 021041.

[174] G. D. Chiara, S. Montangero, P. Calabrese and R. Fazio, *Entanglement entropy dynamics of Heisenberg chains*, J. Stat. Mech. **2006** (2006) P03001.

[175] U. Schollwöck, *The density-matrix renormalization group in the age of matrix product states*, Ann. Phys. **326** (2011) 96, January 2011 Special Issue.

[176] M. B. Hastings, *An area law for one-dimensional quantum systems*, J. Stat. Mech. **2007** (2007) P08024.

[177] J. Eisert, M. Cramer and M. B. Plenio, *Colloquium: Area laws for the entanglement entropy*, Rev. Mod. Phys. **82** (2010) 277.

[178] Y. Zhou, E. M. Stoudenmire and X. Waintal, *What Limits the Simulation of Quantum Computers?*, Phys. Rev. X **10** (2020) 041038.

[179] S. Cheng et al., *Simulating noisy quantum circuits with matrix product density operators*, Phys. Rev. Res. **3** (2021) 023005.

[180] G. H. Low, V. Kliuchnikov and N. Wiebe, *Well-conditioned multiproduct Hamiltonian simulation*, arXiv:1907.11679 (2019).

[181] M. Hauru, M. V. Damme and J. Haegeman, *Riemannian optimization of isometric tensor networks*, SciPost Phys. **10** (2021).

[182] I. A. Luchnikov, M. E. Krechetov and S. N. Filippov, *Riemannian geometry and automatic differentiation for optimization problems of quantum physics and quantum technologies*, New J. Phys. **23** (2021) 073006.

[183] T. Barthel and Q. Miao, *Absence of barren plateaus and scaling of gradients in the energy optimization of isometric tensor network states*, arXiv:2304.00161 (2023).

[184] Q. Miao and T. Barthel, *Isometric tensor network optimization for extensive Hamiltonians is free of barren plateaus*, arXiv:2304.14320 (2023).

[185] Z. Liu, Q. Ye, L.-W. Yu, L. .-.-M. Duan and D.-L. Deng, *Theory on variational high-dimensional tensor networks*, arXiv:2303.17452 (2023).

[186] H.-J. Liao, J.-G. Liu, L. Wang and T. Xiang, *Differentiable Programming Tensor Networks*, Phys. Rev. X **9** (2019).

[187] C. Geng, H.-Y. Hu and Y. Zou, *Differentiable programming of isometric tensor networks*, Mach. Learn.: Sci. Technol. **3** (2022) 015020.

[188] A. Paszke et al., “PyTorch: An Imperative Style, High-Performance Deep Learning Library”, *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Curran Associates Inc., 2019.

[189] M. P. Zaletel and F. Pollmann, *Isometric Tensor Network States in Two Dimensions*, Phys. Rev. Lett. **124** (2020).

[190] E. M. Stoudenmire and S. R. White, *Minimally entangled typical thermal state algorithms*, New J. Phys. **12** (2010) 055026.

[191] S. Anand, K. Temme, A. Kandala and M. Zaletel, *Classical benchmarking of zero noise extrapolation beyond the exactly-verifiable regime*, arXiv:2306.17839 (2023).

[192] N. Schuch, M. M. Wolf, F. Verstraete and J. I. Cirac, *Computational Complexity of Projected Entangled Pair States*, Phys. Rev. Lett. **98** (2007).

[193] M. Lubasch, J. I. Cirac and M.-C. Bañuls, *Algorithms for finite projected entangled pair states*, Phys. Rev. B **90** (2014) 064425.

[194] H. N. Phien, J. A. Bengua, H. D. Tuan, P. Corboz and R. Orús, *Infinite projected entangled pair states algorithm improved: Fast full update and gauge fixing*, Phys. Rev. B **92** (2015).

[195] M. Fannes, B. Nachtergaele and R. F. Werner, *Finitely correlated states on quantum spin chains*, Commun. Math. Phys. **144** (1992) 443.

[196] M. Foss-Feig et al., *Holographic quantum algorithms for simulating correlated spin systems*, Phys. Rev. Res. **3** (2021).

[197] D. Niu et al., *Holographic Simulation of Correlated Electrons on a Trapped-Ion Quantum Processor*, PRX Quantum **3** (2022) 030317.

[198] Y. Zhang, S. Jahanbani, D. Niu, R. Haghshenas and A. C. Potter, *Qubit-efficient simulation of thermal states with quantum tensor networks*, Phys. Rev. B **106** (16 2022) 165126.

[199] S. Anand, J. Hauschild, Y. Zhang, A. C. Potter and M. P. Zaletel, *Holographic Quantum Simulation of Entanglement Renormalization Circuits*, PRX Quantum **4** (3 2023) 030334.

[200] E. Chertkov et al., *Holographic dynamics simulations with a trapped-ion quantum computer*, Nat. Phys. **18** (2022) 1074.

[201] L. Slattery and B. K. Clark, *Quantum Circuits For Two-Dimensional Isometric Tensor Networks*, arXiv:2108.02792 (2021).

[202] Z.-Y. Wei, D. Malz and J. I. Cirac, *Sequential Generation of Projected Entangled-Pair States*, Phys. Rev. Lett. **128** (2022) 010607.

[203] D. Haag, F. Baccari and G. Styliaris, *Typical Correlation Length of Sequentially Generated Tensor Network States*, PRX Quantum **4** (2023).

[204] S.-H. Lin, M. P. Zaletel and F. Pollmann, *Efficient simulation of dynamics in two-dimensional quantum spin systems with isometric tensor networks*, Phys. Rev. B **106** (2022).

[205] T. Soejima et al., *Isometric tensor network representation of string-net liquids*, Phys. Rev. B **101** (2020).

[206] Z. Dai, Y. Wu, T. Wang and M. P. Zaletel, *Fermionic Isometric Tensor Network States in Two Dimensions*, arXiv:2211.00043 (2022).

[207] W. Kadow, F. Pollmann and M. Knap, *Isometric tensor network representations of two-dimensional thermal states*, Phys. Rev. B **107**, 205106 (2023) 205106.

[208] Y. Wu, S. Anand, S.-H. Lin, F. Pollmann and M. P. Zaletel, *Two-dimensional isometric tensor networks on an infinite strip*, Phys. Rev. B **107** (2023).

[209] R. P. Feynman, *Simulating physics with computers*, Int. J. Theor. Phys. **21** (1982) 467.

[210] D. Deutsch, *Quantum theory, the Church-Turing principle and the universal quantum computer*, Proc. Royal Soc. Lon. A **400** (1985) 97.

[211] C. H. Bennett, *Logical Reversibility of Computation*, IBM J. Res. Dev. **17** (1973) 525.

[212] P. Benioff, *The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines*, J. Stat. Phys. **22** (1980) 563.

[213] P. A. Benioff, *Quantum Mechanical Hamiltonian Models of Discrete Processes That Erase Their Own Histories: Application to Turing Machines*, Int. J. Theor. Phys. **21** (1982) 177.

[214] P. Benioff, *Quantum mechanical hamiltonian models of turing machines*, J. Stat. Phys. **29** (1982) 515.

[215] E. Bernstein and U. Vazirani, *Quantum Complexity Theory*, SIAM J. Comput. **26** (1997) 1411.

[216] R. Jozsa, *Characterizing Classes of Functions Computable by Quantum Parallelism*, Proc. Royal Soc. Lon. A **435** (1991) 563.

[217] D. Deutsch and R. Jozsa, *Rapid Solution of Problems by Quantum Computation*, Proc. Royal Soc. Lon. A **439** (1992) 553.

[218] R. P. Feynman, *Quantum mechanical computers*, Found. Phys. **16** (1986) 507.

[219] D. Deutsch, *Quantum Computational Networks*, Proc. Royal Soc. Lond. A **425** (1989) 73.

[220] A. C.-C. Yao, “Quantum Circuit Complexity”, *IEEE Annual Symposium on Foundations of Computer Science*, 1993.

[221] E. Fredkin and T. Toffoli, *Conservative logic*, Int. J. Theor. Phys. **21** (1982) 219.

[222] J. Haferkamp, P. Faist, N. B. T. Kothakonda, J. Eisert and N. Y. Halpern, *Linear growth of quantum circuit complexity*, Nat. Phys. **18** (2022) 528.

[223] D. P. DiVincenzo, *Two-bit gates are universal for quantum computation*, Phys. Rev. A **51** (1995) 1015.

[224] A. Barenco, *A universal two-bit gate for quantum computation*, Proc. Royal Soc. Lon. A **449** (1995) 679.

[225] D. Deutsch, A. Barenco and A. Ekert, *Universality in quantum computation*, Proc. Royal Soc. Lon. A **449** (1995) 669.

[226] S. Lloyd, *Universal Quantum Simulators*, Science **273** (1996) 1073.

[227] A. Barenco et al., *Elementary gates for quantum computation*, Phys. Rev. A **52** (1995) 3457.

[228] E. Knill, *Approximation by Quantum Circuits*, arXiv:quant-ph/9508006 (1995).

[229] A. Y. Kitaev, *Quantum computations: algorithms and error correction*, Russ. Math. Surv. **52** (1997) 1191.

[230] D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner and P. Zoller, *Cold Bosonic Atoms in Optical Lattices*, Phys. Rev. Lett. **81** (1998) 3108.

[231] L.-M. Duan, E. Demler and M. D. Lukin, *Controlling Spin Exchange Interactions of Ultracold Atoms in Optical Lattices*, Phys. Rev. Lett. **91** (2003) 090402.

[232] J. Simon et al., *Quantum simulation of antiferromagnetic spin chains in an optical lattice*, Nature **472** (2011) 307.

[233] A. de Paz et al., *Nonequilibrium Quantum Magnetism in a Dipolar Lattice Gas*, Phys. Rev. Lett. **111** (2013) 185305.

[234] E. Zohar, J. I. Cirac and B. Reznik, *Quantum simulations of lattice gauge theories using ultracold atoms in optical lattices*, Rep. Prog. Phys. **79** (2015) 014401.

[235] A. Mazurenko et al., *A cold-atom Fermi–Hubbard antiferromagnet*, Nature **545** (2017) 462.

[236] M. D. Lukin et al., *Dipole Blockade and Quantum Information Processing in Mesoscopic Atomic Ensembles*, Phys. Rev. Lett. **87** (2001) 037901.

[237] A. Browaeys and T. Lahaye, *Many-body physics with individually controlled Rydberg atoms*, Nat. Phys. **16** (2020) 132.

[238] M. Saffman, T. G. Walker and K. Mølmer, *Quantum information with Rydberg atoms*, Rev. Mod. Phys. **82** (2010) 2313.

[239] M. Saffman, *Quantum computing with atomic qubits and Rydberg interactions: progress and challenges*, J. Phys. B **49** (2016) 202001.

[240] J. Zeiher et al., *Coherent Many-Body Spin Dynamics in a Long-Range Interacting Ising Chain*, Phys. Rev. X **7** (2017) 041063.

[241] P. Scholl et al., *Microwave Engineering of Programmable $XXZ$ Hamiltonians in Arrays of Rydberg Atoms*, PRX Quantum **3** (2022) 020303.

[242] S. Ebadi et al., *Quantum phases of matter on a 256-atom programmable quantum simulator*, Nature **595** (2021) 227.

[243] C. Chen et al., *Continuous symmetry breaking in a two-dimensional Rydberg array*, Nature **616** (2023) 691.

[244] P. Scholl et al., *Quantum simulation of 2D antiferromagnets with hundreds of Rydberg atoms*, Nature **595** (2021) 233.

[245] H. Bernien et al., *Probing many-body dynamics on a 51-atom quantum simulator*, Nature **551** (2017) 579.

[246] D. Barredo, V. Lienhard, S. de Léséleuc, T. Lahaye and A. Browaeys, *Synthetic three-dimensional atomic structures assembled atom by atom*, Nature **561** (2018) 79.

[247] L. Henriet et al., *Quantum computing with neutral atoms*, Quantum **4** (2020) 327.

[248] D. Porras and J. I. Cirac, *Effective Quantum Spin Systems with Trapped Ions*, Phys. Rev. Lett. **92** (2004) 207901.

[249] M. Gärttner et al., *Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet*, Nat. Phys. **13** (2017) 781.

[250] A. Khromova et al.,
*Designer Spin Pseudomolecule Implemented with Trapped Ions in a Magnetic Gradient*,
Phys. Rev. Lett. **108** (2012) 220502.

[251] K. Kim et al., *Entanglement and Tunable Spin-Spin Couplings between Trapped Ions Using Multiple Transverse Modes*, Phys. Rev. Lett. **103** (2009) 120502.

[252] R. Islam et al., *Emergence and Frustration of Magnetism with Variable-Range Interactions in a Quantum Simulator*, Science **340** (2013) 583.

[253] C. Monroe et al., *Programmable quantum simulations of spin systems with trapped ions*,
Rev. Mod. Phys. **93** (2021) 025001.

[254] C. Kokail et al., *Self-verifying variational quantum simulation of lattice models*,
Nature **569** (2019) 355.

[255] D. Jaksch et al., *Fast Quantum Gates for Neutral Atoms*, Phys. Rev. Lett. **85** (2000) 2208.

[256] L. Isenhower et al., *Demonstration of a Neutral Atom Controlled-NOT Quantum Gate*,
Phys. Rev. Lett. **104** (2010) 010503.

[257] H. Weimer, M. Müller, I. Lesanovsky, P. Zoller and H. P. Büchler,
*A Rydberg quantum simulator*, Nat. Phys. **6** (2010) 382.

[258] D. Bluvstein et al., *A quantum processor based on coherent transport of entangled atom arrays*,
Nature **604** (2022) 451.

[259] S. J. Evered et al., *High-fidelity parallel entangling gates on a neutral-atom quantum computer*,
Nature **622** (2023) 268.

[260] C. J. Ballance, T. P. Harty, N. M. Linke, M. A. Sepiol and D. M. Lucas,
*High-Fidelity Quantum Logic Gates Using Trapped-Ion Hyperfine Qubits*,
Phys. Rev. Lett. **117** (2016) 060504.

[261] S. Zhang et al.,
*Error-mitigated quantum gates exceeding physical fidelities in a trapped-ion system*,
Nat. Commun. **11** (2020).

[262] I. Pogorelov et al., *Compact Ion-Trap Quantum Computing Demonstrator*,
PRX Quantum **2** (2021) 020343.

[263] S. A. Moses et al., *A Race Track Trapped-Ion Quantum Processor*, arXiv:2305.03828 (2023).

[264] S. Debnath et al.,
*Demonstration of a small programmable quantum computer with atomic qubits*,
Nature **536** (2016) 63.

[265] B. P. Lanyon et al., *Universal Digital Quantum Simulation with Trapped Ions*,
Science **334** (2011) 57.

[266] E. A. Martinez et al.,
*Real-time dynamics of lattice gauge theories with a few-qubit quantum computer*,
Nature **534** (2016) 516.

[267] M.-H. Yung et al., *From transistor to trapped-ion computers for quantum chemistry*,
Sci. Rep. **4** (2014).

[268] J. Benhelm, G. Kirchmair, C. F. Roos and R. Blatt,
*Towards fault-tolerant quantum computing with trapped ions*, Nat. Phys. **4** (2008) 463.

[269] C. Hempel et al., *Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator*, Phys. Rev. X **8** (2018) 031022.

[270] M. Iqbal et al.,
*Creation of Non-Abelian Topological Order and Anyons on a Trapped-Ion Processor*, arXiv:2305.03766 (2023).

[271] E. Chertkov et al., *Characterizing a non-equilibrium phase transition on a quantum computer*, Nat. Phys. (2023).

[272] M. Iqbal et al.,
*Topological Order from Measurements and Feed-Forward on a Trapped Ion Quantum Computer*, arXiv:2302.01917 (2023).

[273] G. A. Alvarez, D. Suter and R. Kaiser,
*Localization-delocalization transition in the dynamics of dipolar-coupled nuclear spins*, Science **349** (2015) 846.

[274] C. H. Yang et al., *Operation of a silicon quantum processor unit cell above one kelvin*, Nature **580** (2020) 350.

[275] J. J. Pla et al., *A single-atom electron spin qubit in silicon*, Nature **489** (2012) 541.

[276] A. Aspuru-Guzik and P. Walther, *Photonic quantum simulators*, Nat. Phys. **8** (2012) 285.

[277] A. Peruzzo et al., *A variational eigenvalue solver on a photonic quantum processor*, Nat. Commun. **5** (2014).

[278] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin and R. J. Schoelkopf,
*Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation*, Phys. Rev. A **69** (2004) 062320.

[279] J. Koch et al., *Charge-insensitive qubit design derived from the Cooper pair box*, Phys. Rev. A **76** (2007) 042319.

[280] L. DiCarlo et al.,
*Demonstration of two-qubit algorithms with a superconducting quantum processor*, Nature **460** (2009) 240.

[281] A. Blais, A. L. Grimsmo, S. M. Girvin and A. Wallraff, *Circuit quantum electrodynamics*, Rev. Mod. Phys. **93** (2021) 025005.

[282] G. S. Paraoanu, *Microwave-induced coupling of superconducting qubits*, Phys. Rev. B **74** (2006) 140504.

[283] C. Rigetti and M. Devoret, *Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies*, Phys. Rev. B **81** (2010) 134507.

[284] A. Kandala et al.,
*Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets*, Nature **549** (2017) 242.

[285] F. Arute et al., *Quantum Supremacy using a Programmable Superconducting Processor*, Nature **574** (2019) 505.

[286] G. A. Quantum et al., *Hartree-Fock on a superconducting qubit quantum computer*, Science **369** (2020) 1084.

[287] M. P. Harrigan et al., *Quantum approximate optimization of non-planar graph problems on a planar superconducting processor*, Nat. Phys. **17** (2021) 332.

[288] P. Frey and S. Rachel, *Realization of a discrete time crystal on 57 qubits of a quantum computer*, arXiv:2105.06632 (2021).

[289] X. Mi et al., *Time-crystalline eigenstate order on a quantum processor*, Nature **601** (2021) 531.

[290] K. J. Satzinger et al., *Realizing topologically ordered states on a quantum processor*, Science **374** (2021) 1237.

[291] M. Kjaergaard et al., *Demonstration of Density Matrix Exponentiation Using a Superconducting Quantum Processor*, Phys. Rev. X **12** (2022) 011005.

[292] Y. Zhao et al., *Realization of an Error-Correcting Surface Code with Superconducting Qubits*, Phys. Rev. Lett. **129** (2022) 030501.

[293] I. Chen, B. Burdick, Y. Yao, P. P. Orth and T. Iadecola, *Error-mitigated simulation of quantum many-body scars on quantum computers with pulse-level control*, Phys. Rev. Res. **4** (2022) 043027.

[294] E. van den Berg, Z. K. Minev, A. Kandala and K. Temme, *Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors*, Nature Physics **19** (2023) 1116.

[295] J. C. Hoke et al., *Measurement-induced entanglement and teleportation on a noisy quantum processor*, Nature **622** (2023) 481.

[296] Y. Kim et al., *Evidence for the utility of quantum computing before fault tolerance*, Nature **618** (2023) 500.

[297] J. Tindall, M. Fishman, M. Stoudenmire and D. Sels, *Efficient tensor network simulation of IBM's Eagle kicked Ising experiment*, arXiv:2306.14887 (2023).

[298] N. Moll et al., *Quantum optimization using variational algorithms on near-term quantum devices*, Quantum Sci. Technol. **3** (2018) 030503.

[299] N. M. Linke et al., *Experimental comparison of two quantum computing architectures*, Proc. Natl. Acad. Sci. **114** (2017) 3305.

[300] K. W. Plumb et al., *Continuum of quantum fluctuations in a three-dimensional $S = 1$ Heisenberg magnet*, Nat. Phys. **15** (2018) 54.

[301] Y. Ihara et al., *Emergence of Frustrated Short-Range Order above Long-Range Order in the $S = 1/2$ Kagome Antiferromagnet CaCu3(OD)6Cl2·0.6D2O*, J. Phys. Soc. Jpn **90** (2021) 023703.

[302] J. Khatua et al.,
*Experimental signatures of quantum and topological states in frustrated magnetism*, Phys. Rep. **1041** (2023) 1, Experimental signatures of quantum and topological states in frustrated magnetism.

[303] V. V. Shende, I. L. Markov and S. S. Bullock,
*Minimal universal two-qubit controlled-NOT-based circuits*, Phys. Rev. A **69** (2004).

[304] J. J. Vartiainen, M. Möttönen and M. M. Salomaa, *Efficient Decomposition of Quantum Gates*, Phys. Rev. Lett. **92** (2004).

[305] M. Moettoenen, J. J. Vartiainen, V. Bergholm and M. M. Salomaa,
*Quantum Circuits for General Multiqubit Gates*, Phys. Rev. Lett. **93** (2004).

[306] M. Möttönen, J. J. Vartiainen, V. Bergholm and M. M. Salomaa,
*Transformation of Quantum States Using Uniformly Controlled Rotations*,
Quantum Info. Comput. **5** (2005) 467.

[307] V. V. Shende and I. L. Markov,
*Quantum Circuits For Incompletely Specified Two-Qubit Operators*,
Quantum Info. Comput. **5** (2005) 048.

[308] V. Shende, S. Bullock and I. Markov, *Synthesis of quantum-logic circuits*,
IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **25** (2006) 1000.

[309] A. Lye, R. Wille and R. Drechsler, "Determining the minimal number of swap gates for multi-dimensional nearest neighbor quantum circuits",
*The 20th Asia and South Pacific Design Automation Conference*, 2015 178.

[310] R. Wille et al.,
"Look-ahead schemes for nearest neighbor optimization of 1D and 2D quantum circuits",
*2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2016 292.

[311] Y. Nam, N. J. Ross, Y. Su, A. M. Childs and D. Maslov,
*Automated optimization of large quantum circuits with continuous parameters*,
NPJ Quantum Inf. **4** (2018) 23.

[312] D. Venturelli, M. Do, E. Rieffel and J. Frank,
*Compiling quantum circuits to realistic hardware architectures using temporal planners*,
Quantum Sci. Technol. **3** (2018) 025004.

[313] G. G. Guerreschi and J. Park, *Two-step approach to scheduling quantum circuits*,
Quantum Science and Technology **3** (2018) 045003.

[314] A. Cowtan, S. Dilkes, R. Duncan, W. Simmons and S. Sivarajah,
*Phase Gadget Synthesis for Shallow Circuits*, EPTCS **318** (2020) 213.

[315] Y. Zheng and K. Kanno, *Quantum Error Detection with Generalized Syndrome Measurement*,
arXiv:2304.11532 (2023).

[316] E. Farhi, J. Goldstone, S. Gutmann and M. Sipser,
*Quantum Computation by Adiabatic Evolution*, arXiv:quant-ph/0001106 (2000).

[317] D. Aharonov and A. Ta-Shma,
"Adiabatic Quantum State Generation and Statistical Zero Knowledge",
*Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '03,
San Diego, CA, USA: Association for Computing Machinery, 2003 20, ISBN: 1581136749.

[318] E. Farhi and S. Gutmann, *Analog analogue of a digital quantum computation*,
Phys. Rev. A **57** (4 1998) 2403.

[319] A. M. Childs et al., "Exponential algorithmic speedup by a quantum walk",
*Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, ACM, 2003.

[320] E. Farhi, J. Goldstone and S. Gutmann, *A Quantum Algorithm for the Hamiltonian NAND Tree*,
Theory of Computing **4** (2008) 169.

[321] A. M. Childs, R. Cleve, S. P. Jordan and D. Yonge-Mallo,
*Discrete-Query Quantum Algorithm for NAND Trees*, Theory Comput. **5** (2009) 119.

[322] A. M. Childs and R. Kothari, "Simulating Sparse Hamiltonians with Star Decompositions",
*Theory of Quantum Computation, Communication, and Cryptography*,
Springer Berlin Heidelberg, 2011 94.

[323] P. W. Shor, *Scheme for reducing decoherence in quantum computer memory*,
Phys. Rev. A **52** (1995) R2493.

[324] A. R. Calderbank and P. W. Shor, *Good quantum error-correcting codes exist*,
Phys. Rev. A **54** (1996) 1098.

[325] D. Aharonov and M. Ben-Or, "Fault-Tolerant Quantum Computation with Constant Error",
*Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97,
El Paso, Texas, USA: Association for Computing Machinery, 1997 176, ISBN: 0897918886.

[326] P. W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, SIAM J. Comput. **26** (1997) 1484.

[327] J. Preskill, *Quantum Computing in the NISQ era and beyond*, Quantum **2** (2018) 79.

[328] A. M. Childs and N. Wiebe,
*Hamiltonian Simulation Using Linear Combinations of Unitary Operations*,
Quantum Info. Comput. **12** (2012) 901.

[329] A. M. Childs and D. W. Berry, *Black-box Hamiltonian simulation and unitary implementation*,
**12** (2012).

[330] D. W. Berry, A. M. Childs and R. Kothari,
"Hamiltonian Simulation with Nearly Optimal Dependence on all Parameters",
*2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, 2015 792.

[331] A. Carrera Vazquez, D. J. Egger, D. Ochsner and S. Woerner,
*Well-conditioned multi-product formulas for hardware-friendly Hamiltonian simulation*,
Quantum **7** (2023) 1067.

[332] G. H. Low and I. L. Chuang, *Optimal Hamiltonian Simulation by Quantum Signal Processing*,
Phys. Rev. Lett. **118** (2017) 010501.

[333] G. H. Low and I. L. Chuang, *Hamiltonian Simulation by Qubitization*, Quantum **3** (2019) 163.

[334] M. B. Hastings, *Turning Gate Synthesis Errors into Incoherent Errors*, Quantum Info. Comput. **17** (2017) 488.

[335] E. Campbell, *Shorter gate sequences for quantum computing by mixing unitaries*, Phys. Rev. A **95** (2017) 042306.

[336] E. Campbell, *Random Compiler for Fast Hamiltonian Simulation*, Phys. Rev. Lett. **123** (2019) 070503.

[337] A. M. Childs, A. Ostrander and Y. Su, *Faster quantum simulation by randomization*, Quantum **3** (2019) 182.

[338] Y. Ouyang, D. R. White and E. T. Campbell, *Compilation by stochastic Hamiltonian sparsification*, Quantum **4** (2020) 235.

[339] C. H. Cho, D. W. Berry and M.-H. Hsieh, *Doubling the order of approximation via the randomized product formula*, arXiv:2210.11281 (2022).

[340] P. K. Faehrmann, M. Steudtner, R. Kueng, M. Kieferova and J. Eisert, *Randomizing multi-product formulas for Hamiltonian simulation*, Quantum **6** (2022) 806.

[341] K. Heya, Y. Suzuki, Y. Nakamura and K. Fujii, *Variational Quantum Gate Optimization*, arXiv:1810.12745 (2018).

[342] S. Khatri et al., *Quantum-assisted quantum compiling*, Quantum **3** (2019) 140.

[343] K. Sharma, S. Khatri, M. Cerezo and P. J. Coles, *Noise resilience of variational quantum compiling*, New J. Phys. **22** (2020) 043006.

[344] M. Cerezo et al., *Variational quantum algorithms*, Nat. Rev. Phys. **3** (2021) 625.

[345] T. Jones and S. C. Benjamin, *Robust quantum compilation and circuit optimisation via energy minimisation*, Quantum **6** (2022) 628.

[346] K. Mizuta, Y. O. Nakagawa, K. Mitarai and K. Fujii, *Local Variational Quantum Compilation of Large-Scale Hamiltonian Dynamics*, PRX Quantum **3** (2022) 040302.

[347] L. Madden and A. Simonetto, *Best Approximate Quantum Compiling Problems*, ACM Trans. Quantum Comput. **3** (2022).

[348] H. Zhao, M. Bukov, M. Heyl and R. Moessner, *Adaptive Trotterization for time-dependent Hamiltonian quantum dynamics using instantaneous conservation laws*, arXiv:2307.10327 (2023).

[349] H. Zhao, M. Bukov, M. Heyl and R. Moessner, *Making Trotterization Adaptive and Energy-Self-Correcting for NISQ Devices and Beyond*, PRX Quantum **4** (3 2023) 030319.

[350] D. P. DiVincenzo and J. Smolin, *Results on two-bit gate design for quantum computers*, arXiv:cond-mat/9409111 (1994).

[351] C. Mc Keever and M. Lubasch, *Classically optimized Hamiltonian simulation*, Phys. Rev. Res. **5** (2023) 023146.

[352] R. Mansuroglu, F. Fischer and M. J. Hartmann, *Problem-specific classical optimization of Hamiltonian simulation*, Phys. Rev. Res. **5** (2023).

[353] R. Mansuroglu, T. Eckstein, L. Nützel, S. A. Wilkinson and M. J. Hartmann, *Variational Hamiltonian simulation for translational invariant systems via classical pre-processing*, Quantum Sci. Technol. **8** (2023) 025006.

[354] J. R. McClean, J. Romero, R. Babbush and A. Aspuru-Guzik, *The theory of variational hybrid quantum-classical algorithms*, New J. Phys. **18** (2016) 023023.

[355] X. Yuan, S. Endo, Q. Zhao, Y. Li and S. C. Benjamin, *Theory of variational quantum simulation*, Quantum **3** (2019) 191.

[356] R. Haghshenas, J. Gray, A. C. Potter and G. K.-L. Chan, *Variational Power of Quantum Circuit Tensor Networks*, Phys. Rev. X **12** (2022) 011047.

[357] M. B. Dov, D. Shnaiderov, A. Makmal and E. G. D. Torre, *Approximate encoding of quantum states using shallow circuits*, arXiv:2207.00028 (2022).

[358] K. Wada et al., *Simulating time evolution with fully optimized single-qubit gates on parametrized quantum circuits*, Phys. Rev. A **105** (2022) 062421.

[359] S. Barison, F. Vicentini and G. Carleo, *An efficient quantum algorithm for the time evolution of parameterized circuits*, Quantum **5** (2021) 512.

[360] M. Benedetti, M. Fiorentini and M. Lubasch, *Hardware-efficient variational quantum algorithms for time evolution*, Phys. Rev. Res. **3** (2021) 033083.

[361] Y. Li and S. C. Benjamin, *Efficient Variational Quantum Simulator Incorporating Active Error Minimization*, Phys. Rev. X **7** (2017) 021050.

[362] N. F. Berthusen, T. V. Trevisan, T. Iadecola and P. P. Orth, *Quantum dynamics simulations beyond the coherence time on noisy intermediate-scale quantum hardware by variational Trotter compression*, Phys. Rev. Res. **4** (2022) 023097.

[363] S. Wang et al., *Noise-induced barren plateaus in variational quantum algorithms*, Nat. Commun. **12** (2021).

[364] K. Temme, S. Bravyi and J. M. Gambetta, *Error Mitigation for Short-Depth Quantum Circuits*, Phys. Rev. Lett. **119** (2017) 180509.

[365] S. Endo, S. C. Benjamin and Y. Li, *Practical Quantum Error Mitigation for Near-Future Applications*, Phys. Rev. X **8** (2018) 031027.

[366] A. Kandala et al., *Error mitigation extends the computational reach of a noisy quantum processor*, Nature **567** (2019) 491.

## APPENDIX A

# Article 1

PHYSICAL REVIEW LETTERS **130,** 086704 (2023)

# Field-Tunable Berezinskii-Kosterlitz-Thouless Correlations in a Heisenberg Magnet

D. Opherden,[1] M. S. J. Tepaske,[2,3] F. Bärtl,[1,4] M. Weber,[3] M. M. Turnbull,[5] T. Lancaster,[6] S. J. Blundell,[7]
M. Baenitz,[8] J. Wosnitza,[1,4] C. P. Landee,[9] R. Moessner,[3] D. J. Luitz,[2,3] and H. Kühne[1,*]
[1]*Hochfeld-Magnetlabor Dresden (HLD-EMFL) and Würzburg-Dresden Cluster of Excellence ct.qmat,*
*Helmholtz-Zentrum Dresden-Rossendorf, 01328 Dresden, Germany*
[2]*Physikalisches Institut, Universität Bonn, Nussallee 12, 53115 Bonn, Germany*
[3]*Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*
[4]*Institut für Festkörper- und Materialphysik, TU Dresden, 01062 Dresden, Germany*
[5]*Carlson School of Chemistry, Clark University, Worcester, Massachusetts 01610, USA*
[6]*Department of Physics, Centre for Materials Physics, Durham University, Durham DH1 3LE, United Kingdom*
[7]*Clarendon Laboratory, Department of Physics, University of Oxford, Park Road, Oxford OX1 3PU, United Kingdom*
[8]*Max Planck Institute for Chemical Physics of Solids, 01187 Dresden, Germany*
[9]*Department of Physics, Clark University, Worcester, Massachusetts 01610, USA*

(Received 14 September 2022; accepted 18 January 2023; published 23 February 2023)

We report the manifestation of field-induced Berezinskii-Kosterlitz-Thouless (BKT) correlations in the weakly coupled spin-1/2 Heisenberg layers of the molecular-based bulk material $[\mathrm{Cu(pz)_2(2\text{-}HOpy)_2}](\mathrm{PF_6})_2$. At zero field, a transition to long-range order occurs at 1.38 K, caused by a weak intrinsic easy-plane anisotropy and an interlayer exchange of $J'/k_B \approx 1$ mK. Because of the moderate intralayer exchange coupling of $J/k_B = 6.8$ K, the application of laboratory magnetic fields induces a substantial $XY$ anisotropy of the spin correlations. Crucially, this provides a significant BKT regime, as the tiny interlayer exchange $J'$ only induces 3D correlations upon close approach to the BKT transition with its exponential growth in the spin-correlation length. We employ nuclear magnetic resonance measurements to probe the spin correlations that determine the critical temperatures of the BKT transition as well as that of the onset of long-range order. Further, we perform stochastic series expansion quantum Monte Carlo simulations based on the experimentally determined model parameters. Finite-size scaling of the in-plane spin stiffness yields excellent agreement of critical temperatures between theory and experiment, providing clear evidence that the nonmonotonic magnetic phase diagram of $[\mathrm{Cu(pz)_2(2\text{-}HOpy)_2}](\mathrm{PF_6})_2$ is determined by the field-tuned $XY$ anisotropy and the concomitant BKT physics.

DOI: 10.1103/PhysRevLett.130.086704

Cooperative behavior and critical phenomena of strongly correlated magnets are typically dictated by the lattice and spin dimensions, as well as by the symmetry of the underlying Hamiltonian [1–8]. Among the most fascinating examples are two-dimensional (2D) $XY$ spin systems, which are known to undergo a topological Berezinskii-Kosterlitz-Thouless phase transition at a finite temperature $T_{\mathrm{BKT}}$ [9–11], which marks the binding of topological defects in vortex-antivortex pairs. So far, experimental efforts to probe a genuine BKT transition in a bulk material were compromised by the onset of 3D order [12–18] due to the inherent 3D nature of these materials. Still, if the perturbative terms relative to a purely 2D $XY$ model are small enough, the experimental observation of magnetic properties associated with BKT correlations may be possible in the transition regime [19–23].

In particular, a controlled tuning of the $XY$ anisotropy, with associated impact on $T_{\mathrm{BKT}}$, can provide an ideal test bed for experimental studies of BKT physics and their comparison to numerical state-of-the-art modeling. As a possible approach to tune the magnetic correlations away from 2D Heisenberg to a 2D $XY$ symmetry, the application of a uniform magnetic field to the 2D quantum Heisenberg antiferromagnet breaks the $O(3)$ symmetry, but preserves the easy-plane $O(2)$ symmetry, as was confirmed by quantum Monte Carlo (QMC) calculations [24]. Correspondingly, for Zeeman energies of the order of the exchange energy, the effective $XY$-exchange anisotropy can be controlled. The associated BKT transition persists for all fields below saturation, yielding a nonmonotonic magnetic phase diagram [24].

In order to find materials that allow us to study this phenomenology, the chemical engineering of molecular-based bulk magnets is a promising approach. By an appropriate choice of molecular ligands and counterions, the syntheses of several materials that realize a 2D spin-1/2 Heisenberg model on the square lattice were reported [25–35]. In these materials, a moderate

© 2023 American Physical Society

nearest-neighbor exchange interaction of the order of a few K allows for the tunability of the effective exchange anisotropy by experimentally accessible magnetic fields. Indeed, for several $Cu^{2+}$-based molecular materials, a nonmonotonic magnetic phase diagram as a function of the external field was reported [25,30,32,36,37]. The magnetic properties of these molecular-based 2D quantum Heisenberg antiferromagnets were mostly investigated by thermodynamic methods [26,30,32,34,35], thus missing local information about the magnetic correlations in the BKT transition regime.

In this Letter, we report on the field-tunable anisotropy of magnetic correlations in $[Cu(pz)_2(2\text{-HOpy})_2](PF_6)_2$ [with $pz = C_4H_4N_2$, 2-HOpy $= C_5H_4NHO$] (CuPOF in the following), ranging from the almost-isotropic Heisenberg limit at zero field to a substantial $XY$ anisotropy upon increasing the magnetic field strength. We use nuclear magnetic resonance (NMR) as the experimental probe for the dynamic and quasistatic spin correlations. Furthermore, by QMC simulations, we calculate the in-plane spin stiffness, which we use to determine the critical temperatures of the long-range order (LRO) and the BKT transition. Our main findings are (i) that the temperature dependence of the nuclear spin-lattice relaxation rate follows the behavior predicted from 2D BKT theory in a wide range of temperatures, determined by the field-driven anisotropy, (ii) that finite-size scaling of the QMC results permits the extraction of $T_{\mathrm{BKT}}$, which lies below the actual 3D ordering temperature $T_{\mathrm{LRO}}$, and (iii) that both temperatures exhibit a nonmonotonic field dependence, which is analogous to the behavior when instead of the field, the anisotropy of interactions is tuned, a clear signature for the tunability of BKT correlations.

The synthesis and characterization of CuPOF by means of various techniques, including $\mu^+$SR experiments, are described in Ref. [25]. The crystals are flat plates with the crystallographic $c$ axis perpendicular to the plate. The NMR spectra and spin-lattice relaxation time $T_1$ were recorded using a standard Hahn spin-echo pulse sequence and an inversion-recovery method, respectively. The measurements were performed using a commercial phase-coherent spectrometer and a 16 T superconducting magnet, equipped with a $^3$He sample-in-liquid cryostat. A single-axis goniometer was used to align the $c$ axis parallel to the magnetic field.

The magnetic interactions of CuPOF in an applied field are well approximated by the effective Hamiltonian

$$\begin{aligned}\mathcal{H} = J\sum_{\langle i,j\rangle_{\parallel}}[S_i^x S_j^x + S_i^y S_j^y + (1-\Delta)S_i^z S_j^z] \\ + J'\sum_{\langle i,j\rangle_{\perp}}\mathbf{S}_i\cdot\mathbf{S}_j - g\mu_B\mu_0 H\sum_i S_i^z, \qquad (1)\end{aligned}$$

where $\langle i,j\rangle_{\parallel}$ and $\langle i,j\rangle_{\perp}$ denote the intra- and interlayer nearest neighbors, and $J$ and $J'$ are the intra- and interlayer exchange couplings, estimated as $J/k_B = 6.8$ K and $J'/k_B \approx 1$ mK [25]. Whereas $\Delta = 0$ corresponds to the isotropic Heisenberg case, $0 < \Delta \leq 1$ quantifies an easy-plane anisotropy, with a zero-field value of $\Delta \approx 0.01\ldots0.02$ for CuPOF [25].

In the presence of interlayer interactions, any nonfrustrated magnetic quasi-2D lattice inevitably undergoes a transition to long-range order at low temperatures. Because of the very large separation of the magnetic layers in CuPOF, with $J'/J \approx 1.4 \times 10^{-4}$, the very small entropy change associated with the transition to LRO is beyond the experimental resolution of thermodynamic quantities [25,38]. On the other hand, $\mu^+$SR is very sensitive to the local staggered magnetization, and was used to probe the transition to LRO at 1.38 (2) K in CuPOF [25]. This transition occurs under the influence of the weak intrinsic easy-plane anisotropy, which yields a temperature-driven crossover from isotropic to $XY$-type correlations at the crossover temperature $T_{\mathrm{co}} > T_{\mathrm{LRO}}$. An applied magnetic field increases the effective $XY$ anisotropy, which manifests itself as a field-dependent minimum of the uniform bulk susceptibility at $T_{\mathrm{co}}$, as depicted by the pentagons in Fig. 1.

The temperature dependence of the $^{31}$P-NMR spin-lattice relaxation rate at out-of-plane fields up to 16 T is presented

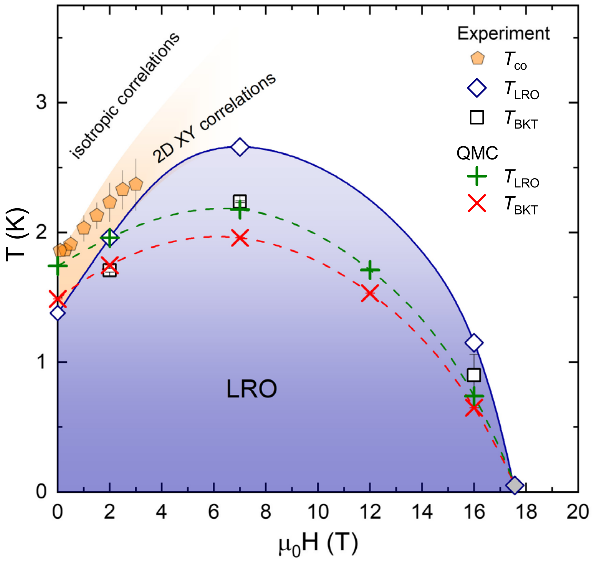


FIG. 1. Phase diagram of CuPOF for out-of-plane magnetic fields from experiment and numerics. The pentagons denote the spin-anisotropy crossover temperature $T_{\mathrm{co}}$ from Ref. [25]. White diamonds indicate the transition temperature $T_{\mathrm{LRO}}$ to long-range order, and squares show the BKT transition temperature $T_{\mathrm{BKT}}$, as obtained from the analysis of the $^{31}$P $1/T_1$ rate (Fig. 2). $T_{\mathrm{LRO}}$ at zero field is determined by $\mu^+$SR measurements [25]. The green pluses and red crosses denote $T_{\mathrm{LRO}}$ and $T_{\mathrm{BKT}}$, respectively, as obtained from QMC calculations (Fig. 3). The diamond at 17.5 T denotes the saturation field, which was determined from magnetization experiments [25], and is in agreement with QMC results. All lines are guides to the eye.

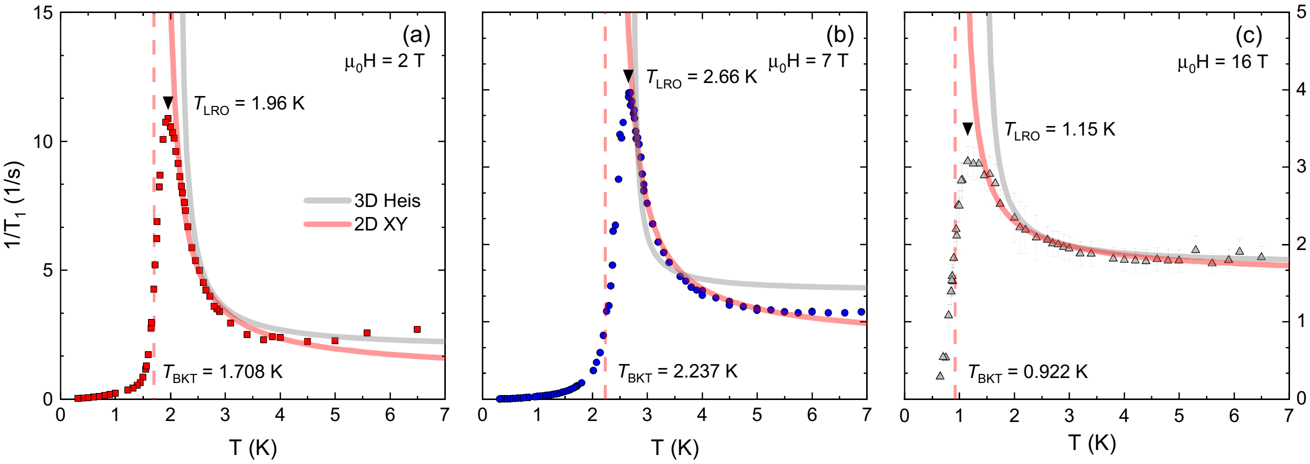


FIG. 2. (a)–(c) Temperature-dependent $^{31}$P nuclear spin-lattice relaxation rate $1/T_1$ of CuPOF, recorded at out-of-plane fields of 2, 7, and 16 T. The solid lines are best fits according to $1/T_1 \propto \xi^{z-\eta}$ for the temperature dependent correlation lengths $\xi_{\mathrm{3DHeis}}$ and $\xi_{\mathrm{2DXY}}$ of the 3D Heisenberg and the 2D $XY$ cases (see main text). The transition temperature $T_{\mathrm{LRO}}$, marked with a downward triangle, is inferred from the $1/T_1$ peak position, and $T_{\mathrm{BKT}}$, marked with a dotted line, is determined from fits according to $1/T_1 \propto \xi_{\mathrm{2DXY}}^{z-\eta}$. At all fields, but most noticeably at 7 T, $1/T_1$ is described best by $\xi_{\mathrm{2DXY}}$ at $T \gtrsim T_{\mathrm{LRO}}$.

in Figs. 2(a)–2(c). The spin-lattice relaxation rate $1/T_1$ has sharp maxima at $T_{\mathrm{LRO}} = 1.96$ and 2.66 K at 2 and 7 T, respectively. In comparison, the maximum amplitude of $1/T_1$ at 16 T ($T_{\mathrm{LRO}} = 1.15$ K) is substantially reduced. The transition temperatures between the 2D $XY$ and the LRO regimes are depicted by diamonds in Fig. 1. The strong dependence of $T_{\mathrm{LRO}}$ on the field strength that we observe in CuPOF clearly indicates a field tunability of the $XY$ anisotropy of the spin correlations [24]. This behavior is confirmed by our QMC simulations.

As previously reported, the $^{31}$P $1/T_1$ rate in CuPOF yields several broad maxima at high temperatures, which are associated with a freezing of the $PF_6$ molecular reorientation modes [39]. Below about 10 K, in the range of interest in the present study, these modes are frozen out and $1/T_1$ becomes temperature independent, indicating predominantly paramagnetic fluctuations. In 2D magnetic lattices, the onset of short-range spin correlations occurs at temperatures $T \simeq J/k_B$ [38], with a correlation length of about one magnetic-lattice constant [16,40].

At temperatures above the onset of LRO, $1/T_1$ can serve as a probe for the dynamic correlation length $\xi$ [19,20,41–44]. As was shown from dynamical scaling arguments [41], $1/T_1$ is proportional to the transverse spin correlation length as $1/T_1 \propto \xi^{z-\eta}$, where $z$ and $\eta$ are characteristic dynamic and critical exponents [3,19,41,45]. By comparing the temperature dependence of $1/T_1$ with the characteristic $\xi$ of different universality classes, we can therefore probe the nature of the predominant correlations in the critical regime, before the system finally undergoes the transition to long-range order. Thus, we compare the BKT correlation length of a 2D easy-plane antiferromagnet, $\xi_{\mathrm{2DXY}} \propto \exp(0.5\pi/\sqrt{T/T_{\mathrm{BKT}} - 1})$ [11,16], with that of a 3D isotropic Heisenberg antiferromagnet, $\xi_{\mathrm{3DHeis}} \propto |T - T_{\mathrm{LRO}}|^{-0.7112}$ [46,47].

To describe $1/T_1$ in the interval $T_{\mathrm{LRO}} \leq T \leq J/k_B$, we note that $\eta = 0.0375$ for the 3D Heisenberg antiferromagnet [46], with the LRO transition residing in the $O(3)$ universality class, whereas the easy-plane model has $\eta = 1/4$ [48–50]. For both models, we use $z = d/2$, with $d$ the spatial dimensionality [45]. The experimental estimates of $T_{\mathrm{BKT}}$ are obtained from fits to the 2D $XY$ form.

In Figs. 2(a)–2(c), we show the measured $1/T_1$ along with both fits, for fields of 2, 7, and 16 T [51]. In contrast to the 3D Heisenberg description, the 2D $XY$ fit accurately captures the increase of $1/T_1$ near $T_{\mathrm{LRO}}$, most noticeably at 7 T. The fits yield $T_{\mathrm{BKT}} = 1.708(14)$, 2.237(7), and 0.90(16) K for applied fields of 2, 7, and 16 T, respectively, with errors determined by bootstrapping. The nonmonotonic dependence of $T_{\mathrm{BKT}}$ on the field tracks that of $T_{\mathrm{LRO}}$, being separated by a few hundred mK for the most part, as shown in the phase diagram in Fig. 1. One should note, however, that the BKT transition is preempted by the LRO that arises from the 3D correlations, stemming from the finite interlayer exchange interaction $J'$. In the Supplemental Material [52], we discuss indications that changing the field strength has similar effects on the spin correlations as changing the exchange anisotropy $\Delta$ [16,62] and argue that hence the field allows us to tune the effective anisotropy. Further, as shown by $\mu^+$SR and $^1$H-NMR spectroscopy, the low-temperature staggered magnetization in CuPOF agrees with a change from Heisenberg behavior at zero field towards that of a 2D $XY$ system at 7 T, see Fig. S2 in the Supplemental Material [52].

In order to shed more light on the experimentally observed phenomenology of mixed Néel and BKT-type correlations, we numerically investigate the Hamiltonian (1) using stochastic series expansion quantum Monte Carlo with directed loops [63]. We consider finite simple-cubic lattices with periodic boundary conditions and dimensions $L \times L \times L/8$, fixing $J/k_B = 6.8$ K, $J'/k_B = 1$ mK, and $\Delta = 0.0185$. To determine $T_{\mathrm{BKT}}$ and $T_{\mathrm{LRO}}$, we calculate the in-plane spin stiffness $\rho = 8L^{-3}\partial^2 F/\partial\phi^2|_{\phi=0}$, which is defined as the second derivative of the free energy $F$ with respect to a uniform in-plane twist angle $\phi$ [64,65]. This quantity is nonzero in the BKT phase and in the thermodynamic limit it should vanish instantly at $T_{\mathrm{BKT}}$. For the finite lattices simulated with QMC, this drop-off is instead continuous, but based on how $\rho$ approaches the instant drop-off with increasing system size, we can determine $T_{\mathrm{BKT}}$. In particular, using finite-size scaling theory, it is predicted that $\rho$ depends on temperature $T$ and system size $L$ as [64]

$$\rho(T, L)/P(L) = f\left(\ln(L) - \frac{a}{\sqrt{T - T_{\mathrm{BKT}}}}\right), \qquad (2)$$

$$P(L) = 1 + \frac{1}{2\ln(L) + c + \ln[c/2 + \ln L]} + \frac{b}{\ln^2 L}, \qquad (3)$$

where $a$, $b$, $c$ are fitting constants and $f$ is a general continuous function which we choose to be a fifth-order polynomial. This parametrization of $\rho$ is fitted closely above $T_{\mathrm{BKT}}$ for simulation data of the $J' = 0$ model to deduce $T_{\mathrm{BKT}}$. Afterwards, we plot $\rho/P$ versus $\ln(L) - a/\sqrt{T - T_{\mathrm{BKT}}}$ in the fitting interval, which should collapse to a single curve if the fit is perfect. We checked that fitting $\rho$ for $J' = 1$ mK in the full 3D model reproduces the 2D $T_{\mathrm{BKT}}$ to within error bars, when fitted at $T > T_{\mathrm{LRO}}$, where the interlayer coupling becomes insignificant such that the 2D scaling ansatz (3) holds. In Fig. 3(a), we show the finite-size collapse of the $\rho$ fit performed at 2 T, for systems with up to $1 \times 10^6$ spins and a temperature grid of $\Delta T = 1$ mK. The fit yields $T_{\mathrm{BKT}} = 1.748(15)$ K.

To determine $T_{\mathrm{LRO}}$, we consider the scaled in-plane stiffness $L\rho$ for the full 3D model with $J' = 1$ mK. At large $L$, this quantity becomes size independent at $T_{\mathrm{LRO}}$ [47]. Hence, by determining the crossings $T^*$ between $L\rho$ curves with two different sizes $L$, and extrapolating this crossing temperature to $L \to \infty$, we obtain $T_{\mathrm{LRO}}$ [66]. In Fig. 3(b), we show the scaling analysis performed for $L\rho$ at 2 T, where the inset shows the $L \to \infty$ scaling of the crossing temperature $T^*$. Here, we used a second-order polynomial, which yields $T_{\mathrm{LRO}} = 1.959(2)$ K. Further calculations of the relevant magnetization components and correlation length are presented in Fig. S3 of the Supplemental Material [52].

Employing these procedures at different magnetic fields, we determined $T_{\mathrm{BKT}} = 1.4877(6)$, $1.7477(15)$, $1.9584(24)$, $1.5323(13)$, and $0.6495(15)$ K, at fields of 0, 2, 7, 12, and 16 T, respectively. We also confirmed that $T_{\mathrm{BKT}} = 0$ when both $\Delta = 0$ and $H = 0$, which emphasizes the strong effect on $T_{\mathrm{BKT}}$ of the seemingly small $\Delta = 0.0185$ for CuPOF. Furthermore, we determined $T_{\mathrm{LRO}} = 1.7425(19)$, $1.9597(20)$, $2.1768(23)$, $1.7110(22)$, and $0.7376(17)$ K. At all fields, our calculations yield $T_{\mathrm{LRO}} > T_{\mathrm{BKT}}$, thus supporting the experimental phenomenology, as can be seen in Fig. 1. We also determined the saturation field to be 17.5 T, in excellent agreement with the experimental value. As in the experiment, the strong dependence of the numerically determined $T_{\mathrm{LRO}}$ on the field strength reflects the effect of the field-induced anisotropy. The quantitative differences to the experimental transition temperatures at elevated fields might be resolved by extending the complexity of the modelling. In Fig. S4 of the Supplemental Material, we obtain a simple estimate of an effective exchange anisotropy $\Delta(H)$ at $H \leq 6$ T and compare it to the low-field results [52].

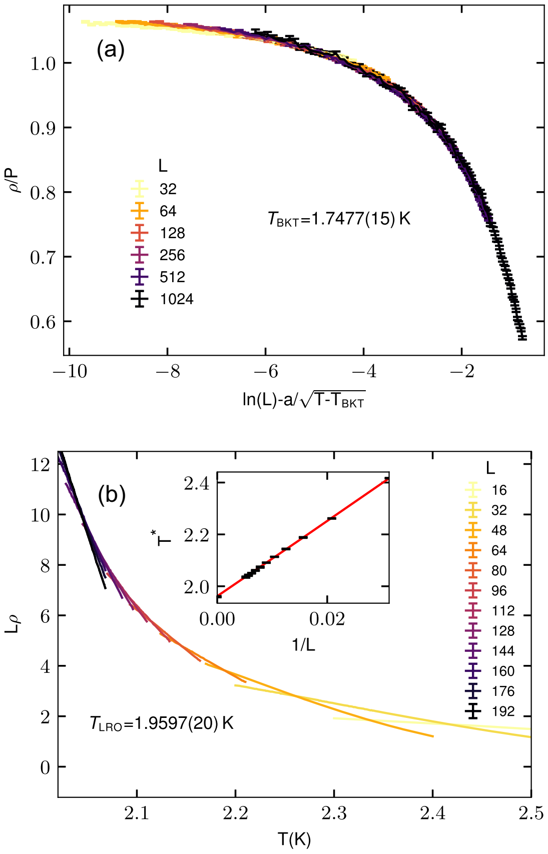


FIG. 3. Finite-size scaling analysis performed to obtain the critical temperatures $T_{\mathrm{BKT}}$ and $T_{\mathrm{LRO}}$ from the QMC simulations at 2 T. (a) Data collapse of the finite-size in-plane spin stiffness $\rho$ fit closely above $T_{\mathrm{BKT}}$, for the $J' = 0$ model, which should collapse to a single curve if the fit is perfect, reaffirming the calculated $T_{\mathrm{BKT}}$. The different curves correspond to different linear sizes $L$. (b) Crossings of the $L\rho$ curves for the $J' = 1$ mK model; the inset shows the $L \to \infty$ scaling of the crossing temperature $T^*$. The red line denotes a second-order polynomial fit, which is extrapolated to $1/L \to 0$ to estimate $T_{\mathrm{LRO}}$.

Our findings suggest the following scenario for the temperature evolution of spin correlations in CuPOF in applied magnetic fields. Decreasing the temperatures from the paramagnetic high-temperature limit, isotropic Heisenberg-type spin correlations develop, which cross over to an anisotropic $XY$-type close to $T_{\mathrm{co}}$. With further decreasing temperature, the correlation length $\xi$ grows exponentially due to the vortex physics described by BKT theory. For $T \gtrsim T_{\mathrm{BKT}}$, a rather low density of these topological excitations is expected [67]. The exponential increase of $\xi$ yields a rapid strengthening of the antiferromagnetic correlations in the $XY$ regime and, therefore, the staggered magnetization appears effectively nonzero even above $T_{\mathrm{LRO}}$ (see Supplemental Material) [52]. With further increase of $\xi$ upon lowering the temperature further, the magnetic correlations, due to the influence of the small but nonzero interlayer interaction $J'$ on the regions with large in-plane correlation lengths, can no longer be treated as 2D, and a transition to long-range order occurs at $T_{\mathrm{LRO}}$. As a consequence of the field-induced BKT-type spin correlations, a concomitant nonmonotonic behavior of the transition temperature $T_{\mathrm{LRO}}$ is observed experimentally and confirmed by our QMC simulations.

In conclusion, the very good agreement between our experimental results and the matching QMC calculations establishes our study of CuPOF as a model case, where the application of a magnetic field allows a controlled tuning of the spin-$1/2$ system from the almost isotropic 2D Heisenberg to the highly anisotropic 2D $XY$ limit. The phenomenology in CuPOF is driven by field-induced Berezinskii-Kosterlitz-Thouless physics under the influence of extremely small interplane interactions, thus providing an attractive opportunity for systematic investigations of the BKT-type topological excitations and calling for further experimental studies by inelastic scattering techniques.

Data presented in this Letter resulting from the UK effort are available from [68].

We acknowledge the support of HLD at HZDR, a member of the European Magnetic Field Laboratory (EMFL). This work was supported by the Deutsche Forschungsgemeinschaft (DFG) through SFB 1143 (project ID 247310070), the Würzburg-Dresden Cluster of Excellence on Complexity and Topology in Quantum Matter—*ct.qmat* (EXC 2147, project ID 390858490) and the cluster of excellence ML4Q (EXC2004, project ID 390534769). Our QMC code is based on the ALPS libraries [69]. Part of this work was carried out at the STFC ISIS facility and we are grateful for provision of beam time. This work is supported by EPSRC (UK).

D. O. and M. S. J. T. contributed equally to this work.

---

[*] Corresponding author.
h.kuehne@hzdr.de

[1] L. Onsager, Phys. Rev. **65**, 117 (1944).
[2] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).
[3] *Magnetic Properties of Layered Transition Metal Compounds*, edited by L. J. d. Jongh (Springer, Netherlands, 1990).
[4] B. Widom, J. Chem. Phys. **43**, 3898 (1965).
[5] M. E. Fisher, Rev. Mod. Phys. **46**, 597 (1974).
[6] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).
[7] M. J. Feigenbaum, Physica (Amsterdam) **7D**, 16 (1983).
[8] P. W. Anderson, Science **235**, 1196 (1987).
[9] V. L. Berezinskii, J. Exp. Theor. Phys. **32**, 493 (1971).
[10] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).
[11] J. M. Kosterlitz, J. Phys. Condens. Matter **7**, 1046 (1974).
[12] B.-G. Liu, Phys. Rev. B **41**, 9563 (1990).
[13] N. Majlis, S. Selzer, and G. C. Strinati, Phys. Rev. B **45**, 7872 (1992).
[14] N. Majlis, S. Selzer, and G. C. Strinati, Phys. Rev. B **48**, 957 (1993).
[15] L. Siurakshina, D. Ihle, and R. Hayn, Phys. Rev. B **61**, 14601 (2000).
[16] H.-Q. Ding, Phys. Rev. Lett. **68**, 1927 (1992).
[17] H.-Q. Ding, J. Phys. Condens. Matter **2**, 7979 (1990).
[18] A. Cuccoli, T. Roscilde, R. Vaia, and P. Verrucchi, Phys. Rev. Lett. **90**, 167205 (2003).
[19] B. J. Suh, F. Borsa, L. L. Miller, M. Corti, D. C. Johnston, and D. R. Torgeson, Phys. Rev. Lett. **75**, 2212 (1995).
[20] D. Waibel, G. Fischer, T. Wolf, H. v. Löhneysen, and B. Pilawa, Phys. Rev. B **91**, 214412 (2015).
[21] E. S. Klyushina, J. Reuther, L. Weber, A. T. M. N. Islam, J. S. Lord, B. Klemke, M. Må nsson, S. Wessel, and B. Lake, Phys. Rev. B **104**, 064402 (2021).
[22] N. Caci, L. Weber, and S. Wessel, Phys. Rev. B **104**, 155139 (2021).
[23] P. Carretta, M. Filibian, R. Nath, C. Geibel, and P. J. C. King, Phys. Rev. B **79**, 224432 (2009).
[24] A. Cuccoli, T. Roscilde, R. Vaia, and P. Verrucchi, Phys. Rev. B **68**, 060402(R) (2003).
[25] D. Opherden *et al.*, Phys. Rev. B **102**, 064431 (2020).
[26] F. M. Woodward, P. J. Gibson, G. B. Jameson, C. P. Landee, M. M. Turnbull, and R. D. Willett, Inorg. Chem. **46**, 4256 (2007).
[27] P. A. Goddard, J. Singleton, P. Sengupta, R. D. McDonald, T. Lancaster, S. J. Blundell, F. L. Pratt, S. Cox, N. Harrison, J. L. Manson, H. I. Southerland, and J. A. Schlueter, New J. Phys. **10**, 083025 (2008).
[28] F. Xiao, F. M. Woodward, C. P. Landee, M. M. Turnbull, C. Mielke, N. Harrison, T. Lancaster, S. J. Blundell, P. J. Baker, P. Babkevich, and F. L. Pratt, Phys. Rev. B **79**, 134412 (2009).
[29] J. L. Manson, K. H. Stone, H. I. Southerland, T. Lancaster, A. J. Steele, S. J. Blundell, F. L. Pratt, P. J. Baker, R. D. McDonald, P. Sengupta, J. Singleton, P. A. Goddard, C. Lee, M.-H. Whangbo, M. M. Warter, C. H. Mielke, and P. W. Stephens, J. Am. Chem. Soc. **131**, 4590 (2009).
[30] E. Čižmár, S. A. Zvyagin, R. Beyer, M. Uhlarz, M. Ozerov, Y. Skourski, J. L. Manson, J. A. Schlueter, and J. Wosnitza, Phys. Rev. B **81**, 064422 (2010).

[31] A. J. Steele, T. Lancaster, S. J. Blundell, P. J. Baker, F. L. Pratt, C. Baines, M. M. Conner, H. I. Southerland, J. L. Manson, and J. A. Schlueter, Phys. Rev. B **84**, 064412 (2011).
[32] Y. Kohama, M. Jaime, O. E. Ayala-Valenzuela, R. D. McDonald, E. D. Mun, J. F. Corbey, and J. L. Manson, Phys. Rev. B **84**, 184402 (2011).
[33] P. A. Goddard, J. L. Manson, J. Singleton, I. Franke, T. Lancaster, A. J. Steele, S. J. Blundell, C. Baines, F. L. Pratt, R. D. McDonald, O. E. Ayala-Valenzuela, J. F. Corbey, H. I. Southerland, P. Sengupta, and J. A. Schlueter, Phys. Rev. Lett. **108**, 077208 (2012).
[34] P. A. Goddard *et al.*, Phys. Rev. B **93**, 094430 (2016).
[35] V. Selmani, C. P. Landee, M. M. Turnbull, J. L. Wikaira, and F. Xiao, Inorg. Chem. Commun. **13**, 1399 (2010).
[36] P. Sengupta, C. D. Batista, R. D. McDonald, S. Cox, J. Singleton, L. Huang, T. P. Papageorgiou, O. Ignatchik, T. Herrmannsdörfer, J. L. Manson, J. A. Schlueter, K. A. Funk, and J. Wosnitza, Phys. Rev. B **79**, 060409(R) (2009).
[37] N. A. Fortune, S. T. Hannahs, C. P. Landee, M. M. Turnbull, and F. Xiao, J. Phys. Conf. Ser. **568**, 042004 (2014).
[38] P. Sengupta, A. W. Sandvik, and R. R. P. Singh, Phys. Rev. B **68**, 094423 (2003).
[39] D. Opherden, F. Bärtl, S. Yamamoto, Z. T. Zhang, S. Luther, S. Molatta, J. Wosnitza, M. Baenitz, I. Heinmaa, R. Stern, C. P. Landee, and H. Kühne, Phys. Rev. B **103**, 014428 (2021).
[40] H.-Q. Ding and M. S. Makivić, Phys. Rev. Lett. **64**, 1449 (1990).
[41] F. Borsa, M. Corti, T. Goto, A. Rigamonti, D. C. Johnston, and F. C. Chou, Phys. Rev. B **45**, 5756 (1992).
[42] F. Tabak, A. Lascialfari, and A. Rigamonti, J. Phys. Condens. Matter **5**, B31 (1993).
[43] L. Bossoni, P. Carretta, R. Nath, M. Moscardini, M. Baenitz, and C. Geibel, Phys. Rev. B **83**, 014412 (2011).
[44] P. Gaveau, J. P. Boucher, L. P. Regnault, and Y. Henry, J. Appl. Phys. **69**, 6228 (1991).
[45] P. C. Hohenberg and B. I. Halperin, Rev. Mod. Phys. **49**, 435 (1977).
[46] M. Campostrini, M. Hasenbusch, A. Pelissetto, P. Rossi, and E. Vicari, Phys. Rev. B **65**, 144520 (2002).
[47] A. W. Sandvik, Phys. Rev. Lett. **80**, 5196 (1998).
[48] D. R. Nelson and J. M. Kosterlitz, Phys. Rev. Lett. **39**, 1201 (1977).
[49] M. S. Makivić, Phys. Rev. B **46**, 3167 (1992).
[50] H.-Q. Ding, Phys. Rev. B **45**, 230 (1992).
[51] We note that a constant was added in the fits to account for the paramagnetic contribution to $1/T_1$.
[52] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.130.086704 for additional information about the staggered spin correlations investigated by means of NMR, $\mu^+$SR, and QMC, as well as an estimate of the field-induced exchange anisotropy, which includes Refs. [53–61].
[53] J. Als-Nielsen, S. T. Bramwell, M. T. Hutchings, G. J. McIntyre, and D. Visser, J. Phys. Condens. Matter **5**, 7871 (1993).
[54] S. V. Isakov and R. Moessner, Phys. Rev. B **68**, 104409 (2003).
[55] S. T. Bramwell and P. C. W. Holdsworth, J. Phys. Condens. Matter **5**, L53 (1993).
[56] S. T. Bramwell and P. C. W. Holdsworth, Phys. Rev. B **49**, 8811 (1994).
[57] A. Taroni, S. T. Bramwell, and P. C. W. Holdsworth, J. Phys. Condens. Matter **20**, 275233 (2008).
[58] M. Troyer, M. Imada, and K. Ueda, J. Phys. Soc. Jpn. **66**, 2957 (1997).
[59] J.-K. Kim and M. Troyer, Phys. Rev. Lett. **80**, 2705 (1998).
[60] M. Troyer, Universality in two-dimensional quantum Heisenberg antiferromagnets, in *Open Problems in Strongly Correlated Electron Systems*, edited by J. Bonča, P. Prelovšek, A. Ramšak, and S. Sarkar, NATO Science Series Vol. 15 (Springer, Dordrecht, 2001).
[61] P. Hasenfratz and F. Niedermayer, Phys. Lett. B **268**, 231 (1991).
[62] A. Cuccoli, T. Roscilde, V. Tognetti, R. Vaia, and P. Verrucchi, Phys. Rev. B **67**, 104414 (2003).
[63] O. F. Syljuåsen and A. W. Sandvik, Phys. Rev. E **66**, 046701 (2002).
[64] Y.-D. Hsieh, Y.-J. Kao, and A. W. Sandvik, J. Stat. Mech. (2013) P09001.
[65] A. W. Sandvik, AIP Conf. Proc. **1297**, 135 (2010).
[66] D. J. Luitz, F. Alet, and N. Laflorencie, Phys. Rev. B **89**, 165106 (2014).
[67] K. Yosida, *Theory of Magnetism* (Springer-Verlag, Berlin, Heidelberg, 1996).
[68] https://wrap.warwick.ac.uk/139957.
[69] B. Bauer *et al.*, J. Stat. Mech. (2011) P05001.

# Supplemental Material: Field-tunable Berezinskii-Kosterlitz-Thouless correlations in a Heisenberg magnet

D. Opherden,[1] M. S. J. Tepaske,[2,3] F. Bärtl,[1,4] M. Weber,[3] M. M. Turnbull,[5] T. Lancaster,[6]
S. J. Blundell,[7] M. Baenitz,[8] J. Wosnitza,[1,4] C. P. Landee,[9] R. Moessner,[3] D. J. Luitz,[2,3] and H. Kühne[1]
[1] *Hochfeld-Magnetlabor Dresden (HLD-EMFL) and Würzburg-Dresden Cluster of Excellence ct.qmat, Helmholtz-Zentrum Dresden-Rossendorf, 01328 Dresden, Germany*
[2] *Physikalisches Institut, Universität Bonn, Nussallee 12, 53115 Bonn, Germany*
[3] *Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*
[4] *Institut für Festkörper- und Materialphysik, TU Dresden, 01062 Dresden, Germany*
[5] *Carlson School of Chemistry, Clark University, Worcester, MA 01610, USA*
[6] *Durham University, Centre for Materials Physics, South Road, Durham DH1 3LE, UK*
[7] *Clarendon Laboratory, Department of Physics, University of Oxford, Park Road, Oxford OX1 3PU, UK*
[8] *Max Planck Institute for Chemical Physics of Solids, 01187 Dresden, Germany*
[9] *Department of Physics, Clark University, Worcester, MA 01610, USA*

(Dated: January 15, 2023)

### Staggered spin correlations probed by NMR and $\mu^+$SR

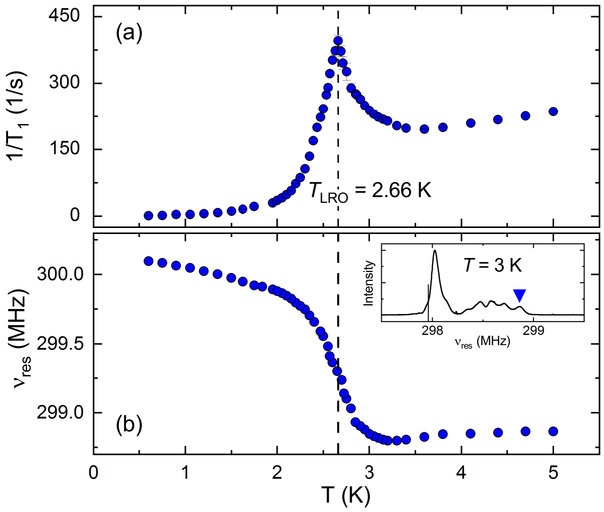


Figure S1. Temperature-dependent $^1$H-NMR (a) spin-lattice relaxation rate and (b) resonance frequency of CuPOF at 7 T. The vertical dashed line indicates the transition temperature $T_{\rm LRO} = 2.66$ K. A representative $^1$H-NMR spectrum at 3 K is shown in the inset of (b). The blue triangle marks the spectral line under investigation. The solid vertical line indicates the Larmor frequency $\nu_{\rm L} = 297.96$ MHz, given by the external field.

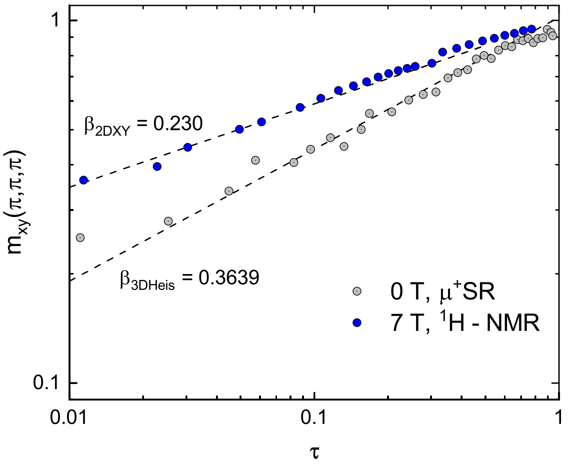


Figure S2. Log-log plot of the normalized $^1$H-NMR and $\mu^+$SR frequencies ($\mu^+$SR data are from Ref. [1]), probing the staggered magnetization $m_{xy}(\pi,\pi,\pi)$, and plotted as a function of the reduced temperature $\tau = 1 - T/T_c$, with $T_c = T_{\rm LRO}$. The dashed lines are plots of $m_{xy} \propto \tau^\beta$, where $\beta$ denotes an effective critical exponent.

In order to investigate the effect of the field-tunable $XY$ anisotropy on the quasi-static spin correlations, we probed the evolution of the staggered magnetization as an effective order parameter. As reported previously, the $^1$H-NMR spectra of CuPOF yield a distinct line splitting at low temperatures, which provides a direct probe of the local staggered magnetization [1]. More precisely, the NMR spectrum represents a histogram of the quasi-static fields, probed at the positions of the resonantly excited nuclear moments, on the time scale of the measuring process (a few ten $\mu$s here). The temperature-dependent resonance frequency $\nu_{\rm res}$ (determined as the first spectral moment) of a line at the high-frequency end of the $^1$H-NMR spectrum, recorded at 7 T, is presented in Fig. S1(b). The corresponding $^1$H nuclear spin-lattice relaxation rate is presented in Fig. S1(a), and yields $T_{\rm LRO} = 2.66$ K, identical to $T_{\rm LRO}$ as determined from the $^{31}$P spin-lattice relaxation rate, see Fig. 2(b) in the main text. A deviation of $\nu_{\rm res}$ from an almost constant value at high temperatures occurs at about 3 K $\simeq T_{\rm co} > T_{\rm LRO}$.

Whereas the staggered magnetization already becomes non-zero in the $XY$ regime, we define the maximum temperature of the $^1$H spin-lattice relaxation rate as critical temperature, see Fig. S1(a), as supported by the

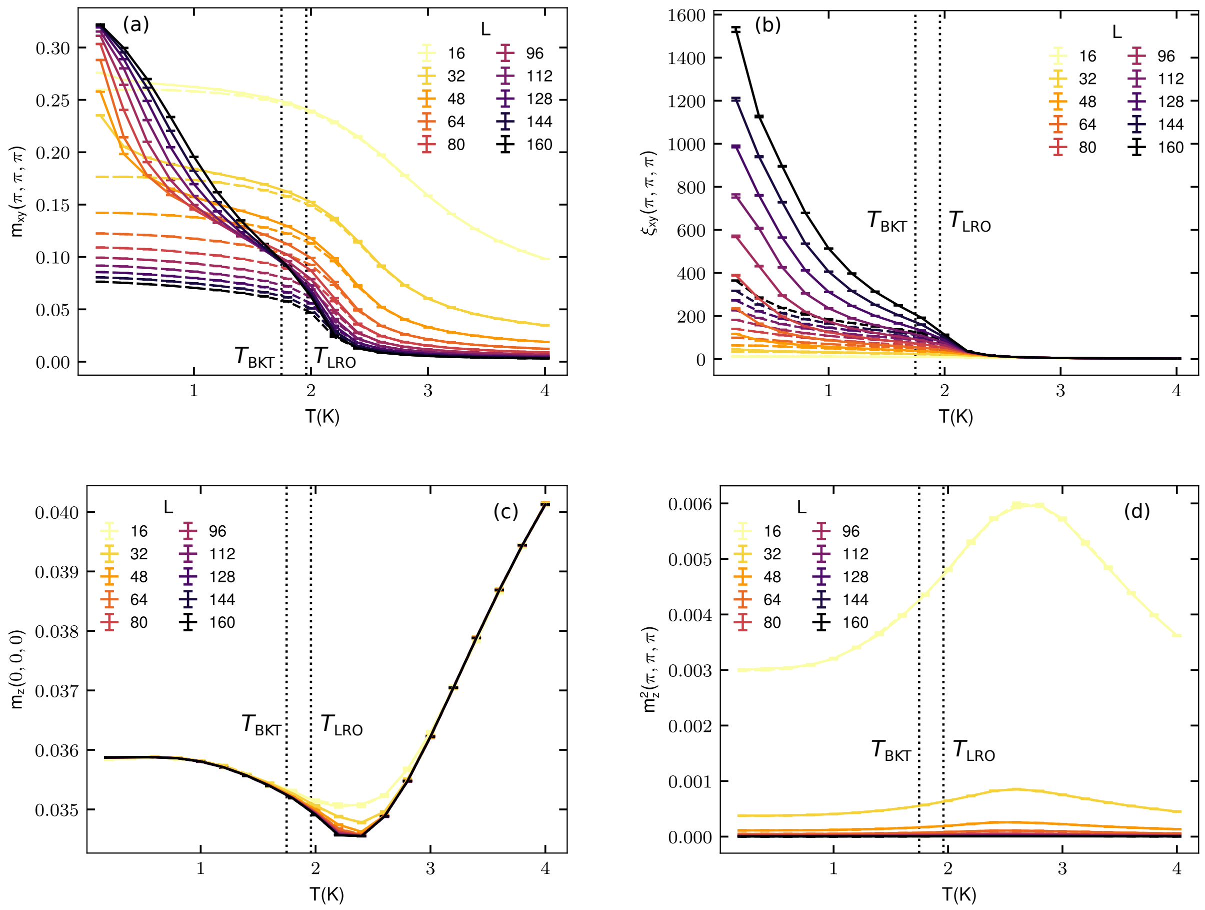


Figure S3. (a) The staggered in-plane magnetization $m_{xy}(\pi,\pi,\pi)$, (b) the corresponding correlation length $\xi_{xy}(\pi,\pi,\pi)$, (c) the uniform out-of-plane magnetization $m_z(0,0,0)$, and (d) the squared staggered out-of-plane magnetization $m_z^2(\pi,\pi,\pi)$, all determined from the structure factor. The lattices have sizes $L\times L\times L/8$ and we simulated the Hamiltonian (1) in the main text with intralayer coupling $J/k_{\rm B}=6.8$ K, interlayer coupling $J'/k_{\rm B}=1$ mK, intrinsic exchange anisotropy $\Delta=0.0185$, and magnetic field strength $\mu_0 H=2$ T. The solid lines are for the $J'=1$ mK model and the dashed lines for $J'=0$. The vertical dashed lines denote the critical temperatures $T_{\rm BKT}$ and $T_{\rm LRO}$ as determined in Fig. 3 in the main text.

following reasoning. In a layered anisotropic magnetic lattice, the correlation length significantly increases with decreasing temperature, following an exponential growth described as $\xi_{\rm 2DXY}\propto\exp\left(0.5\pi/\sqrt{T/T_{\rm BKT}-1}\right)$ [2, 3]. In the presence of a finite interlayer coupling $J'$, the transition to long-range order is expected at $\xi^2 J'/J\simeq 1$ [4]. According to our QMC simulations, at $T_{\rm LRO}$, the in-plane correlation length is of the order of 100 lattice spacings, as can be seen for $\mu_0 H=2$ T in Fig. S3(b). With $J'/J\simeq 1.4\times10^{-4}$ for CuPOF [1], the condition $\xi^2 J'/J\simeq 1$ is satisfied at $T_{\rm LRO}$.

Thus, the $1/T_1$ maximum precisely represents $T_{\rm LRO}$, whereas the previously reported splitting of the $^1$H spectral lines, stemming from the evolution of a finite staggered magnetization, occurs at a slightly higher temperature [1]. Further, although a small Dzyaloshinskii-Moriya (DM) interaction with a vector parallel to the crystallographic $b$ axis is not strictly excluded from symmetry considerations, our ESR data place an upper limit of 50 GHz on the presence of an excitation gap [1].

Closely below $T_c$, the staggered magnetization $m_{xy}(\pi,\pi,\pi)$ scales with the reduced temperature $\tau=(1-T/T_c)$ as $m_{xy}\propto\tau^\beta$, where $\beta$ may be interpreted as an effective critical exponent. Here, we study to what extent the expected universality classes of an (an)isotropic spin system in (non-)zero field are reflected in the experimental data, but note that there may in practise be a broad crossover region "interpolating" between the two (Heisenberg and XY) cases [5].

Employing $T_c=T_{\rm LRO}=2.66$ K at 7 T, we plot the normalized $^1$H resonance frequency as a function of the reduced temperature in a log-log plot, see Fig. S2. We find a good agreement when comparing with the critical exponent $\beta_{\rm 2DXY}=3\pi^2/128\simeq 0.23$ of a finite-size

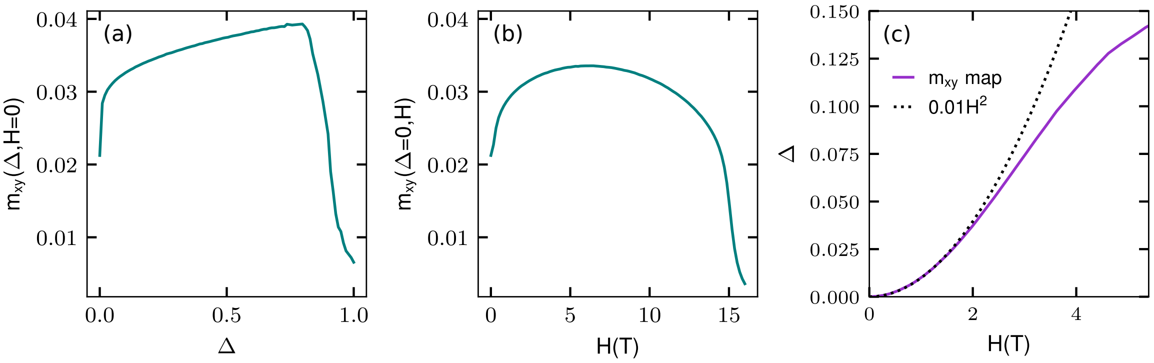


Figure S4. (a) The staggered in-plane magnetizations $m_{xy}(\Delta, H = 0)$ and (b) $m_{xy}(\Delta = 0, H)$ for an $L \times L = 100 \times 100$ Heisenberg system at $T = 1.2$ K. In (c), we show the numerically determined field-induced exchange anisotropy $\Delta(H)$, that was determined for $\mu_0 H \leq 6$ T by solving $m_{xy}(\Delta, H = 0) = m_{xy}(\Delta = 0, H)$. The black dotted line shows the quadratic dependence with $\Delta = 0.01(\mu_0 H)^2$.

2D $XY$ model [6, 7]. Similar observations were made for other materials that realize a planar $XY$ lattice [8–10]. The same analysis was applied to the $\mu^+$SR frequency at zero field [1], using $T_c = 1.38(2)$ K, giving a good agreement when comparing to the critical exponent $\beta_{\text{3DHeis}} = 0.3639(35)$ [11] of the 3D Heisenberg model, and, similarly well, to the critical exponent $\beta_{\text{3DXY}} = 0.33$ [4] of the 3D $XY$ model. These observations further support the scenario of the enhanced anisotropy of intralayer spin correlations at elevated fields.

## Staggered spin correlations calculated with QMC

To infer the pattern of the long-range order that is associated to $T_{\text{LRO}}$, which was calculated with QMC by examining the $L\rho$ crossings in the lower panel of Fig. 3 in the main text, we also calculate the in-plane structure factor $S_{xy}$ and the corresponding correlation length $\xi$ [12], i.e.,

$$S_{xy}(\vec{k}) = \sum_j e^{i\vec{k}\cdot\vec{r}_j} \left(\langle S_0^x S_j^x\rangle + \langle S_0^y S_j^y\rangle\right), \tag{1}$$

$$\xi_{xy}(\vec{k}) = \frac{L}{2\pi}\sqrt{\frac{S_{xy}(\vec{k})}{S_{xy}(\vec{k}+\mathrm{d}\vec{k})} - 1}\,. \tag{2}$$

Here, we introduced a staggering phase based on the lattice position $\vec{r}_j$ of site $j$ and a staggering vector $\vec{k}$ with nearest-by vector $\vec{k} + \mathrm{d}\vec{k}$. This structure factor can be used to define a magnetization via $m_{xy}^2 = S_{xy}(\vec{k})/N$, where $N$ denotes the amount of spins. In Fig. S3(a), we show the staggered in-plane magnetization $m_{xy}(\pi,\pi,\pi)$ at 2 T as a function of temperature for various values of $L$, and in Fig. S3(b) we show the corresponding correlation length $\xi_{xy}(\pi,\pi,\pi)$. The $J' = 1$ mK results are shown as solid lines and the $J' = 0$ results as dashed lines. We see an onset of the magnetization and in-plane spin-correlations at $T_{\text{LRO}}$ that does not scale to zero with system size when $J'$ is nonzero.

In Fig. S3(c), we show the uniform out-of-plane magnetization $m_z(0,0,0)$, which depicts the field dependency of the out-of-plane canting of the in-plane antiferromagnetic order and converges with system size. To verify that the magnetic order is only in-plane staggered, we show the squared staggered out-of-plane magnetization $m_z(\pi,\pi,\pi)$ in Fig. S3(d), which clearly scales to zero for large system sizes.

## Field-induced exchange anisotropy

As a simple estimate of the field-induced exchange anisotropy, we compute the staggered in-plane magnetizations $m_{xy}(\Delta, H = 0)$ and $m_{xy}(\Delta = 0, H)$ for an $L \times L = 100 \times 100$ Heisenberg system at $T = 1.2$ K, and find $\Delta(H)$ such that $m_{xy}(\Delta, H = 0) = m_{xy}(\Delta = 0, H)$ is satisfied. This condition should hold if the Hamiltonians $\mathcal{H}(\Delta, H = 0)$ and $\mathcal{H}(\Delta = 0, H)$ can be mapped onto each other, thereby giving an estimate of the field-induced exchange anisotropy $\Delta(H)$. In Figs. S4(a) and S4(b), we show $m_{xy}(\Delta, H = 0)$ and $m_{xy}(\Delta = 0, H)$, and in Fig. S4(c) we show the numerically determined field-induced exchange anisotropy $\Delta(H)$ that was found at $H \leq 6$ T. To compare with the perturbative quadratic field dependence estimate from [13], we also plot the quadratic curve $\Delta = 0.01(\mu_0 H)^2$, showing excellent agreement at small $H$.

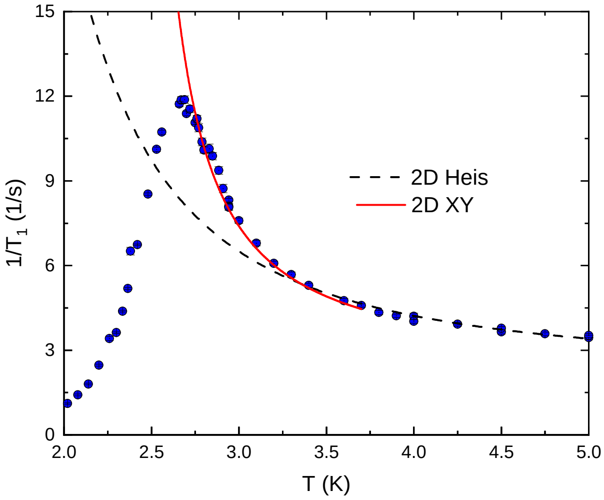


Figure S5. Temperature-dependent ${}^{31}$P NMR $1/T_1$ rate of CuPOF at 7 T. The dashed line indicates a fit with a 2D Heisenberg model between $T_{\mathrm{co}} \approx 3.2$ K and $J/k_{\mathrm{B}} = 6.8$ K. The red line indicates a fit with a 2D XY model between between $T_{\mathrm{LRO}}$ and $T_{\mathrm{co}}$.

## Crossover in growth of dynamic correlation length at $T_{\mathrm{co}}$

At temperatures above the $XY$ crossover regime, the dynamic spin correlations in CuPOF are isotropic, i.e., of 2D Heisenberg type. At around $T_{\mathrm{co}}$, a crossover to the low-temperature 2D $XY$ regime takes place. To probe this crossover, we described the ${}^{31}$P $1/T_1$ data of CuPOF, recorded at 7 T, between $T_{\mathrm{co}} \approx 3.2$ K and $J/k_{\mathrm{B}} = 6.8$ K with a 2D Heisenberg model, according to $\xi_{\mathrm{2DHeis}} \propto \exp\left[2\pi 0.178 J/\left(k_{\mathrm{B}} T\right)\right]$ [11, 14–17]. In contrast, we described the data between $T_{\mathrm{LRO}}$ and $T_{\mathrm{co}}$ with a 2D $XY$ model, according to $\xi_{\mathrm{2DXY}} \propto \exp\left(0.5\pi/\sqrt{T/T_{\mathrm{BKT}} - 1}\right)$ [2, 3]. Clearly, this comparison shows the formation of 2D $XY$ correlations below $T_{\mathrm{co}}$, manifested as an increased slope of the temperature-dependent $1/T_1$ rate as compared to that according to the 2D Heisenberg model.

---

[1] D. Opherden, N. Nizar, K. Richardson, J. C. Monroe, M. M. Turnbull, M. Polson, S. Vela, W. J. A. Blackmore, P. A. Goddard, J. Singleton, E. S. Choi, F. Xiao, R. C. Williams, T. Lancaster, F. L. Pratt, S. J. Blundell, Y. Skourski, M. Uhlarz, A. N. Ponomaryov, S. A. Zvyagin, J. Wosnitza, M. Baenitz, I. Heinmaa, R. Stern, H. Kühne, and C. P. Landee, Phys. Rev. B **102**, 064431 (2020).
[2] J. M. Kosterlitz, J. Phys. Condens. Matter **7**, 1046 (1974).
[3] H.-Q. Ding, Phys. Rev. Lett. **68**, 1927 (1992).
[4] J. Als-Nielsen, S. T. Bramwell, M. T. Hutchings, G. J. McIntyre, and D. Visser, J. Phys.: Condens. Matter **5**, 7871 (1993).
[5] S. V. Isakov and R. Moessner, Phys. Rev. B **68**, 104409 (2003).
[6] S. T. Bramwell and P. C. W. Holdsworth, J. Phys.: Condens. Matter **5**, L53 (1993).
[7] S. T. Bramwell and P. C. W. Holdsworth, Phys. Rev. B **49**, 8811 (1994).
[8] A. Taroni, S. T. Bramwell, and P. C. W. Holdsworth, J. Phys.: Condens. Matter **20**, 275233 (2008).
[9] P. Carretta, M. Filibian, R. Nath, C. Geibel, and P. J. C. King, Phys. Rev. B **79**, 224432 (2009).
[10] E. S. Klyushina, J. Reuther, L. Weber, A. T. M. N. Islam, J. S. Lord, B. Klemke, M. Månsson, S. Wessel, and B. Lake, Phys. Rev. B **104**, 064402 (2021).
[11] M. Troyer, M. Imada, and K. Ueda, J. Phys. Soc. Jpn. **66**, 2957 (1997).
[12] A. W. Sandvik, Aip Conference Proceedings **1297**, 135 (2010).
[13] P. Sengupta, C. D. Batista, R. D. McDonald, S. Cox, J. Singleton, L. Huang, T. P. Papageorgiou, O. Ignatchik, T. Herrmannsdörfer, J. L. Manson, J. A. Schlueter, K. A. Funk, and J. Wosnitza, Phys. Rev. B **79**, 060409(R) (2009).
[14] J.-K. Kim and M. Troyer, Phys. Rev. Lett. **80**, 2705 (1998).
[15] Troyer, M. (2001). Universality in Two-Dimensional Quantum Heisenberg Antiferromagnets. In: J. Bonča, P. Prelovšek, A. Ramšak, and S. Sarkar (eds) Open Problems in Strongly Correlated Electron Systems. NATO Science Series, vol 15. (Springer, Dordrecht).
[16] P. Hasenfratz and F. Niedermayer, Phys. Lett. B **268**, 231 (1991).
[17] H.-Q. Ding and M. S. Makivić, Phys. Rev. Lett. **64**, 1449 (1990).

# APPENDIX B

# Article 2

Low Temperature Physics ARTICLE scitation.org/journal/ltp

# Berezinskii–Kosterlitz–Thouless correlations in copper-based quasi-2D spin systems (Review Article)

Cite as: Fiz. Nizk. Temp. **49**, 899–907 (July 2023); doi: 10.1063/10.0019692
Submitted: 29 May 2023

View Online Export Citation CrossMark

D. Opherden,[1] F. Bärtl,[1,2] M. S. J. Tepaske,[3] C. P. Landee,[4] J. Wosnitza,[1,2] and H. Kühne[1,a)]

## AFFILIATIONS

[1]Hochfeld-Magnetlabor Dresden (HLD-EMFL) and Würzburg-Dresden Cluster of Excellence ct.qmat, Helmholtz-Zentrum Dresden-Rossendorf, Dresden 01328, Germany
[2]Institut für Festkörper- und Materialphysik, TU Dresden, Dresden 01062, Germany
[3]Physikalisches Institut, Universität Bonn, Bonn 53115, Germany
[4]Department of Physics, Clark University, Worcester, Massachusetts 01610, USA

[a)]**Author to whom correspondence should be addressed:** h.kuehne@hzdr.de

## ABSTRACT

We present an overview of selected copper-based quasi-2D square-lattice spin-1/2 materials with an easy-plane anisotropy, providing the possibility to study emergent Berezinskii-Kosterlitz-Thouless (BKT) correlations. In particular, in those materials with a comparatively small exchange coupling, the effective *XY* anisotropy of the low-temperature spin correlations can be controlled by an applied magnetic field, yielding a systematic evolution of the BKT correlations. In cases where the residual interlayer correlations are small enough, dynamical BKT correlations in the critical regime may be observed experimentally, whereas the completion of the genuine BKT transition is preempted by the onset of long-range order.

13 July 2023 07:58:09

*Published under an exclusive license by AIP Publishing.* https://doi.org/10.1063/10.0019692

## 1. INTRODUCTION

Due to its importance as a fundamental model system in quantum magnetism, the two-dimensional quantum Heisenberg spin-1/2 antiferromagnet (2DQHAF) on a square lattice and its ground-state properties, as well as the influence of small perturbations and magnetic fields on the spin correlations, were investigated in numerous works over the past decades. In the presence of an *XY* anisotropy, the occurrence of the topological Berezinskii-Kosterlitz-Thouless phase transition at a finite temperature $T_{BKT}$,[1–3] which marks the binding of topological defects in vortex-antivortex pairs, is a phenomenon of particular interest. Whereas the BKT transition is often studied for the spin-1/2 case, it occurs also in the classical limit.[4,5] In the reported experimental studies, the investigation of a genuine BKT transition in bulk materials was compromised by the onset of long-range order (LRO).[6–12] Typically, a residual interlayer coupling $J'$ stabilizes LRO at temperatures above $T_{BKT}$, thus preventing the manifestation of the BKT transition in material realizations of the 2D *XY* model. Still, if the perturbations relative to the 2D *XY* model are small enough, the spin system may yield experimental signatures of BKT-type correlations that develop at temperatures approaching $T_{LRO}$.[12–19]

The 2D spin-1/2 Heisenberg square-lattice antiferromagnet with weak easy-plane anisotropy in an applied magnetic field can be described by the Hamiltonian

$$\mathcal{H} = J\sum_{\langle i,j\rangle_{\parallel}}[S_i^xS_j^x + S_i^yS_j^y + (1-\Delta)S_i^zS_j^z] + J'\sum_{\langle i,j\rangle_{\perp}}\mathbf{S}_i\cdot\mathbf{S}_j - g\mu_B\mu_0H\sum_i S_i^z, \qquad (1)$$

where $\langle i,j\rangle_{\parallel}$ and $\langle i,j\rangle_{\perp}$ denote the intra- and interlayer nearest neighbors, and $J$ and $J'$ are the intra- and interlayer exchange couplings, respectively. Whereas $\Delta = 0$ corresponds to the isotropic Heisenberg case, $0 < \Delta \leq 1$ denotes a nonzero *XY* anisotropy.

For a small exchange coupling $J$ of a few K, the application of experimentally available magnetic fields of several T offers the possibility to continuously tune the low-temperature spin correlations from the 2D Heisenberg to the 2D *XY* limit.[12,20–26] As was shown by numerical and analytical calculations, a uniform magnetic field breaks the $O(3)$ symmetry of the 2D quantum Heisenberg antiferromagnet, but preserves the easy-plane $O(2)$ symmetry.[21] Correspondingly, for Zeeman energies of the order of the exchange energy, the resulting effective $XY$ anisotropy can be controlled. The associated BKT transition persists for all fields below the saturation field, yielding a nonmonotonic magnetic phase diagram.[21] This provides an excellent possibility for experimental studies of BKT physics and their comparison to theoretical predictions.

## 2. CLASSIFICATION OF QUASI-2D MATERIALS

In contrast to an ideal 2D Heisenberg spin system, the magnetic layers in bulk materials yield a nonzero interlayer exchange $J'$, and typically also a weak intrinsic anisotropy $\Delta_{\text{int}}$, which is described by the anisotropy parameter $\Delta$ in Eq. (1). Thus, in order to characterize how well the spin system in a given material can be approximated by the 2D Heisenberg model, the parameters $J'$ and $\Delta_{\text{int}}$ need to be determined. Nonzero values of $J'$ and $\Delta_{\text{int}}$ lead to long-range order at a critical temperature $T_{LRO}$, in contrast to the ideal 2D spin-1/2 Heisenberg case, which does not reveal long-range order at finite temperatures.[27] Therefore, a first qualification of a material as a realization of the 2D spin-1/2 Heisenberg antiferromagnet may be defined by the ratio $k_B T_{LRO}/J$, which varies between zero and values approaching unity for the 2D and the 3D spin-1/2 Heisenberg cases, respectively.[28]

In order to estimate the interlayer interaction $J'$, an empirical form, proposed by Yasuda *et al.*, is often used:[29]

$$k_B T_{LRO} = \frac{4\pi\rho_s}{2.43 - \ln(J'/J)}, \tag{2}$$

where $\rho_s = 0.183\,J$ is the renormalized spin-stiffness. Since this estimate assumes $\Delta_{\text{int}} = 0$, it represents only an upper boundary of $J'$ for weakly anisotropic materials. Both, the interlayer coupling $J'$ as well as the easy-plane anisotropy $\Delta_{\text{int}}$, may drive long-range order at nonzero temperatures. Therefore, the ratio $J'/J$, calculated for quasi-2DQHAFs with a finite $\Delta_{\text{int}}$ by means of Eq. (2), represents an upper limit of the interlayer coupling.

For a nonzero $\Delta_{\text{int}}$ and $J' = 0$, quantum Monte Carlo (QMC) calculations showed that, even for anisotropies as small as $10^{-3}$, the critical behavior of the magnetic lattice resembles that of the Berezinskii-Kosterlitz-Thouless universality class. A weak logarithmic decrease of the $T_{BKT}$ temperature with reduction of the spin

TABLE I. Selected $Cu^{2+}$-based quasi-2D spin-1/2 Heisenberg square-lattice antiferromagnets.[24,32–41] The exchange interaction $J$, the ordering temperature $T_{LRO}$, the $g$-factor (where available, both the in-plane and the out-of-plane components are listed), the anisotropy field $H_A$, as well as the saturation field $H_{\text{sat}}$ are presented. The ratio $J'/J$ of the inter- to intralayer coupling is estimated using Eq. (2). The anisotropy parameter $\Delta_{\text{int}}$ is calculated from the dc susceptibility minimum at $T_{\text{co}}$ using Eq. (5). Indirect estimates of $\Delta_{\text{int}}$ from (*) ESR,[36] (**) specific heat,[24] (†) and renormalization of the spin-wave dispersion analysis[12] are shown as well.

| | $J/k_B$, K | $T_{LRO}$, K | $k_B T_{LRO}/J$ | $J'/J$ | $g$-factor | $H_A$, T | $H_{sat}$, T | $H_A/H_{sat}$ | $\Delta_{\text{int}}$ |
|---|---|---|---|---|---|---|---|---|---|
| $Cu(pz)_2(ClO_4)_2$ | 18.1 | 4.21 | 0.232 | $5.6 \cdot 10^{-4}$ | 2.25<br>2.04 | 0.28 | 51.1 | $5.5 \cdot 10^{-3}$ | $4.6 \cdot 10^{-3}$ |
| $Cu(pz)_2(BF_4)_2$ | 15.3 | 3.8 | 0.248 | $1.1 \cdot 10^{-3}$ | … | 0.25 | 43 | $5.8 \cdot 10^{-3}$ | $6.2 \cdot 10^{-3}$ |
| $Cu(pz)_2(ReO_4)_2$ | 15.1 | 4.2 | 0.278 | $2.9 \cdot 10^{-3}$ | 2.13 | … | 42.7 | … | … |
| $Cu(pz)_2(H_2O)_2Cr_2O_7$ | 4.7 | $< 1.6$ | $< 0.34$ | $\lesssim 1.3 \cdot 10^{-2}$ | 2.13 | … | 13.3 | … | … |
| $[Cu(pz)_2(NO_3)](PF_6)$ | 10.8 | 3.05 | 0.282 | $3.3 \cdot 10^{-3}$ | … | $7 \cdot 10^{-3}$ | 30 | $2.3 \cdot 10^{-4}$ | $1.2 \cdot 10^{-2}$ |
| $[Cu(pz)_2(HF_2)](BF_4)$ | 6.3 | 1.54 | 0.244 | $9.1 \cdot 10^{-4}$ | 2.13 | … | 18.0 | … | … |
| $[Cu(pz)_2(HF_2)](ClO_4)$ | 7.2 | 1.91 | 0.265 | $1.9 \cdot 10^{-3}$ | 2.26<br>2.07 | 0.08 | 20.2 | $4 \cdot 10^{-3}$ | … |
| $[Cu(pz)_2(HF_2)](PF_6)$ | 12.8 | 4.38 | 0.342 | $1.4 \cdot 10^{-2}$ | 2.11 | … | 35.5 | … | $3 \cdot 10^{-3}$ (*) |
| $[Cu(pz)_2(HF_2)](AsF_6)$ | 12.8 | 4.34 | 0.339 | $1.3 \cdot 10^{-2}$ | 2.13 | … | 36.1 | … | … |
| $[Cu(pz)_2(HF_2)](SbF_6)$ | 13.3 | 4.31 | 0.324 | $9.4 \cdot 10^{-3}$ | 2.14 | … | 37.6 | … | … |
| $[Cu(pz)_2(4\text{-}phpy\text{-}O)_2](ClO_4)$ | 7.5 | 1.63 | 0.217 | $2.8 \cdot 10^{-4}$ | 2.26<br>2.04 | 0.11 | 21.1 | $5.2 \cdot 10^{-3}$ | … |
| $[Cu(pz)_2(pyO)_2](ClO_4)$ | 7.7 | 1.70 | 0.220 | $3.3 \cdot 10^{-4}$ | 2.26<br>2.04 | 0.11 | 21.9 | $5 \cdot 10^{-3}$ | … |
| $[Cu(pz)_2(pyO)_2](PF_6)_2$ | 8.1 | 1.71 | 0.211 | $2.1 \cdot 10^{-4}$ | 2.25<br>2.05 | … | 23.7 | … | $7 \cdot 10^{-3}$ (**) |
| **$[Cu(pz)_2(2\text{-}HOpy)_2](PF_6)_2$ (= CuPOF)** | **6.8** | **1.38(2)** | **0.203** | **$1.4 \cdot 10^{-4}$** | **2.29<br>2.07** | **0.36 (4)** | **17.57<br>19.5** | **$1.85 \cdot 10^{-2}$** | **$0.9 \cdot 10^{-2}$** |
| $Sr_2CuO_2Cl_2$ | ∼ 1450 | 255 | 0.176 | $2.4 \cdot 10^{-5}$ | … | 0.7 | ∼ 4000 | $1.8 \cdot 10^{-4}$ | $1.4 \cdot 10^{-3}$(†) |
| $La_2CuO_4$ | ∼ 1600 | 320 | 0.200 | $1.2 \cdot 10^{-4}$ | … | … | ∼ 4500 | … | $2 \cdot 10^{-4}$ |

13 July 2023 07:58:09

Published under an exclusive license by AIP Publishing

anisotropy was determined as[30]

$$k_B T_{BKT} = \frac{4\pi\rho_s}{\ln(330/\Delta)}, \qquad (3)$$

with the spin stiffness $\rho_s = 0.177\,J$.

In Table 1, we present a list of several Cu-based materials and their experimentally determined parameters, i.e., the intralayer coupling $J$, the zero-field ordering temperature $T_{LRO}$, the ratio $J'/J$, the electronic $g$-factor, the anisotropy field $H_A$ (corresponding to a spin-flop anomaly), the saturation field $H_{\rm sat}$, and the estimated easy-plane anisotropy $\Delta_{\rm int}$.

Metal-organic materials, in which the magnetic $Cu^{2+}$ ions are embedded into a matrix of organic molecules, represent a pathway for realizing a 2DQHAF on a square lattice with small exchange coupling $J/k_B$ of several K. By an appropriate choice of molecular ligands and counterions, the syntheses of several such materials were reported.[24,31–39,42] Often, pyrazine (pz = $C_4H_4N_2$) molecules are used as ligands to link the $Cu^{2+}$ ions in the magnetic quasi-2D layers. The canting of the pyrazine rings[43–45] and the choice of the counterions[46] determine the strength of the exchange interaction. For several $Cu^{2+}$-based molecular materials, a magnetic phase diagram in applied magnetic fields was reported.[24,31,36,47,48] Therein, the magnetic properties were mostly investigated by thermodynamic methods,[24,32,36,39,42] thus missing local information about the magnetic correlations at low temperatures.

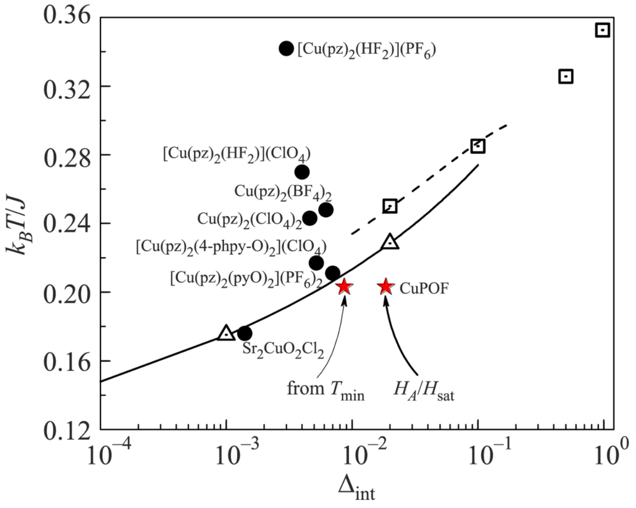


**FIG. 1.** Characteristic temperatures versus the anisotropy $\Delta_{\rm int}$. The open symbols denote QMC calculations of $k_BT_{BKT}/J$ for weakly-anisotropic 2DQHAFs from Ref. 10 (open squares) and Ref. 30 (open triangles). Solid and dashed lines denote fits to the QMC data by empirical formulas from Refs. 30 and 10, respectively. Full circles denote $k_BT_{LRO}/J$ of several quasi-2DQHAF materials with weak $XY$ anisotropy $\Delta_{\rm int}$.[34,39,50] The full red stars represent $k_BT_{LRO}/J$ of the material CuPOF versus estimates of the anisotropy parameter from the dc susceptibility and the anisotropy field, respectively.[31]

As mentioned above, a 2D square-lattice spin-1/2 system with $XY$ anisotropy exhibits a topological transition at a finite temperature $T_{BKT}$.[10,30] QMC results for the dependence of this BKT transition temperature on the anisotropy parameter $\Delta_{\rm int}$ are presented as open symbols and solid or dashed lines in Fig. 1.[10,30] The black solid circles denote the experimentally obtained values of $k_BT_{LRO}/J$ versus $\Delta_{\rm int}$ for selected quasi-2D weakly anisotropic QHAFs.[34,39,50] Where possible, $\Delta_{\rm int}$ was determined by Eq. (5) from a characteristic minimum of the dc susceptibility at $T_{\rm min} = T_{\rm co}$, compare Table 1. For $Cu(pz)_2(ClO_4)_2$ and $Cu(pz)_2(BF_4)_2$, very similar values of $\Delta_{\rm int}$ were found from measurements of the field- and temperature-dependent magnetization. For $[Cu(pz)_2(HF_2)](ClO_4)$, the experimentally determined value of $\Delta_{\rm int} = H_A/H_{\rm sat}$ was used. $[Cu(pz)_2(4\text{–}phpy\text{–}O)_2](ClO_4)$ and $[Cu(pz)_2(pyO)_2](ClO_4)$ are very similar in their properties and composition. Hence, only the slightly better isolated second material is presented in Fig. 1. For $[Cu(pz)_2(HF_2)](PF_6)$ and $[Cu(pz)_2(pyO)_2](PF_6)_2$, an indirect estimate of $\Delta_{\rm int}$ using ESR and specific-heat measurements from Refs. 24 and 36 was used.

For the inorganic compound $Sr_2CuO_2Cl_2$, which is a well-known realization of the 2DQHAF model, the anisotropy parameter, evaluated from the spin-dispersion analysis,[12,51] is an order of magnitude larger than the estimate $\Delta_{\rm int} = H_A/H_{\rm sat} = 1.8\cdot10^{-4}$. However, the latter estimate of $\Delta_{\rm int}$ suffers from the uncertainty of $H_{sat}$. From the dc-susceptibility minimum at $T_{co} \simeq 320$ K, reported in Ref. 50, $\Delta_{\rm int} = 9.9\cdot10^{-4}$ was found by use of Eq. (5). Further, $Sr_2CuO_2Cl_2$ hosts extremely well-isolated magnetic layers, with a ratio $k_BT_{LRO}/J = 0.176$, from which $J'/J = 2.4\cdot10^{-5}$ can be estimated.[13,34,49,50] However, the antiferromagnetic intralayer coupling of 1450 K yields extremely large saturation fields, which prohibits experimental studies of field-induced effects on the spin correlations.

In this context, the material $[Cu(pz)_2(2\text{–}HOpy)_2](PF_6)_2$ (CuPOF in the following) is of particular interest. The values of $J$, $T_{LRO}$, $J'/J$, the $g$-factor, the saturation field $H_{\rm sat}$, and the anisotropy field $H_{sat}$, as well as the anisotropy parameter $\Delta_{\rm int}$ were determined by various experimental probes.[31,52,53] When comparing CuPOF with the other materials listed in Table 1, it can be characterized as an excellent realization of a quasi-2DQHAF with a small nearest-neighbor interaction of $J/k_B = 6.8$ K.

## 3. CHARACTERISTIC TEMPERATURES VS *XY* ANISOTROPY

As shown in Fig. 1, for selected quasi-2D materials with a very weak interlayer interaction, such as $Sr_2CuO_2Cl_2$ and $[Cu(pz)_2(pyO)_2](PF_6)_2$, the experimentally determined values of $k_BT_{LRO}/J$ are very close to those of the QMC calculations of $k_BT_{BKT}/J$ for a weakly anisotropic 2DQHAF.

This indicates that the critical spin correlations of the BKT transition and those underlying the formation of long-range order are closely related in these materials. Following Refs. 12 and 30, a comparison of the experimentally observed spin-anisotropy crossover temperature $T_{\rm co}$ and the LRO transition temperature $T_{LRO}$ with QMC calculations of $T_{\rm co}$ and $T_{BKT}$ for a weakly anisotropic square-lattice 2DQHAF is presented in Fig. 2 for $Sr_2CuO_2Cl_2$ and CuPOF. For both materials, excellent agreement is observed for the

13 July 2023 07:58:09

Published under an exclusive license by AIP Publishing

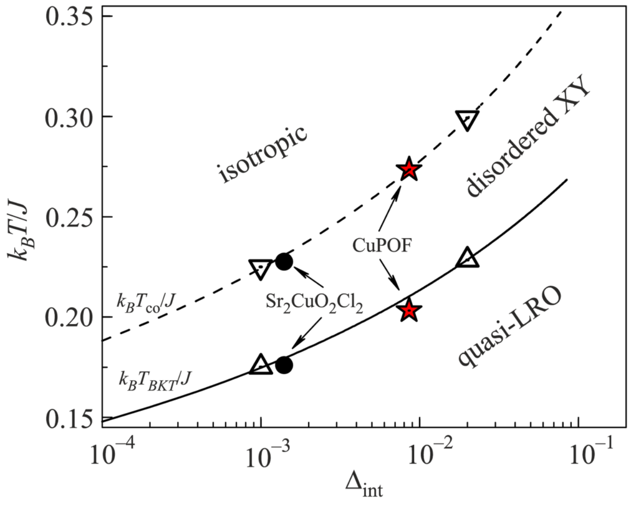


**FIG. 2.** Phase diagram for weakly anisotropic 2D spin-1/2 square-lattice Heisenberg antiferromagnets from Refs. 12 and 30. The calculated BKT transition and spin-anisotropy crossover temperatures $k_BT_{BKT}/J$ and $k_BT_{\rm co}/J$, respectively, are presented as open up and down triangles. The temperatures $k_BT_{\rm co}/J$ and $k_BT_{LRO}/J$ for $Sr_2CuO_2Cl_2$ and CuPOF are denoted by black circles and red stars, respectively. The dashed and solid lines are plots according to Eqs. (5) and (3), respectively.

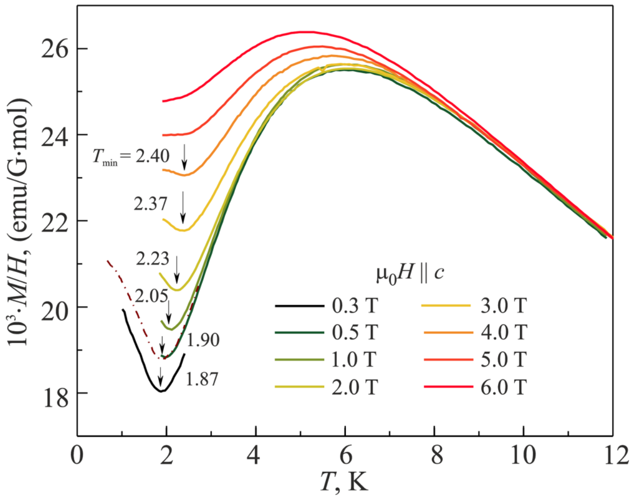


**FIG. 3.** Molar dc susceptibility of single-crystalline CuPOF at different magnetic fields applied perpendicular to the crystallographic planes.[31] The black downward arrows indicate the crossover temperature, as discussed in the text.

characteristic temperatures $T_{\rm co}$, $T_{BKT}$, and $T_{\rm LRO}$. The larger value of $k_BT_{LRO}/J$ = 0.203 for CuPOF, as compared to $k_BT_{LRO}/J = 0.176$ for $Sr_2CuO_2Cl_2$, is attributed only to the comparatively stronger intrinsic spin anisotropy in CuPOF. Therefore, the magnetic layers in CuPOF are concluded to be similarly well isolated as in $Sr_2CuO_2Cl_2$.

Due to the very weak coupling of the magnetic layers in CuPOF, with $J'/J \approx 1.4\cdot10^{-4}$, the small entropy change expected at the transition at $T_{LRO}$ is beyond the experimental resolution of thermodynamic probes.[31,54] On the other hand, $\mu^+$SR is very sensitive to the local staggered magnetization, and was used to probe the transition to LRO at 1.38(2) K in CuPOF.[31] This transition occurs under the influence of the weak intrinsic easy-plane anisotropy, which yields a temperature-driven crossover from isotropic to *XY* -type correlations at the crossover temperature $T_{\rm co} > T_{LRO}$. An applied magnetic field increases the effective *XY* anisotropy, which manifests itself as a field-dependent minimum of the uniform bulk susceptibility at $T_{\rm min} = T_{\rm co}$, as depicted in Fig. 3.

## 4. *XY* ANISOTROPY AND CHARACTERISTIC TEMPERATURES VS FIELD

A weak intrinsic anisotropy $\Delta_{\rm int}$ can be determined from, e.g., measurements of the temperature- and field-dependent magnetization.[12,30,34,39,50,56] A qualitatively different behavior of the field-dependent magnetization for magnetic fields applied parallel and perpendicular to the easy plane is expected at temperatures below $T_{\rm co}$. At the anisotropy field $H_A$, a step-like feature of the in-plane magnetization occurs. Accordingly, $H_A$ represents a measure of the spin-anisotropy and $\Delta_{\rm int}$ can be evaluated as[57]

$$\Delta_{\rm int} = \frac{2Sg\mu_BH_A}{zJ} = \frac{H_A}{H_{\rm sat}}, \tag{4}$$

where $z = 4$ is the coordination number for a square lattice. For CuPOF, $H_A$ = 0.36(1) $T$, so that $\Delta_{\rm int} = H_A/H_{\rm sat} = 1.85(5)\cdot10^{-2}$ is obtained as an estimate of the intrinsic easy-plane anisotropy.[31] The intrinsic anisotropy, caused by the combination of crystal electric field effects and residual spin-orbit coupling of the $Cu^{2+}$ ions, is comparable for all metal-organic compounds in Table 1, resulting in similar values of the anisotropy fields $H_A$ and components of the electronic *g*-factor.

As mentioned above, at fields above $H_A$, both the in- and out-of-plane dc susceptibility exhibit a broad minimum as a function of temperature. This anisotropic behavior can be understood in terms of the 2D Heisenberg model in the presence of a weak easy-plane anisotropy.[12,30,34] Cuccoli *et al.* reported an empirical formula for estimating the anisotropy parameter $\Delta_{\rm int}$ from the temperature of the minimum of the out-of-plane susceptibility at $T_{\rm min} = T_{\rm co}$:[12,30]

$$k_BT_{\rm co} = \frac{4\pi\rho_s}{\ln(160/\Delta)}, \tag{5}$$

with the spin stiffness $\rho_s = 0.214\,J$. For CuPOF, employing Eq. (5) with $J/k_B = 6.8$ K and $T_{\rm co}$ =1.86(5) K, we determine an anisotropy of $\Delta_{\rm int} = 0.9(2)\cdot10^{-2}$.

In applied magnetic fields, a weakly-anisotropic quasi-2D Heisenberg spin system is described by the Hamiltonian (1). From analytical arguments[23,58] and Monte Carlo simulations,[20] it was

13 July 2023 07:58:09

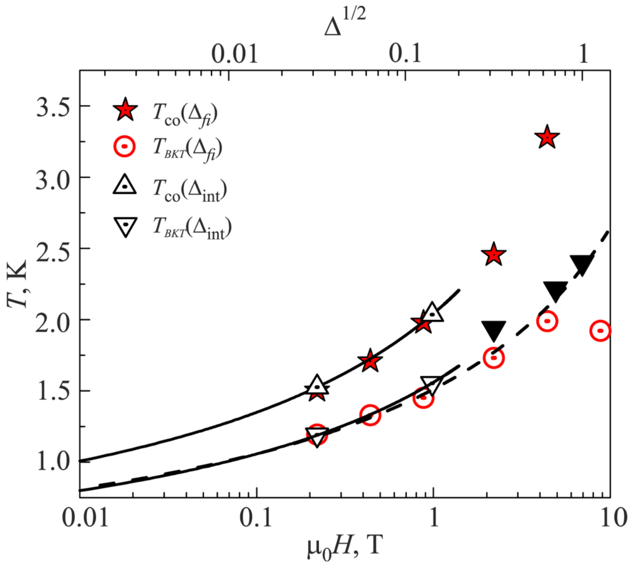


**FIG. 4.** QMC results of $T_{\rm co}$ and $T_{BKT}$ for a 2DQHAF with (i) a field-induced anisotropy $\Delta_{fi}$ and (ii) an intrinsic anisotropy $\Delta_{\rm int}$.[10,12,21,30,59] Red stars and circles refer to the field-dependent temperatures $T_{\rm co}$ and $T_{BKT}$ of the isotropic 2DQHAF.[21] Up and down triangles refer to $T_{\rm co}$ and $T_{BKT}$ of the intrinsically anisotropic 2DQHAF from Ref. 30 (open symbols) and Refs. 10 and 59 (solid symbols). The black solid lines denote the empirical expressions (5) and (3) for $T_{\rm co}$ and $T_{BKT}$, respectively, in the weakly-anisotropic regime.[12,30] The magnetic field and the anisotropy parameter $\Delta$ are shown on the bottom and top horizontal scale, respectively. The temperatures and magnetic fields are scaled with the parameters of CuPOF, i.e., $J/k_B$ =6.8 K and $H_{{\rm sat},c}$ = 17.57(5) T.

established that, in the low-field regime, the 2DQHAF with $J' = \Delta_{\rm int} = 0$ in uniform magnetic fields exhibits an $XY$-like phase below $T_{\rm co}$. Thus, the isotropic 2DQHAF in applied magnetic fields can be mapped to a 2DQHAF with a weak easy-plane anisotropy by defining an effective field-induced anisotropy parameter $\Delta_{fi}$. From comparing the QMC results of $T_{BKT}$ for both these cases of the 2DQHAF, Cuccoli *et al.*[21] proposed a field-dependent anisotropy of the form $\Delta_{fi} \simeq h^2$ for the low-field regime, where $h = g\mu_BH/(JS)$ is the reduced magnetic field, with $h = 8$ at magnetic saturation. A comparison of the resulting characteristic temperatures is presented in Fig. 4. Here, the field-dependent temperatures $T_{\rm co}$ and $T_{BKT}$ are presented as red stars and open circles, respectively. The black dashed line denotes the field dependence according to $t_{BKT}(h) \simeq (4\pi\rho_s/J)/\ln(C/h^2)$, where $\rho_s$ is the spin stiffness and $C$ a constant.[21]

Further, in Fig. 4, $T_{\rm co}$ and $T_{BKT}$ for the weakly anisotropic 2DQHAF are plotted versus $\Delta^{1/2}$, since $\Delta_{fi} \simeq h^2$. Here, the solid down triangles are the QMC results from Refs. 10, 59, the open up and down triangles are the results from Ref. 30. In the low-field regime, a very good agreement between the QMC calculations for both cases of the 2DQHAF is found for both $T_{\rm co}$ and $T_{BKT}$. Thus, the effect of an increasing magnetic field applied to the isotropic 2DQHAF yields very similar spin correlations as an increase of the intrinsic easy-plane anisotropy in zero field.

An isotropic 2DQHAF in applied magnetic fields can be described by a renormalized planar rotator model with progressively decreasing rotator length.[21] By means of a detailed finite-size scaling analysis, Cuccoli *et al.*[21] verified an $XY$-type anisotropy of the spin system with a subsequent BKT transition for all magnetic fields up to the saturation field, regardless of the gradually reduced planar spin projection. Thus, the competing mechanisms of the field-induced $XY$ anisotropy and the progressive spin canting in field direction govern the field-dependent evolution of $T_{BKT}$.

## 5. FIELD-DRIVEN EVOLUTION OF THE STAGGERED MAGNETIZATION

For a 2DQHAF in applied magnetic fields, upon cooling from the paramagnetic regime, Heisenberg-type spin correlations develop below $T \simeq J/k_B$, and cross over to a $XY$-type in the regime of $T_{co}$. With further decreasing temperature, the spin correlation length $\xi$ grows exponentially approaching the BKT transition at $T_{BKT}$. For $T > T_{BKT}$, a rather low density of vortices is expected.[60] Furthermore, whereas Skyrmion-type textures may develop between about $J/k_B$ and $T_{\rm co}$, the associated correlation length is rather short in this regime.[53] The exponential increase of $\xi$ yields a rapid strengthening of the antiferromagnetic correlations in the $XY$ regime and, therefore, the staggered magnetization becomes nonzero in a finite-size system. With further increase of $\xi$ upon lowering the temperature, the magnetic correlations, due to the influence of the small but nonzero interlayer interaction $J'$ on the regions with large in-plane correlation lengths, can no longer be treated as 2D. Therefore, a transition to long-range order occurs at $T_{LRO}$.

In order to investigate the effect of the field-tuned $XY$ anisotropy on the static spin correlations in CuPOF, we probed the evolution of the staggered magnetization. The linewidth of the $^{31}$P-NMR spectra, probing the local-field distribution at the $^{31}$P sites, provides a measure of the staggered magnetization $m_{xy}$ ($\pi$, $\pi$, $\pi$). The temperature-dependent $^{31}$P-NMR linewidth $\Delta\nu^*$. recorded at 2 and 7 T, is presented in Fig. 5(a). The linewidth is normalized by the low-temperature limit, $\Delta\nu \to 0$, as well as the $x$ axis by the onset temperature $T^*$ of a steeply increasing $\Delta\nu$. This onset occurs close to $T_{LRO}$, which is determined as the maximum of the temperature-dependent $^{31}$P $1/T_1$ rate at given field, compare Fig. 6. In a 2D $XY$ magnetic lattice, the correlation length increases exponentially as $\xi_{2DXY} \propto \exp\left(0.5\pi/\sqrt{T/T_{BKT}-1}\right)$.[2,10] In the presence of a finite interlayer coupling $J'$. the transition to long-range order is expected at $\xi^2J'/J \simeq 1$.[61] With an inplane correlation length of the order of 100 lattice spacings at $T_{LRO}$ and $J'/J \simeq 1.4 \cdot 10^{-4}$ for CuPOF, the condition $\xi^2J'/J \simeq 1$ is satisfied, and $T_{LRO}$ may be interpreted as critical temperature $T_c$.[53]

Closely below $T_c$, the staggered magnetization $m_{xy}$ ($\pi$, $\pi$, $\pi$) scales with the reduced temperature $\tau = (1-T/T_c)$ as $m_{xy} \propto \tau^\beta$,where $\beta$ may be interpreted as an effective critical exponent. Employing $T^* \simeq T_c = 2.25$ K at 2 T and $T^* \simeq T_c = 2.66$ K at 7 T, we plot the normalized $^{31}$P linewidth as a function of the reduced temperature in a log-log plot, see Fig. 5(b). For the 7 T data, we find good agreement with the critical exponent $\beta_{2DXY} = 3\pi^2/128 \simeq 0.23$ of a finite-size 2D $XY$ model.[62,63] Similar observations were made for other

13 July 2023 07:58:09

Published under an exclusive license by AIP Publishing

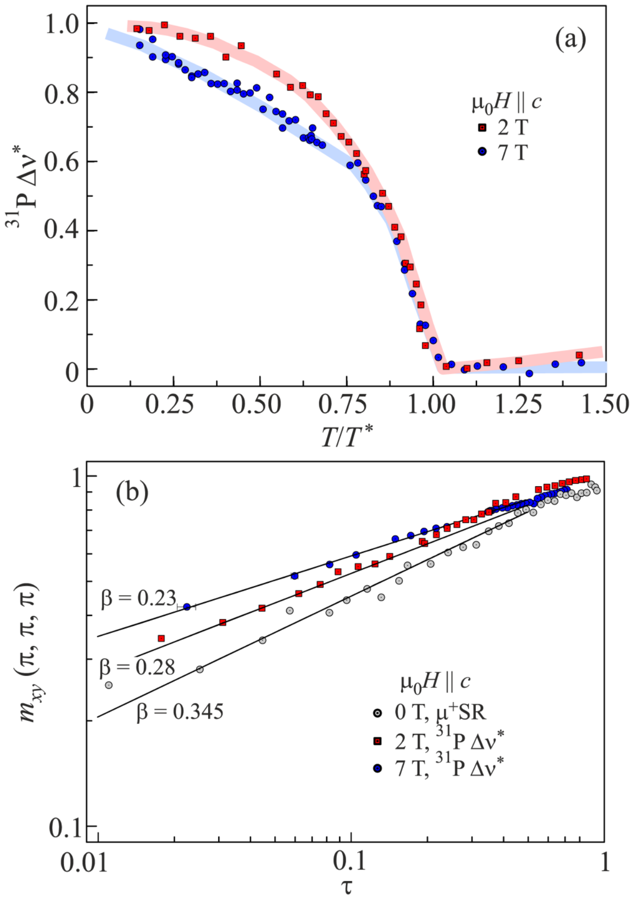


**FIG. 5.** (a) Normalized $^{31}$P-NMR linewidth Δν*, recorded at 2 and 7 T, plotted versus the reduced temperature $T/T^*$. (b) Double-logarithmic plot of the staggered magnetization $m_{xy}$ (π, π, π), plotted versus the reduced temperature $\tau = (1 - T/T_c)$. At zero field, $m_{xy}$ is probed by the normalized $\mu^+$SR precession frequency. At 2 and 7 T, $m_{xy}$ is probed by the normalized $^{31}$P-NMR resonance linewidth Δν*. The black solid lines denote a power-law behavior according to $m_{xy} \propto \tau^{\beta}$.

materials that realize a planar *XY* lattice.[17,19,64] The same analysis was applied to data of the $\mu^+$SR precession frequency at zero field,[31] using $T_c$ =1.38(2) K, giving β = 0.345, which is in good agreement with the theoretical critical exponent $\beta_{3\mathrm{DHeis}} = 0.3639(35)$[65] of the 3D Heisenberg model, and, similarly well, with the critical exponent $\beta_{3DXY} = 0.33$[61] of the 3D *XY* model. At 2 T, we find β = 0.28, which may be interpreted in terms of a crossover between the Heisenberg and *XY* cases.[66]

## 6. FIELD-DRIVEN EVOLUTION OF THE DYNAMIC CORRELATIONS

The temperature-dependent $^{31}$P nuclear spin-lattice relaxation rate $1/T_1$ for out-of-plane magnetic fields of 2, 7, and 16 T is presented in Fig. 6. At high temperatures, $1/T_1$ is almost temperature-independent, indicating predominantly paramagnetic fluctuations. At temperatures above the onset of LRO, $1/T_1$ probes the dynamic correlation length ξ.[15,16,67–70] As was shown from dynamical scaling arguments,[67] $1/T_1$ is proportional to the transverse spin correlation length as $1/T_1 \propto \xi^{z-\eta}$, where $z$ and η are characteristic dynamic and critical exponents.[15,25,67,71] In 2D magnetic lattices, the onset of short-range spin correlations occurs at temperatures $T \simeq J/k_B$,[54] with a correlation length of about one magnetic-lattice constant.[10,72] To probe the crossover at $T_{\mathrm{co}}$, we describe the $^{31}$P $1/T_1$ rate above $T_{\mathrm{co}} \approx 3.2$ K with a 2D Heisenberg model, according to $\xi_{2\mathrm{DHeis}} \propto \exp[2\pi 0.178 J/(k_B T)]$.[65,72–75] In contrast, we describe the data between $T_{LRO}$ and $T_{\mathrm{co}}$ with a 2D *XY* model, according to $\xi_{2DXY} \propto \exp\left(0.5\pi/\sqrt{T/T_{BKT}-1}\right)$.[2,10] Clearly, this comparison shows the formation of 2D *XY* correlations below $T_{\mathrm{co}}$ as a stronger increase of the temperature-dependent $1/T_1$ rate, compared to the increase according to the 2D Heisenberg model. A fit of $1/T_1$ in the interval $T_{LRO} \le T \le J/k_B$ with the 2D *XY* model yields $T_{BKT}$ = 1.708(14), 2.237(7), and 0.90(16) K for applied fields of 2, 7, and 16 T, respectively, with errors determined by bootstrapping.[53] The BKT transition itself is preempted by the LRO that arises from the 3D correlations, stemming from the finite interlayer exchange interaction $J'$. At $T_{LRO}$, a sharp maximum of $1/T_1$ is observed for the two lower fields of 2 and 7 T. The amplitude of the $1/T_1$ maximum at 16 T, closely below the saturation field of 17.5 T, is substantially reduced in comparison.

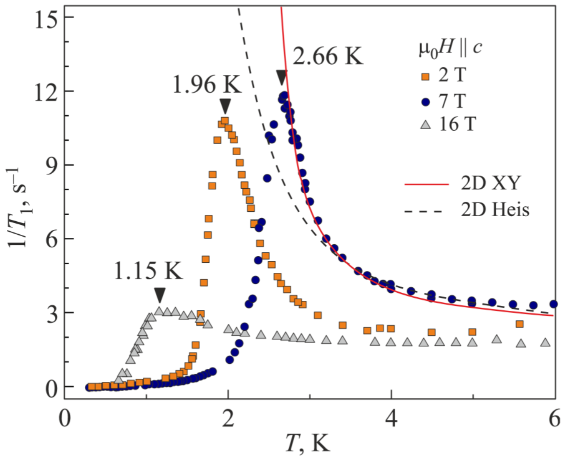


**FIG. 6.** Temperature-dependent $^{31}$P nuclear spin-lattice relaxation rate of CuPOF at 2, 7, and 16 T.[53] The downward triangles indicate the long-range ordering temperature $T_{LRO}$, determined as the rate maximum. For the data at 7 T, the dashed line indicates a fit with a 2D Heisenberg model between $T_{\mathrm{co}} \approx 3.2$ K and $J/k_B$ = 6.8 K. The red line indicates a fit with a 2D *XY* model between $T_{LRO}$ and $T_{\mathrm{co}}$.

With similar reasoning, for the case of $Sr_2CuO_2Cl_2$, it was argued that the LRO transition is induced by the incipient intralayer BKT transition at $T_{BKT} < T_{LRO}$.[10,12] Moreover, a spin-anisotropy crossover at $T_{\mathrm{co}} \sim 320\,\tilde{\mathrm{K}}$ was detected in measurements of the correlation length by means of neutron scattering[55]

13 July 2023 07:58:09

Published under an exclusive license by AIP Publishing

and NMR.[14] This is a very similar scenario of a spin-anisotropy crossover from isotropic Heisenberg correlations at $T > T_{co}$ to $XY$-like planar anisotropy at $T < T_{co}$ as we found for CuPOF.

## 7. SUMMARY

We presented an overview and classification of several copper-based quasi-2D spin-1/2 square-lattice materials. These provide the opportunity to study Berezinskii-Kosterlitz-Thouless correlations, which emerge from a nonzero $XY$ anisotropy. If the residual interlayer correlations are small enough, dynamic BKT correlations may be probed in the transition regime, whereas the completion of the genuine BKT transition is preempted by the onset of long-range order.

In particular, we discussed the material CuPOF as a model case for a 2DQHAF with small exchange coupling $J$, for which the application of a magnetic field allows a controlled tuning of the spin correlations from the almost isotropic 2D Heisenberg to the highly-anisotropic 2D $XY$ limit. As a consequence of the field-induced BKT-type spin correlations, a concomitant nonmonotonic behavior of the transition temperature $T_{LRO}$ is observed. The phenomenology in CuPOF is driven by field-induced Berezinskii-Kosterlitz-Thouless physics under the influence of extremely small interplane interactions, thus providing an opportunity for systematic investigations of BKT-type topological excitations.

## ACKNOWLEDGMENTS

We acknowledge the support of HLD at HZDR, member of the European Magnetic Field Laboratory (EMFL). This work was supported by the Deutsche Forschungsgemeinschaft (DFG) through SFB 1143 (project ID 247310070), the Würzburg-Dresden Cluster of Excellence on Complexity and Topology in Quantum Matter - *ct.qmat* (EXC 2147, project ID 390858490), and the cluster of excellence ML4Q (EXC 2004, project ID 390534769).

## REFERENCES

[1] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).
[2] J. M. Kosterlitz, J. Phys. C **7**, 1046 (1974).
[3] V. L. Berezinskii, J. Exp. Theor. Phys. **32**, 493 (1971).
[4] S. B. Khokhlachev, Zh. Exp. Teor. Fiz. **70**, 265 (1976).
[5] L. Capriotti, A. Cuccoli, V. Tognetti, R. Vaia, and P. Verrucchi, Physica D **119**, 68 (1998).
[6] B.-G. Liu, Phys. Rev. B **41**, 9563 (1990).
[7] N. Majlis, S. Selzer, and G. C. Strinati, Phys. Rev. B **45**, 7872 (1992).
[8] N. Majlis, S. Selzer, and G. C. Strinati, Phys. Rev. B **48**, 957 (1993).
[9] L. Siurakshina, D. Ihle, and R. Hayn, Phys. Rev. B **61**, 14601 (2000).
[10] H.-Q. Ding, Phys. Rev. Lett. **68**, 1927 (1992).
[11] H.-Q. Ding, J. Phys.: Condens. Matter **2**, 7979 (1990).
[12] A. Cuccoli, T. Roscilde, R. Vaia, and P. Verrucchi, Phys. Rev. Lett. **90**, 167205 (2003).
[13] M. Greven, R. J. Birgeneau, Y. Endoh, M. A. Kastner, B. Keimer, M. Matsuda, G. Shirane, and T. R. Thurston, Phys. Rev. Lett. **72**, 1096 (1994).
[14] B. J. Suh, F. Borsa, L. L. Miller, D. C. Johnston, D. R. Torgeson, and M. Corti, J. Appl. Phys. **79**, 5084 (1996).
[15] B. J. Suh, F. Borsa, L. L. Miller, M. Corti, D. C. Johnston, and D. R. Torgeson, Phys. Rev. Lett. **75**, 2212 (1995).
[16] D. Waibel, G. Fischer, T. Wolf, H. v. Löhneysen, and B. Pilawa, Phys. Rev. B **91**, 214412 (2015).
[17] E. S. Klyushina, J. Reuther, L. Weber, A. T. M. N. Islam, J. S. Lord, B. Klemke, M. Månsson, S. Wessel, and B. Lake, Phys. Rev. B **104**, 064402 (2021).
[18] N. Caci, L. Weber, and S. Wessel, Phys. Rev. B **104**, 155139 (2021).
[19] P. Carretta, M. Filibian, R. Nath, C. Geibel, and P. J. C. King, Phys. Rev. B **79**, 224432 (2009).
[20] D. P. Landau and K. Binder, Phys. Rev. B **24**, 1391 (1981).
[21] A. Cuccoli, T. Roscilde, R. Vaia, and P. Verrucchi, Phys. Rev. B **68**, 060402(R) (2003).
[22] W. J. M. de Jonge, J. P. A. M. Hijmans, F. Boersma, J. C. Schouten, and K. Kopinga, Phys. Rev. B **17**, 2922 (1978).
[23] L. J. de Jongh and H. J. M. de Groot, Solid State Commun. **53**, 731 (1985).
[24] Y. Kohama, M. Jaime, O. E. Ayala-Valenzuela, R. D. McDonald, E. D. Mun, J. F. Corbey, and J. L. Manson, Phys. Rev. B **84**, 184402 (2011).
[25] *Magnetic Properties of Layered Transition Metal Compounds*, edited by L. J. de Jongh (Springer, Netherlands, 1990).
[26] A. Iaizzi, H. D. Scammell, O. P. Sushkov, and A. W. Sandvik, Phys. Rev. B **101**, 104412 (2020).
[27] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).
[28] A. W. Sandvik, Phys. Rev. Lett. **80**, 5196 (1998).
[29] C. Yasuda, S. Todo, K. Hukushima, F. Alet, M. Keller, M. Troyer, and H. Takayama, Phys. Rev. Lett. **94**, 217201 (2005).
[30] A. Cuccoli, T. Roscilde, V. Tognetti, R. Vaia, and P. Verrucchi, Phys. Rev. B **67**, 104414 (2003).
[31] D. Opherden, N. Nizar, K. Richardson, J. C. Monroe, M. M. Turnbull, M. Polson, S. Vela, W. J. A. Blackmore, P. A. Goddard, J. Singleton, E. S. Choi, F. Xiao, R. C. Williams, T. Lancaster, F. L. Pratt, S. J. Blundell, Y. Skourski, M. Uhlarz, A. N. Ponomaryov, S. A. Zvyagin, J. Wosnitza, M. Baenitz, I. Heinmaa, R. Stern, H. Kühne, and C. P. Landee, Phys. Rev. B **102**, 064431 (2020).
[32] F. M. Woodward, P. J. Gibson, G. B. Jameson, C. P. Landee, M. M. Turnbull, and R. D. Willett, Inorg. Chem. **46**, 4256 (2007).
[33] P. A. Goddard, J. Singleton, P. Sengupta, R. D. McDonald, T. Lancaster, S. J. Blundell, F. L. Pratt, S. Cox, N. Harrison, J. L. Manson, H. I. Southerland, and J. A. Schlueter, New J. Phys. **10**, 083025 (2008).
[34] F. Xiao, F. M. Woodward, C. P. Landee, M. M. Turnbull, C. Mielke, N. Harrison, T. Lancaster, S. J. Blundell, P. J. Baker, P. Babkevich, and F. L. Pratt, Phys. Rev. B **79**, 134412 (2009).
[35] J. L. Manson, K. H. Stone, H. I. Southerland, T. Lancaster, A. J. Steele, S. J. Blundell, F. L. Pratt, P. J. Baker, R. D. McDonald, P. Sengupta, J. Singleton, P. A. Goddard, C. Lee, M.-H. Whangbo, M. M. Warter, C. H. Mielke, and P. W. Stephens, J. Am. Chem. Soc. **131**, 4590 (2009).
[36] E. Čižmár, S. A. Zvyagin, R. Beyer, M. Uhlarz, M. Ozerov, Y. Skourski, J. L. Manson, J. A. Schlueter, and J. Wosnitza, Phys. Rev. B **81**, 064422 (2010).
[37] A. J. Steele, T. Lancaster, S. J. Blundell, P. J. Baker, F. L. Pratt, C. Baines, M. M. Conner, H. I. Southerland, J. L. Manson, and J. A. Schlueter, Phys. Rev. B **84**, 064412 (2011).
[38] P. A. Goddard, J. L. Manson, J. Singleton, I. Franke, T. Lancaster, A. J. Steele, S. J. Blundell, C. Baines, F. L. Pratt, R. D. McDonald, O. E. Ayala-Valenzuela, J. F. Corbey, H. I. Southerland, P. Sengupta, and J. A. Schlueter, Phys. Rev. Lett. **108**, 077208 (2012).
[39] P. A. Goddard, J. Singleton, I. Franke, J. S. Möller, T. Lancaster, A. J. Steele, C. V. Topping, S. J. Blundell, F. L. Pratt, C. Baines, J. Bendix, R. D. McDonald, J. Brambleby, M. R. Lees, S. H. Lapidus, P. W. Stephens, B. W. Twamley, M. M. Conner, K. Funk, J. F. Corbey, H. E. Tran, J. A. Schlueter, and J. L. Manson, Phys. Rev. B **93**, 094430 (2016).
[40] T. Thio, T. R. Thurston, N. W. Preyer, P. J. Picone, M. A. Kastner, H. P. Jenssen, D. R. Gabbe, C. Y. Chen, R. J. Birgeneau, and A. Aharony, Phys. Rev. B **38**, 905 (1988).
[41] N. S. Headings, S. M. Hayden, R. Coldea, and T. G. Perring, Phys. Rev. Lett. **105**, 247001 (2010).

13 July 2023 07:58:09

Published under an exclusive license by AIP Publishing

[42] V. Selmani, C. P. Landee, M. M. Turnbull, J. L. Wikaira, and F. Xiao, Inorg. Chem. Commun. **13**, 1399 (2010).
[43] M. S. Haddad, D. N. Hendrickson, J. P. Cannady, R. S. Drago, and D. S. Bieksza, J. Am. Chem. Soc. **101**, 898 (1979).
[44] H. W. Richardson, J. R. Wasson, and W. E. Hatfield, Inorg. Chem. **16**, 484 (1977).
[45] J. Darriet, M. S. Haddad, E. N. Duesler, and D. N. Hendrickson, Inorg. Chem. **18**, 2679 (1979).
[46] S. Vela, J. Jornet-Somoza, M. M. Turnbull, R. Feyerherm, J. J. Novoa, and M. Deumal, Inorg. Chem. **52**, 12923 (2013).
[47] P. Sengupta, C. D. Batista, R. D. McDonald, S. Cox, J. Singleton, L. Huang, T. P. Papageorgiou, O. Ignatchik, T. Herrmannsdörfer, J. L. Manson, J. A. Schlueter, K. A. Funk, and J. Wosnitza, Phys. Rev. B **79**, 060409(R) (2009).
[48] N. A. Fortune, S. T. Hannahs, C. P. Landee, M. M. Turnbull, and F. Xiao, J. Phys. Conf. Ser. **568**, 042004 (2014).
[49] M. Greven, R. J. Birgeneau, Y. Endoh, M. A. Kastner, B. Keimer, M. Matsuda, G. Shirane, and T. R. Thurston, Physica B **199**, 642 (1994).
[50] D. Vaknin, S. K. Sinha, C. Stassis, L. L. Miller, and D. C. Johnston, Phys. Rev. B **41**, 1926 (1990).
[51] J. Igarashi, J. Phys.: Condens. Matter **4**, 10265 (1992).
[52] D. Opherden, F. Bärtl, S. Yamamoto, Z. T. Zhang, S. Luther, S. Molatta, J. Wosnitza, M. Baenitz, I. Heinmaa, R. Stern, C. P. Landee, and H. Kühne, Phys. Rev. B **103**, 014428 (2021).
[53] D. Opherden, M. S. J. Tepaske, F. Bärtl, M. Weber, M. M. Turnbull, T. Lancaster, S. J. Blundell, M. Baenitz, J. Wosnitza, C. P. Landee, R. Moessner, D. J. Luitz, and H. Kühne, Phys. Rev. Lett. **130**, 086704 (2023).
[54] P. Sengupta, A. W. Sandvik, and R. R. P. Singh, Phys. Rev. B **68**, 094423 (2003).
[55] D. Vaknin, L. L. Miller, J. L. Zarestky, and D. C. Johnston, Physica C **274**, 331 (1997).
[56] M. Matsuura, K. Gilijamse, J. E. W. Sterkenburg, and D. J. Breed, Phys. Lett. A **33**, 363 (1970).
[57] M. E. Lines, J. Appl. Phys. **40**, 1352 (1969).
[58] Y. Okwamoto, J. Phys. Soc. Jpn. **53**, 2434 (1984).
[59] H.-Q. Ding and M. S. Makivić, Phys. Rev. B **42**, 6827 (1990).
[60] K. Yosida, *Theory of Magnetism* (Springer-Verlag, Berlin, Heidelberg, 1996).
[61] J. Als-Nielsen, S. T. Bramwell, M. T. Hutchings, G. J. McIntyre, and D. Visser, J. Phys.: Condens. Matter **5**, 7871 (1993).
[62] S. T. Bramwell and P. C. W. Holdsworth, J. Phys.: Condens. Matter **5**, L53 (1993).
[63] S. T. Bramwell and P. C. W. Holdsworth, Phys. Rev. B **49**, 8811 (1994).
[64] A. Taroni, S. T. Bramwell, and P. C. W. Holdsworth, J. Phys.: Condens. Matter **20**, 275233 (2008).
[65] M. Troyer, M. Imada, and K. Ueda, J. Phys. Soc. Jpn. **66**, 2957 (1997).
[66] S. V. Isakov and R. Moessner, Phys. Rev. B **68**, 104409 (2003).
[67] F. Borsa, M. Corti, T. Goto, A. Rigamonti, D. C. Johnston, and F. C. Chou, Phys. Rev. B **45**, 5756 (1992).
[68] F. Tabak, A. Lascialfari, and A. Rigamonti, J. Phys.: Condens. Matter **5**, B31 (1993).
[69] L. Bossoni, P. Carretta, R. Nath, M. Moscardini, M. Baenitz, and C. Geibel, Phys. Rev. B **83**, 014412 (2011).
[70] P. Gaveau, J. P. Boucher, L. P. Regnault, and Y. Henry, J. Appl. Phys. **69**, 6228 (1991).
[71] P. C. Hohenberg and B. I. Halperin, Rev. Mod. Phys. **49**, 435 (1977).
[72] H.-Q. Ding and M. S. Makivić, Phys. Rev. Lett. **64**, 1449 (1990).
[73] J.-K. Kim and M. Troyer, Phys. Rev. Lett. **80**, 2705 (1998).
[74] M. Troyer, "Universality in two-dimensional quantum Heisenberg antiferromagnets," in *Open Problems in Strongly Correlated Electron Systems*, edited by J. Bonča, P. Prelovšek, A. Ramšak, and S. Sarkar (NATO Science Series, Springer, Dordrecht, 2001), Vol. 15.
[75] P. Hasenfratz and F. Niedermayer, Phys. Lett. B **268**, 231 (1991).

13 July 2023 07:58:09

Published under an exclusive license by AIP Publishing

# APPENDIX C

# Article 3

SciPost Phys. 14, 073 (2023)

# Optimal compression of quantum many-body time evolution operators into brickwall circuits

**Maurits S. J. Tepaske**[1], **Dominik Hahn**[2] and **David J. Luitz**[1,2⋆]

**1** Physikalisches Institut, Universität Bonn, Nussallee 12, 53115 Bonn, Germany
**2** Max-Planck-Institut for the Physics of Complex Systems,
Nöthnitzer Straße 38, 01187 Dresden, Germany

⋆ david.luitz@uni-bonn.de

## Abstract

**Near term quantum computers suffer from a degree of decoherence which is prohibitive for high fidelity simulations with deep circuits. An economical use of circuit depth is therefore paramount. For digital quantum simulation of quantum many-body systems, real time evolution is typically achieved by a Trotter decomposition of the time evolution operator into circuits consisting only of two qubit gates. To match the geometry of the physical system and the CNOT connectivity of the quantum processor, additional SWAP gates are needed. We show that optimal fidelity, beyond what is achievable by simple Trotter decompositions for a fixed gate count, can be obtained by compiling the evolution operator into optimal brickwall circuits for the $S = 1/2$ quantum Heisenberg model on chains and ladders, when mapped to one dimensional quantum processors without the need of additional SWAP gates.**

Copyright M. S. J. Tepaske *et al*.
This work is licensed under the Creative Commons
Attribution 4.0 International License.
Published by the SciPost Foundation.

Received 16-05-2022
Accepted 10-01-2023
Published 17-04-2023
doi:10.21468/SciPostPhys.14.4.073

---

## Contents

**1 Introduction** **2**

**2 Model and Method** **3**
  2.1 Model 3
  2.2 Trotter circuits 4
  2.3 Optimization 6
  2.4 Stacking circuits 8
  2.5 Quantities of interest 8

**3 Results** **9**
  3.1 Infidelity 9
  3.2 Out-of-time-ordered correlators 13
  3.3 Analysis of the compressed circuit 15

**4 Conclusion and Outlook** **18**

A Gate count equations 19

B Convergence of the optimization 20

C OTOC details 21

D Emergence of lattice symmetries 22

References 25

## 1 Introduction

Quantum processors are a rapidly evolving technology which is expected to be pivotal for many classically hard problems like integer factorization, database search, optimization and many others [1–4]. While truly universal quantum computing is still a long shot, one of the most promising near-term applications is the simulation of complex quantum systems due to their relative similarity to the quantum hardware itself. The simulation of such systems on classical computers is extremely hard due to the exponential complexity in terms of storage and computer time, while both problems are naturally solved on quantum hardware.

There are two different approaches: analog and digital quantum simulators. Analog simulators are specifically engineered systems to mimic the corresponding dynamics of the target system and are often based on quantum optical setups. This technique has been successfully applied to condensed matter systems [4–9] and lattice gauge theories [10–12] and is in principle extremely powerful but requires a tailored experimental setup for a given type of problem.

In contrast, digital quantum simulators [13] rely on a discrete representation of the wave function on an array of two level systems (dubbed qubits), which can be fully controlled by a universal set of quantum gates which allows in principle for the representation of any unitary operation on the many-body wave function, represented as a sequence of gates. Due to the universal representation of the wave function, this is an attractive approach which is extremely flexible once a suitable mapping of the system of interest to qubits is devised. Recent applications include condensed matter systems [14–19], simulations from quantum chemistry [11,20–22] and high-energy physics [23,24]. Digital quantum simulations were also used to realize exotic phases of matter like time crystals [25,26] and quantum spin liquids [27].

The state-of-the-art method for simulating the real time dynamics of complex quantum systems involves a factorization of the time evolution operator into a sequence of gates using Trotter decompositions of different orders [28–32], introducing discrete time steps to get an approximation of the exact time evolution of the system. This introduces a discretization error, which can be systematically controlled by using smaller step sizes. As a downside, small step sizes require a larger number of gates. Due to the fragility of the quantum state stored in the machine, and due to hardware imperfections, each additional gate potentially introduces new sources of error due to dissipation processes. Hence a trade-off between discretization errors and errors due to intrinsic machine noise during the simulation is required. To achieve optimal fidelity in light of this tradeoff, it is therefore important to minimize the resource costs for a given simulation. Recent work yielded tighter bounds for the discretization errors [33]. Furthermore, it was also argued recently that beyond a certain step size the fidelity of the Trotter decomposition breaks down in a universal fashion, leading to a regime of quantum chaos [34,35]. This sets also upper bounds for possible step sizes. It remains however unclear, whether better alternatives to Trotter decompositions exist.

One promising approach in this regard are quantum variational algorithms. The main idea of them is to approximate a time-evolved state using a parametrized circuit [36–39]. The parameters are then fixed using optimization algorithms on a quantum computer. Recent numerics suggest that the number of parameters needed to describe time-evolved states or ground states scales favorable even in comparison to matrix-product states [40, 41]. Most of these algorithms involve optimization where gradients are measured directly on the quantum devices, or they use deep learning approaches. However, the measurement of gradients on a quantum device is currently infeasible due to the high error rates, while optimization using deep neural networks is not controlled.

In this paper we take a more universal approach. Rather than focussing on the wave function, we directly target the time evolution operator, aiming at a compact representation as a shallow circuit. We use brickwall circuits in which the gates are parametrized two qubit unitaries, connecting neighboring qubits in the architecture of the quantum processor as an ansatz for the time evolution operator. This parametrized circuit can be optimized classically to represent the time evolution operator for a given time step with high fidelity. The resulting circuit can then be repeated to evolve the quantum state to later times. We show that such an optimized circuit can yield significantly higher fidelity time evolution for a fixed gate count compared to the traditional Trotter decomposition and is thus superior for digital quantum simulation.

We also show that this strategy allows us to obtain similar accuracy using significantly less gates, even for systems where the physical geometry does not coincide with the proposed circuit architecture, essentially “baking in” the otherwise required SWAP gates to match geometries into the circuit. As an interesting benchmark problem, we use our approach to compute out-of-time-ordered correlators (OTOCs) and show that we achieve better accuracy than Trotter methods with similar resource cost. Finally, we analyze the gate structure of the optimized gates, as a first step towards further improvements of this approach.

## 2 Model and Method

### 2.1 Model

For concreteness and simplicity, we focus on simulating finite systems of $s = 1/2$ spins on a lattice with $L$ sites, designed to be performed on a quantum processor with an identical Hilbert space $\mathcal{H}$, which is the product space of $L$ two-level quantum systems (qubits) $\bigotimes_{i=1}^{L} \mathcal{Q}_i$ and has an exponentially growing dimension $\dim H = 2^L$. Specifically, we discuss spin-1/2 systems with SU(2) symmetric Heisenberg couplings

$$h_{ij} = \sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + \sigma_i^z \sigma_j^z \,, \tag{1}$$

between nearest neighbor (NN) spins on a chain $\mathfrak{c}$ and a triangular ladder $\mathfrak{l}$, both with open boundary conditions, i.e.

$$H_{\mathfrak{c}} = \sum_{\langle i,j \rangle} h_{ij} \,, \qquad H_{\mathfrak{l}} = \sum_{\langle\langle i,j \rangle\rangle} h_{ij} \,. \tag{2}$$

Here $\sigma_{x,y,z}$ are the usual Pauli matrices while $\langle i, j \rangle$ and $\langle\langle i, j \rangle\rangle$ denote the NN sites of the chain, or the NN sites of our triangular ladder geometry (note that this is identical to a chain with nearest and next nearest neighbor (NNN) interactions). These lattice geometries are illustrated in Fig. 1.

Most current quantum devices using superconducting qubits are not capable of all-to-all connectivity, i.e. due to the chip setup two qubit gates can only be applied between neighboring

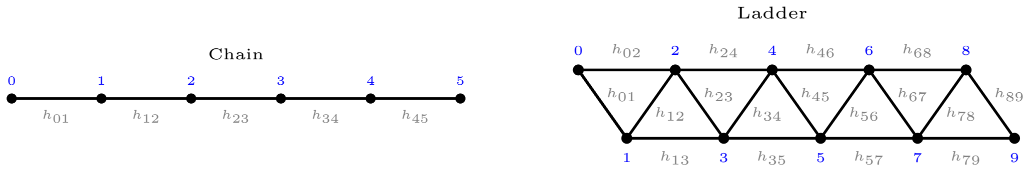


Figure 1: The chain (left) and triangular ladder (right) lattice geometries used in this work.

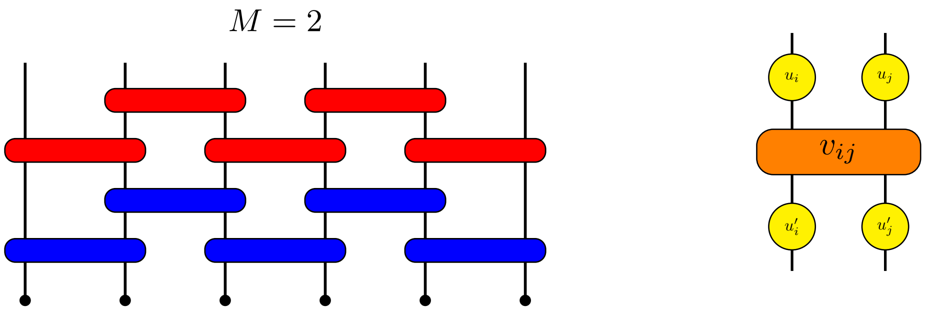


Figure 2: Left: A brickwall circuit with depth $M = 2$ for six qubits, with each color representing a $M = 1$ layer. Circles represent the initial state of the qubits and boxes indicate a two qubit unitary gate applied to a pair of neighboring qubits. Right: Parametrization of a two qubit unitary as a product of four single qubit gates and one two qubit gate.

qubits, which are arranged in different geometries [42–44] In order to apply gates between distant qubits, one has to use a sequence of swap gates, which exchange the quantum state of neighboring qubits, such that effectively the states of distant qubits are moved to neighboring qubits in the processor geometry. On these, any two qubit gate can be applied and then the swap sequence needs to be applied in reverse order. This requires a great number of additional gates and therefore introduces further possible sources of errors.

Our goal is therefore to find the best unitary circuit $\mathcal{C}$ of a given depth $M$ to approximate the time evolution operator $\mathcal{U}(t) = \exp(-itH_{c/l})$. In order to mimic the limited connectivity of current quantum devices, we choose $\mathcal{C}$ to consist only of NN two-qubit gates on a 1d chain, arranged in a brickwall pattern, i.e. we model our quantum processor as an open chain of qubits, while one of our physical models we want to simulate on this machine has a different, triangular ladder, geometry. This allows us to investigate whether it is possible to compile the time evolution operator in a nearest neighbor, brickwall circuit (exemplified in the left panel of Fig. 2) without the need for additional swap gates, which are generally costly on superconducting platforms.

## 2.2 Trotter circuits

To benchmark the performance of the brickwall circuits we will compare them with the first-, second- and fourth-order Trotter circuits that are based on the well known Trotter decompositions [45]. Here we introduce these circuits for the Hamiltonians (2) that are used in this work.

For the chain Hamiltonian $H_c$ we have two non-commuting parts, namely the bond Hamil-

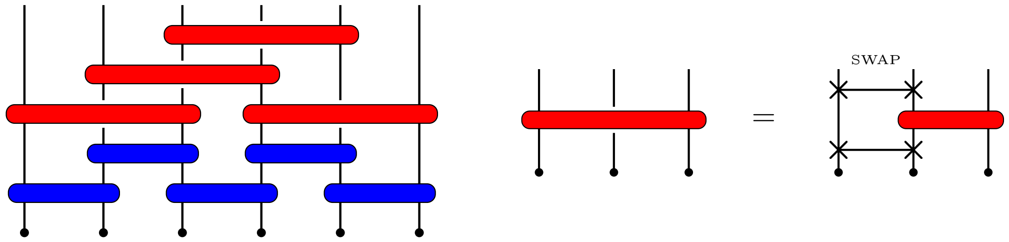


Figure 3: Left: The blue brickwall layer encodes the first-order Trotter decomposition for NN interacting Hamiltonians. The combination of the blue and red layers encodes the first-order Trotter decomposition for NNN interacting Hamiltonians, where the blue gates act on NN qubits whereas the red gates act on NNN qubits. Right: The decomposition involving SWAP gates, displayed as the crossed line, which is used to convert the NNN two-qubit gate into a circuit involving only two-qubit gates.

tonians $h_{i,i+1}$ (1) on alternating bonds, such that we can split $H_\mathfrak{c}$ in two commuting parts as

$$H_\mathfrak{c} = H_1 + H_2 = \sum_{i=0,2,\ldots} h_{i,i+1} + \sum_{i=1,2,\ldots} h_{i,i+1}\,. \tag{3}$$

For the ladder Hamiltonian we have on top of this three extra non-commuting parts due to the NNN couplings, i.e. we can split $H_\mathfrak{l}$ into five commuting parts as

$$H_\mathfrak{l} = H_\mathfrak{c} + H_3 + H_4 + H_5 = H_\mathfrak{c} + \sum_{i=0,3,\ldots} h_{i,i+2} + \sum_{i=1,4,\ldots} h_{i,i+2} + \sum_{i=2,5,\ldots} h_{i,i+2}\,. \tag{4}$$

By writing the Hamiltonians in this way we can define the $M = 1$ first-order Trotter circuits for $H_\mathfrak{c}$ and $H_\mathfrak{l}$ as [45]

$$\begin{align}
\mathcal{U}_\mathfrak{c}^{1\text{st}}(t) &= \mathcal{U}_1(t)\mathcal{U}_2(t) \tag{5}\\
&= \exp(-\mathrm{i}tH_1)\exp(-\mathrm{i}tH_2)\,, \tag{6}\\
\mathcal{U}_\mathfrak{l}^{1\text{st}}(t) &= \mathcal{U}_1(t)\mathcal{U}_2(t)\mathcal{U}_3(t)\mathcal{U}_4(t)\mathcal{U}_5(t) \tag{7}\\
&= \exp(-\mathrm{i}tH_1)\exp(-\mathrm{i}tH_2)\exp(-\mathrm{i}tH_3)\exp(-\mathrm{i}tH_4)\exp(-\mathrm{i}tH_5)\,. \tag{8}
\end{align}$$

These circuits approximate the exact $\mathcal{U}(t) = \exp(-\mathrm{i}tH_{\mathfrak{c}/\mathfrak{l}})$ with error $\mathcal{O}(t^2)$ [33]. Note that depth $M = 1$ for the Trotter circuits does not mean one brickwall layer, but instead one Trotter step $\mathcal{U}_{\mathfrak{c}/\mathfrak{l}}^{1\text{nd}}(t)$. While these coincide for the first-order Trotter circuit for the chain, this is not the case for the first-order Trotter circuit for the ladder, and for the second- and fourth-order Trotter circuits which we introduce below. The circuit diagram for $\mathcal{U}_\mathfrak{c}^{1\text{st}}(t)$ is shown as the blue brickwall layer in the left panel of Fig. 3, where $\mathcal{U}_1(t)$ is the half-brickwall layer on odd bonds and $\mathcal{U}_2(t)$ is the half-brickwall layer on even bonds. The circuit diagram for $\mathcal{U}_\mathfrak{l}^{1\text{st}}(t)$ is the full circuit in this figure, where $\mathcal{U}_1(t)$ and $\mathcal{U}_2$ again form the blue brickwall layer while $\mathcal{U}_3(t)$, $\mathcal{U}_4(t)$ and $\mathcal{U}_5(t)$ form the red layer, containing two-qubit gates that act on NNN instead of NN qubits. To turn this into a circuit that involves only NN two-qubit gates we introduce the SWAP gate and decompose every NNN gate as in the right panel of Fig. 3.

The circuit layers $\mathcal{U}_1, \mathcal{U}_2, \mathcal{U}_3, \mathcal{U}_4, \mathcal{U}_5$ form the building blocks of the second- and fourth-order Trotter circuits. The $M = 1$ second-order Trotter circuits are composed as [45]

$$\begin{align}
\mathcal{U}_\mathfrak{c}^{2\text{nd}}(t) &= \mathcal{U}_1(t/2)\mathcal{U}_2(t)\mathcal{U}_1(t/2)\,, \tag{9}\\
\mathcal{U}_\mathfrak{l}^{2\text{nd}}(t) &= \mathcal{U}_1(t/2)\mathcal{U}_2(t/2)\mathcal{U}_3(t/2)\mathcal{U}_4(t/2)\mathcal{U}_5(t)\mathcal{U}_4(t/2)\mathcal{U}_3(t/2)\mathcal{U}_2(t/2)\mathcal{U}_1(t/2)\,, \tag{10}
\end{align}$$

which approximate the exact evolution operators with error $\mathcal{O}(t^3)$ [33]. Using these second-order Trotter circuits we can define the $M = 1$ fourth-order Trotter circuits as [45]

$$\mathcal{U}^{\text{4th}}_{\mathfrak{c}/\mathfrak{l}}(t) = \mathcal{U}^{\text{2nd}}_{\mathfrak{c}/\mathfrak{l}}(t_1)\mathcal{U}^{\text{2nd}}_{\mathfrak{c}/\mathfrak{l}}(t_1)\mathcal{U}^{\text{2nd}}_{\mathfrak{c}/\mathfrak{l}}(t_2)\mathcal{U}^{\text{2nd}}_{\mathfrak{c}/\mathfrak{l}}(t_1)\mathcal{U}^{\text{2nd}}_{\mathfrak{c}/\mathfrak{l}}(t_1)\,, \tag{11}$$

where we defined the time steps

$$t_1 = \frac{1}{4-4^{1/3}}t\,, \qquad t_2 = (1-4t_1)t\,. \tag{12}$$

These circuits approximate the exact evolution operators with error $\mathcal{O}(t^5)$ [33].

Because we are concerned with circuits that are implemented on a quantum processor with only NN qubit connectivity, we have to convert every NNN two-qubit gate that appears in $\mathcal{U}^{\text{1st}}_{\mathfrak{l}}, \mathcal{U}^{\text{2nd}}_{\mathfrak{l}}, \mathcal{U}^{\text{4th}}_{\mathfrak{l}}$ to three NN two-qubit gates, as shown in the right panel of Fig. 3. The gate counts $N_g$ of the resulting NN Trotter circuits are given in Sec. A, also for the chain geometry.

## 2.3 Optimization

Each two-qubit gate $U_{ij} \in \mathbb{C}^{4\times 4}$ of the circuit $\mathcal{C}$, acting on two neighboring qubits $i$ and $j$, can be decomposed into a product of one-qubit gates $u_i \in \mathbb{C}^{2\times 2}$ and a two-qubit gate $v_{ij} \in \mathbb{C}^{4\times 4}$ [46]

$$U_{ij} = (u_i \otimes u_j)v_{ij}(u_i' \otimes u_j')\,. \tag{13}$$

Here $v_{ij}$ is parameterized as

$$v_{ij}(\lambda_0,\lambda_1,\lambda_2) = e^{-i(\lambda_0\sigma_i^x\otimes\sigma_j^x+\lambda_1\sigma_i^y\otimes\sigma_j^y+\lambda_2\sigma_i^z\otimes\sigma_j^z)}\,, \tag{14}$$

with three real parameters $\lambda_{0,1,2}$, and the $u_i$ are parameterized up to a global phase as

$$u_i(\phi_0,\phi_1,\phi_2) = \begin{pmatrix} e^{i\phi_1}\cos(\phi_0) & e^{i\phi_2}\sin(\phi_0) \\ -e^{-i\phi_2}\sin(\phi_0) & e^{-i\phi_1}\cos(\phi_0) \end{pmatrix}\,, \tag{15}$$

each containing three real parameters $\phi_{0,1,2}$. Hence this decomposition of $U_{ij}$ contains 15 real parameters, and it can be visualised as in the right panel of Fig. 2. To represent the unitary gate as a global unitary matrix, acting on the full wave function, we introduce its matrix form

$$\text{mat}(U_{ij}) = I_{2^{i-1}} \times U_{ij} \times I_{2^{L-j}}\,, \tag{16}$$

by taking the Kronecker product with identity matrices on the qubits on which the gate does not act (and implicitly encoding the nearest neighbor condition $j = i + 1$). The entire circuit is a product of such unitaries and can formally be expressed by

$$\mathcal{C} = \prod_{k=0}^{N_g} \text{mat}(U_{i_k,j_k})\,, \tag{17}$$

where $N_g$ is the total number of gates in the circuit. Since each gate is parametrized by $\vec{\theta}_{i_k} = (\vec{\lambda}_{i_k}, \vec{\phi}_{i_k})$, the circuit depends on all these $15N_g$ parameters $\vec{\theta} = (\vec{\theta}_{i_0}, \vec{\theta}_{i_1} \ldots) \in \mathbb{R}^{15N_g}$

$$\mathcal{C}(\vec{\theta}) = \prod_{k=0}^{N_g} \text{mat}(U_{i_k,j_k}(\vec{\theta}_{i_k}))\,. \tag{18}$$

In practice, when stacking the gates to form the circuit, we merge two one-qubit unitaries into a single one-qubit unitary where possible, since a product of general one-qubit unitaries can be written as a single general one-qubit unitary. This reduces the amount of circuit parameters.

We would like to find an optimal parameter set $\vec{\theta}$ for a given circuit architecture, such that the distance between the unitary represented by the circuit $\mathcal{C}(\vec{\theta})$ and the targeted time evolution operator $\mathcal{U}$ of the system up to time $t$ is minimized. For two unitary operators $\mathcal{U}$ and $\mathcal{C}$, we therefore define a measure of distance in terms of the normalized Frobenius norm, namely the "infidelity" $\epsilon$, given by

$$\epsilon = \frac{1}{2}\frac{\|\mathcal{U}-\mathcal{C}\|_F^2}{2^L} = \frac{1}{2^{L+1}}\mathrm{Tr}\big[(\mathcal{U}-\mathcal{C})^\dagger(\mathcal{U}-\mathcal{C})\big] = 1 - \frac{\mathrm{Re}\mathrm{Tr}\big[\mathcal{U}^\dagger\mathcal{C}\big]}{2^L}\,. \tag{19}$$

We use this infidelity as an objective function, such that we obtain a minimization problem for a fixed circuit architecture (number and sequence of two qubit gates). In our case the target unitary $\mathcal{U}$ is an approximation of an exact time-evolution operator, where the error stems from the tensor network methods that make the optimization tractable.

The objective function $\epsilon$ needs to be evaluated many times during the optimization and we find that it is efficient to first compress the time evolution operator $\mathcal{U}$ into a matrix product operator (MPO) $\mathcal{M}_\chi$ of bond dimension $\chi$, such that we can calculate $\epsilon$ via efficient standard tensor network methods. For the local systems we investigate here and for short times, this is always efficient, due to the low operator entanglement of the time evolution operator [47]. In particular, we discard the smallest singular values for which the squares sum to a tiny number, since their contribution is negligible, such that lowly entangled operators do not saturate the maximum bond dimension $\chi$. To obtain the (truncated) MPO representation of $\mathcal{U}$ with negligible discretization error, we take an identity MPO and perform time-evolving block decimation [45,48] with a small timestep $\delta t = 10^{-4}$ and fourth-order Trotter decomposition, such that the introduced error is negligible.[1]

To optimize the parameters $\vec{\theta}$ of the circuit such that $\epsilon$ is minimal, we employ the paradigm of differentiable programming [49]. Here the gradient $\nabla_{\vec{\theta}}\,\epsilon$ is calculated in a similar fashion as the original backpropagation algorithm used for deep neural networks [50], which has been generalized to arbitrary programs, including tensor network algorithms [49]. To this end, a program is represented as a computational graph through which the local gradients are propagated, which requires each computational component to have a well-defined gradient. In particular, for the tensor network algorithm in this work, the SVD is a crucial component, and so it is important to construct a stable SVD gradient [49]. Fortunately, differentiable programming inherits the cost from its base algorithm, i.e. in our case from the $M$ SVDs that are performed when obtaining the circuit MPO at every iteration. As a result our algorithm has the scaling $\mathcal{O}(N_i L M d^6 \chi^3)$, where $N$ is the amount of gradient descent iterations. Importantly, even though the cost scales linearly with system size $L$ and circuit depth $M$, the amount of parameters grows as $\mathcal{O}(LM)$, such that the amount of iterations required to reach a low-lying minimum also grows, because local minima prolifrate with growing parameter count [51].

Using the global gradient $\nabla_{\vec{\theta}}\,\epsilon$ we then perform gradient descent. We use this global optimization procedure instead of the local optimization from [40] because we found that this yields significantly higher fidelity when an Adam-like adaptive learning rate is used [52]. Here it is crucial not to stop optimizing when the infidelity appears to have stagnated, since we have often found that the optimization gets stuck in such a "local minimum" for some time before it jumps out and converges to a lower minimum. This is possibly related to the "barren plateau" problem that often occurs when performing gradient descent for quantum circuits with a large parameter space, where the optimization reaches a set of circuit parameters for which the majority of its gradients become very small such that the optimization (temporarily) halts [53]. In Sec. B we review the Adam method and discuss the mentioned convergence behavior in more detail.

---

[1] We compare the results for our circuits to Trotter circuits with comparable gate counts, and in all instances of the involved Trotter circuits, the timesteps are several orders of magnitude larger than the stepsize used to approximate the target unitary $\mathcal{U}$.

At small $M$ the optimized circuits in a sense compress the targeted time evolution operator, especially when its time-step is large, and therefore they are called "compressed circuits". In Sec. D we check if the lattice symmetries of the targeted unitary emerge in the gates of the optimized circuits.

## 2.4 Stacking circuits

The general strategy we implement is the following: For some (short) timestep $t$, we find an optimal circuit $\mathcal{C}(\vec{\theta})$ which best approximates the targeted time evolution operator $U_t$. In principle, $t$ is arbitrary, with the general logic that shorter $t$ unitaries can be encoded by shallower circuits (lower $M$). In practice, $t$ will be also governed by the time grid, on which observables should be evaluated, although this could be achieved also by working with two or more different optimized circuits with different $t$, a case we do not further discuss in this work. To propagate the wave function to longer times, which are multiples of $t$, we then use the circuit

$$\mathcal{C}(\vec{\theta})^n \approx \mathcal{U}_t^n \,. \tag{20}$$

It is interesting to investigate how well this stacked circuit performs for time evolution to longer times and we will confront these results to benchmarks for the circuits discussed in Sec. 2.2 that result from traditional Trotter decompositions.

## 2.5 Quantities of interest

Having obtained the compressed circuits for short times, for which the relatively low entanglement allows for an accurate description with truncated MPOs, we then compute $\epsilon$ for long times using the stacked circuits as approximation. If we now were to use the same MPO formalism that was used during the optimization, the growing of entanglement as we stack the circuit multiple times results in either an unfeasible amount of required computational resources or significant truncation errors. In particular, the stacked circuit represents a target unitary at large times, which generally has large entanglement, such that an accurate MPO representation requires a saturated bond dimension, i.e. the central tensors would require bond dimension $2^L$ to prevent significant truncation errors.

For a highly entangled MPS $|\psi_i\rangle$ this central bond dimension is instead $2^{L/2}$, which is still managable for the system sizes considered in this work. Hence, to probe the true representablity of the stacked circuit, without having to deal with artefacts of the tensor network method, we use typicality [54]. Here the trace in Eq. 19 is replaced by the average over $N_\psi$ Haar random states $|\psi_i\rangle$, i.e.

$$\mathrm{Tr}\big[\mathcal{U}^\dagger \mathcal{C}\big] \approx \frac{1}{N_\psi} \sum_i \langle \psi_i | \mathcal{U}^\dagger \mathcal{C} | \psi_i \rangle \,. \tag{21}$$

This allows us to calculate $\epsilon$ in an unbiased manner for the system sizes considered in this work.

Besides using the infidelity $\epsilon$ as a measure of the performance of the circuits, we will also use the circuits to compute out-of-time-ordered correlators (OTOCs) [55]. For spin-1/2 $\sigma^z$ operators, the OTOC $C_{ij}$ between lattice sites $i$ and $j$ is defined with the Frobenius norm as

$$C_{ij}(t) = \Big\| \big[ \sigma_i^z(t), \sigma_j^z \big] \Big\|_{\mathrm{F}}^2 \,, \tag{22}$$

where $\sigma_i^z(t) = \mathcal{C}^\dagger \sigma_i^z \mathcal{C}$ is the spin operator on site $i$ evolved by the circuit. As for the infidelity, it is important to use typicality instead of the truncated MPO formalism when calculating $C_{ij}$ for a circuit that is stacked many times.

To calculate (22) we invoke the hermiticity of the spin operators $\sigma^z$, such that by expanding the commutator in (22) we can write the OTOC as

$$C_{ij}(t) = 1 - \frac{1}{4}\mathrm{Tr}\big[\sigma_j^z \sigma_i^z(t) \sigma_j^z \sigma_i^z(t)\big], \tag{23}$$

which is readily calculated in the MPO formalism. Concretely, we take an identity MPO and put a $z$-spin operator $\sigma^z$ at site $i$, which is then evolved in the Heisenberg picture by the circuit $\mathcal{C}$, yielding a different MPO. Then we again take an identity MPO and put a $z$-spin operator on site $j$, which we do not evolve. Then we calculate the trace in (23) via a full contraction of four MPOs, which can be done efficiently.

## 3 Results

To benchmark the performance of the compression strategy outlined in Sec. 2, we systematically analyze the infidelity $\epsilon$ as a function of simulation time step $t$, total gate count $N_g$ and system size, in direct comparison to Trotter decompositions of different orders, and present these results in Sec. 3.1. In Sec. 3.2 we extend this systematic analysis to out-of-time-ordered correlators (OTOCs) (22). Furthermore, in Sec. 3.3 we probe the structure of the gates that make up the optimized circuits, in an attempt to uncover the structures that allow these circuits to outperform their Trotter counterparts.

### 3.1 Infidelity

As a first test of the circuit optimization algorithm outlined in Sec. 2, we compare the optimal infidelities of compressed circuits to those of comparable Trotter circuits. Concretely, we consider time evolution operators of the chain and ladder Heisenberg Hamiltonians (2) at three system sizes $L = 8, 12, 16$ and two time-steps $t = 1, 2$. For each Hamiltonian, system size and time-step, we determine the time evolution operator $\mathcal{U}$ with numerically negligible discretization error for a certain bond dimension $\chi$, and perform the global optimization as outlined in Sec 2 to minimize the infidelity $\epsilon$ of the compressed circuit. For $L = 8, 12, 16$ we have taken $\chi = 256, 150, 100$ as a compromise between precision and practical efficiency. We note that our main concern here is not to get a numerically exact MPO representation, but rather a reasonably good approximation of the time evolution operator. We call this our target time-evolution operator, which we want to approximate with our circuits.

As a first benchmark, we take for each of our parameter sets various circuit depths $M = 1, 2, 4, 8, 16$, where $M$ is the number of elementary layers of $L-1$ gates, and consider $\epsilon$ as a function of the corresponding gate count $N_g$ (see Sec. A for details on how to obtain the number of gates). We compare this with first-, second- and fourth-order Trotter circuits [45].

The results are shown in Fig. 4. The left pair of panel columns is for the chain and the right pair is for the ladder. The first and third columns are for time-step $t = 1$ and the second and fourth are for $t = 2$. The upper row is for system size $L = 8$, the middle row is for $L = 12$, and the bottom row is for $L = 16$. Each panel contains the infidelities of the optimized compressed circuits (CC) as a red line, and the infidelities of the Trotter circuits as blue lines. The infidelities of the Trotter circuits are calculated for the same depths $M$ as the compressed circuit, where it should be remembered from Sec. 2.2 that in this case $M$ is not necessarily equal to the amount of brickwall layers in the Trotter circuit, but is instead equal to the amount of Trotter steps that compose the circuit. The time-step of the Trotter step is chosen as $t/M$, such that $M$ subsequent steps correspond to a total time-step $t$. The gate counts of the Trotter circuits were calculated with the expressions in Sec. A, which take into account the number of swap gates required to map the ladder geometry to a chain of qubits.

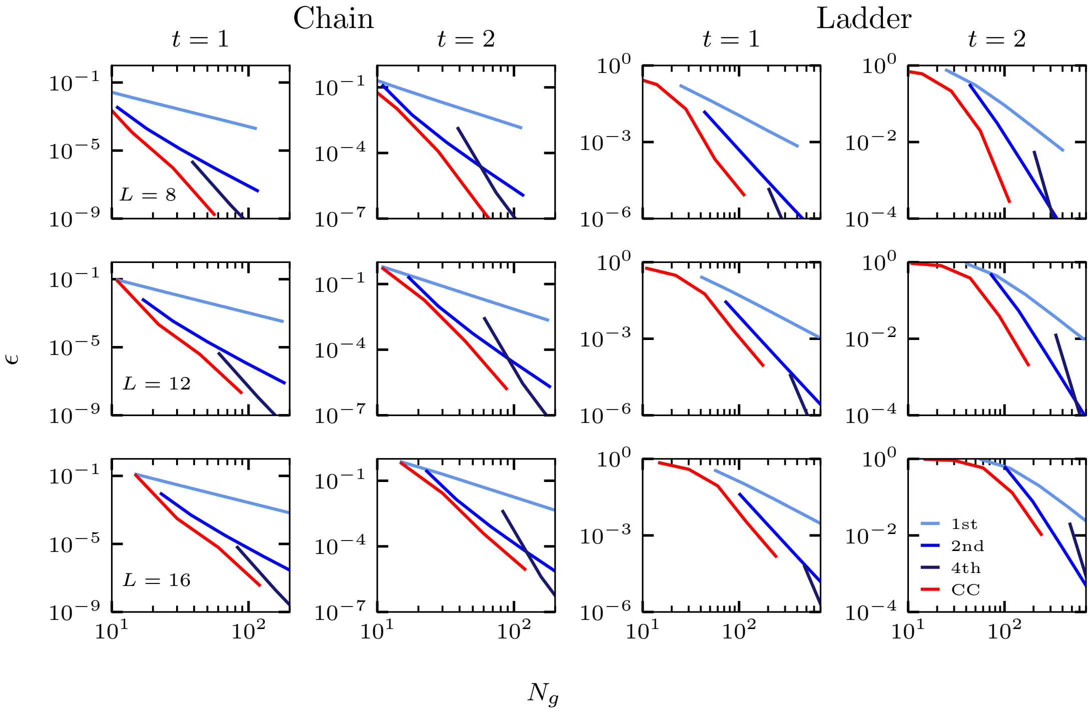


Figure 4: The infidelity $\epsilon$ versus gate count $N_g$ for the time evolution operator of the Heisenberg model on a chain (left panels) and ladder (right panels) in log-log scale. The first and third columns are for $t = 1$ while the second and fourth columns are for $t = 2$. The top panels are for $L = 8$ and a time evolution MPO with $\chi = 256$, the middle panels are for $L = 12$ with $\chi = 150$, and the bottom panels are for $L = 16$ with $\chi = 100$. The blue curves represent the Trotter circuits and the red curve represents the compressed circuit (CC).

From Fig. 4 it becomes clear that per gate the compressed circuit outperforms the Trotter circuits for all considered parameter sets. Moreover, it appears that for $L = 8$ the infidelity of the compressed circuit roughly scales with $N_g$ like the best Trotter order, but with a more favorable prefactor, i.e. at intermediate gate counts it scales as second-order whereas at the highest probed gate count it scales as fourth-order. We have found that the same picture emerges when plotting $\epsilon$ versus the $t$ at which the circuit was optimized, where $M = 1$ scales like first-order Trotter, and by increasing $M$ we approach the fourth-order scaling, passing through the second-order scaling. We expect the same to hold for $L = 12$ and $L = 16$, if we could reach a lower minimum, but here the optimization is more expensive.

Having considered the infidelities of the compressed circuits at the time-step for which they were optimized, we now quantify how these infidelities grow when the circuits are stacked, which we do for the same systems as in Fig. 4. To this end we select a compressed circuit that was optimized at $t = 2$, and take for every Trotter order a circuit of depth $M$ with a gate count as close as possible to that of the compressed circuit, and choose its time-step to be $t/M$.

Concretely, for the chain we take a compressed circuit with $M = 8$, in which case we have to take first-, second-, and fourth-order Trotter circuits with $M = 8, 7, 1$. Using the gate count equations from Sec. A we find that for $L = 8$ the circuits have $N_g = 56, 56, 53, 39$, for $L = 12$ they have $N_g = 88, 88, 83, 61$, and for $L = 16$ they have $N_g = 120, 120, 113, 83$. For the ladder we take a compressed circuit with $M = 16$, such that we have to take first-, second-, and fourth-order Trotter circuits with $M = 4, 3, 1$. The corresponding gate counts are $N_g = 112, 100, 124, 204$ for $L = 8$, $N_g = 176, 164, 207, 341$ for $L = 12$,

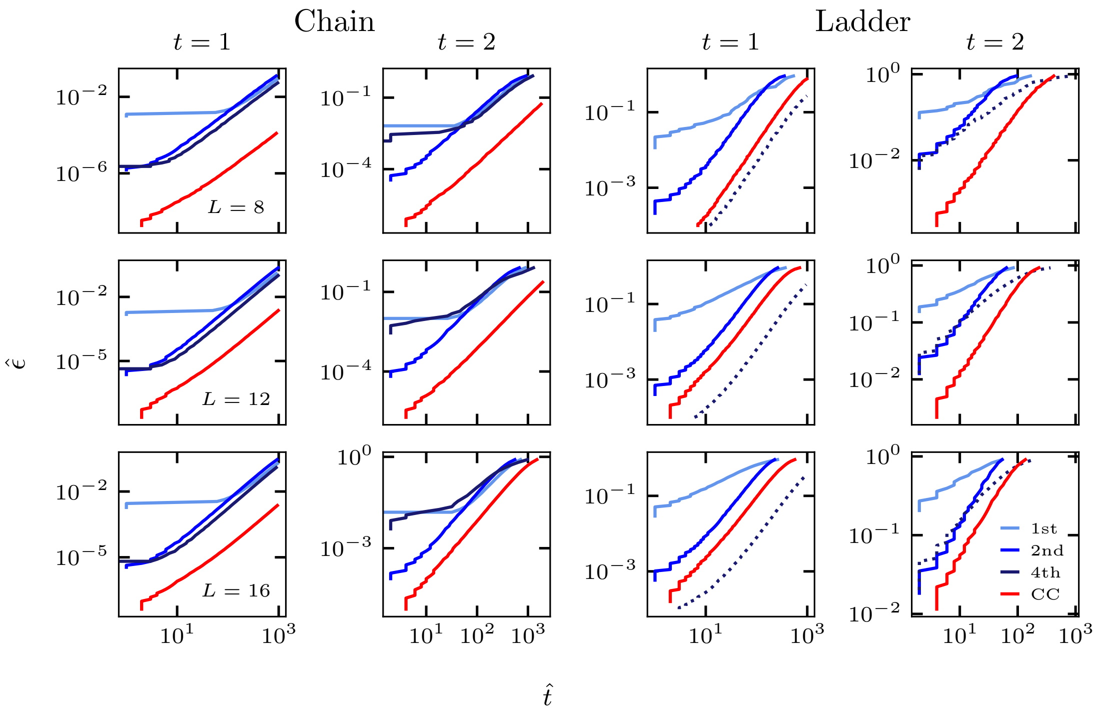


Figure 5: The time $\hat{t}$ after which the stacked circuits exceed the infidelity threshold $\hat{\epsilon}$, for the time evolution operator of the Heisenberg model on a chain (left panels) and ladder (right panels) in log-log scale. The first and third columns are for circuits optimized at $t = 1$ while the second and fourth columns are for $t = 2$, with the circuits being stacked up to a thousand times. The circuits were chosen such that they have similar gate counts, with $M = 8, 8, 7, 1$ for the chain and $M = 16, 4, 3, 1$ for the ladder, for the compressed circuit and first-, second- and fourth-order Trotter circuits, respectively. The top panels are for $L = 8$ with $\chi = 256$, the middle panels are for $L = 12$ with $\chi = 150$, and the bottom panels are for $L = 16$ with $\chi = 100$. The blue curves represent the Trotter circuits and the red curve represents the compressed circuit (CC). The fourth-order Trotter circuit for the ladder is displayed as a dashed line, since it contains roughly twice as many gates as the compressed circuit and is therefore not necessarily indicative of their relative performance.

and $N_g = 240, 228, 290, 478$ for $L = 16$.

To quantify the quality of the compressed and Trotter circuits under stacking, we take various infidelity thresholds $\hat{\epsilon}$ and stack the circuits up to a thousand times until they cross this threshold at some time $\hat{t}$, i.e. we determine $\epsilon(\hat{t}) = \hat{\epsilon}$. As mentioned in Sec. 2 we utilize typicality (21) to calculate the stacked infidelities.

In Fig. 5 we plot $\hat{\epsilon}$ versus $\hat{t}$ in log-log scale. The used color coding is identical to that of Fig. 4, except that the fourth-order Trotter circuit for the ladder is now represented with a dashed line, to emphasize that its infidelity relative to that of the compressed circuit is not necessarily indicative of the relative performance, because it contains roughly twice as many gates as the compressed circuit. From these plots it is clear that the advantage of the compressed circuits from Fig. 4 is not lost when stacking it many times. In particular, in all considered cases the compressed circuits are able to go to significantly larger times, at all infidelity thresholds, than the Trotter counterparts. The only exception is for the ladder at $t = 1$, where the fourth-order Trotter circuit performs better, but as mentioned this Trotter circuit has twice as many gates as the compressed circuit and is therefore not a fair comparison.

From the plots we extract the universal quadratic power-law $\hat{\epsilon} \propto \hat{t}^2$, for both the compressed and the Trotter circuits. This error scaling is analogous to first-order Trotter decomposition. The only exception is the ladder with $L = 16$ at $t = 2$, where the infidelity reaches $\epsilon \approx 1$ rather quickly, such that it is situated in the rounding part that is also observed for the $t = 1$ ladder curves at the high-infidelity end. The gap between the compressed circuits and the best performing Trotter circuits is thus found to grow quadratically with $\hat{t}$. Concretely, for the chain with $L = 12$ and timestep $t = 1$, we find that for $\hat{\epsilon} = 10^{-3}$ the compressed circuit has $\hat{t} = 644$ whereas the best Trotter circuit (i.e. of fourth-order) has $\hat{t} = 94$. For $\hat{\epsilon} = 10^{-4}$ we instead get $\hat{t} = 201$ for the compressed circuit and $\hat{t} = 29$ for the best Trotter circuit. For the same system at timestep $t = 2$, we find that at $\hat{\epsilon} = 10^{-3}$ the compressed circuit has $\hat{t} = 116$ while the best Trotter circuit has $\hat{t} = 14$. At $\epsilon = 10^{-2}$ we have $\hat{t} = 378$ for the compressed circuit and $\hat{t} = 46$ for the best Trotter circuit. From these values it is clear that for the chain we can go roughly eight times further in time than the best Trotter circuit with similar gate count. These values are for $L = 12$, and the same analysis at $L = 8$ reveals that here we can go fourteen to twenty times as far, while for $L = 16$ we can go three to eight times as far, with the lower bounds for $t = 2$ and the upper bounds for $t = 1$. These values emphasize that the larger we choose $\hat{\epsilon}$, the larger the gap between $\hat{t}$ of the compressed and Trotter circuits becomes, which grows quadratically as stated above. This implies that the superiority of the compressed circuits over Trotter circuits becomes especially apparent when we set a relatively high error threshold, which for the compressed circuits is reached at much larger time than for Trotter circuits which have comparable gate count.

Repeating this analysis for the ladder, again starting off with $L = 12$ and $t = 1$, we find at $\hat{\epsilon} = 10^{-2}$ that the compressed circuit has $\hat{t} = 34$ whereas the best Trotter circuit, excluding the fourth-order Trotter with double the gate count, has $\hat{t} = 14$. With $\hat{\epsilon} = 10^{-1}$ the compressed circuit has $\hat{t} = 125$ whereas the second-order Trotter circuit has $\hat{t} = 57$. For the same system at $t = 2$ and with $\hat{\epsilon} = 10^{-1}$, we have $\hat{t} = 40$ for the compressed circuit and $\hat{t} = 10$ for the second-order Trotter circuit. Hence for the ladder we can go roughly two to four times as far than the best Trotter circuit with comparable gate count. Repeating this analysis for $L = 8$ we find that we can go five to two times farther, and for $L = 16$ we can go three to two times farther, again with the lower bounds for $t = 1$ and the upper bounds for $t = 2$.

Instead of examining the stacking behavior of compressed and Trotter circuits with comparable gate count, we now compare how circuits with comparable optimized infidelity stack, to see whether similar fidelities are achievable with compressed circuits that have only a fraction of the gates of Trotter circuits. To this end we consider the chain and ladder for a single system size $L = 12$, with time-step $t = 2$ for the chain and $t = 1$ for the ladder, and we stack the circuits up to $t = 20$. For simplicity we compare only with second-order Trotter circuits, as we find analogous results for the other Trotter orders. For the chain we take compressed circuits with $M = 4, 8$, in which case the second-order Trotter circuits with similar optimized infidelity have $M = 5, 16$. Imporantly, while these compressed and Trotter circuits have similar fidelity, the $M = 5$ Trotter circuit has 1.4 times the gate count of the $M = 4$ compressed circuit, whereas the $M = 16$ Trotter circuit has 2.1 times the gate count of the $M = 8$ compressed circuit. For the ladder we take compressed circuits with $M = 8, 16$, such that the corresponding second-order Trotter circuits have $M = 2, 4$, i.e. they contain 1.6 times as many gates.

The results are displayed in Fig. 6 in log-log scale, where in the left panel we show the stacked infidelities for the chain and in the right panel for the ladder. The red dashed lines are for the power laws $\epsilon \propto t^n$ with the best fitting power $n$. It is seen that the infidelity increases similarly for all considered pairs of compressed and Trotter circuits, which like Fig. 5 emphasizes that the compression strategy expounded in Sec. 2 has no drawbacks at long times, relative to the Trotter circuits. Moreover, the mentioned discrepancy in gate counts, with in all cases the Trotter circuit having significantly more gates, makes the compressed

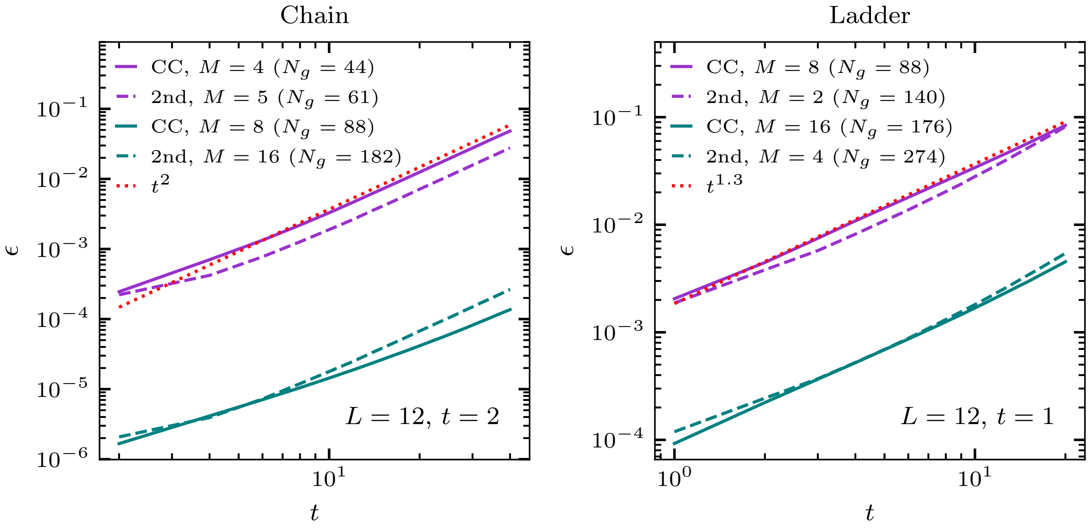


Figure 6: The infidelity $\epsilon$ versus stacking time $t$ for the time evolution operator of the $L = 12$ Heisenberg model on a chain at $t = 2$ (left panels) and ladder at $t = 1$ (right panels), for compressed and second-order circuits that are stacked twenty times. The circuits were chosen such they have similar $\epsilon$ at the optimized $t$, with $M = 4, 8$ and $M = 5, 16$ for compressed and second-order Trotter circuits on the chain, and $M = 8, 16$ and $M = 2, 4$ for the ladder. As a result the compressed circuits have significantly lower gate count than the corresponding Trotter circuits. The red dashed lines are for the power laws $\epsilon \propto t^n$ with the best fitting power $n$.

circuits especially favorable for simulation on real quantum devices, where the error due to gate imperfections and decoherence noise hampers time evolution.

## 3.2 Out-of-time-ordered correlators

Having studied the infidelity and its behavior under stacking in detail in Sec. 3.1, we now use the compressed circuits to determine the behavior of a quantity that does not enter the objective function (19), namely the OTOC (22).

In Fig. 7 we show the absolute $C_{i=2,j}(t)$ errors, relative to the targeted time-evolution operator, for compressed circuits which were optimized for $L = 8, 12, 16$ chains and ladders at $t = 2$ and stacked up to ten times, along with the errors for Trotter circuits with gate counts similar to these compressed circuits. For the chain we let $j$ run over all sites, whereas for the ladder it runs over all rungs. The upper three rows are for the chain while the lower three rows are for the ladder. The first and fourth row are for $L = 8$, the second and fifth row are for $L = 12$, and the third and sixth row are for $L = 16$. The left column is for the compressed circuit while the second, third and fourth columns are for the first-, second- and fourth-order Trotter circuits. As in Fig. 5 the depths are $M = 8, 8, 7, 1$ for the chain and $M = 16, 4, 3, 1$ for the ladder, for the compressed circuit and first-, second- and fourth-order Trotter circuits, respectively.

For the chain it is clear that the compressed circuit works better than the Trotter circuits within the lightcone, whereas it is slightly worse than the second- and fourth-order Trotter circuits at approximating the small values outside of the lightcone. For the ladder the compressed circuit is better everywhere, even better than the fourth-order Trotter circuit which has twice as many gates. Hence we draw the same conclusion as from Fig. 5: With a similar amount of gates we are able to go farther in time with the compressed circuits than with the

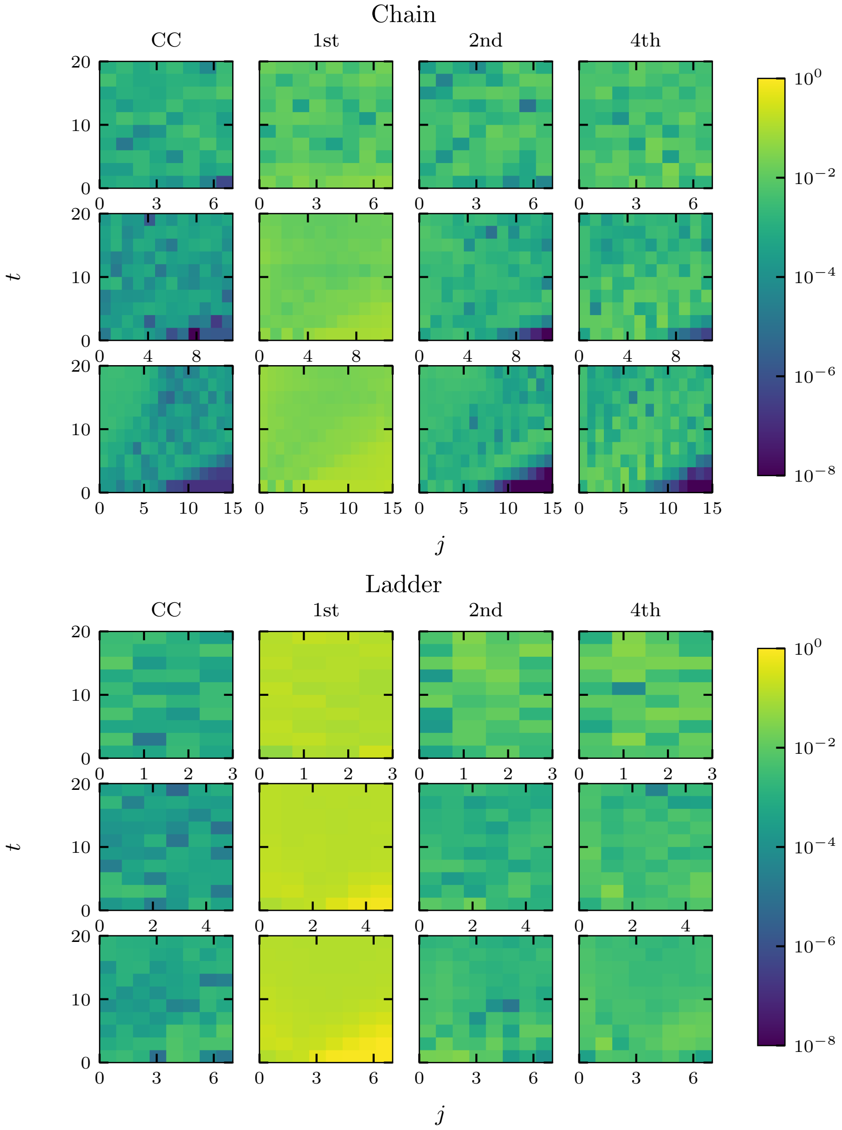


Figure 7: The absolute $C_{i=2,j}(t)$ errors for the chain (top three rows) and ladder (bottom three rows) for a compressed circuit optimized at $t = 2$ and stacked up to ten times, along with the errors for Trotter circuits with similar gate counts. For the chain $j$ labels the sites and for the ladder it labels the rungs. The first and fourth row are for $L = 8$ with $\chi = 256$, the second and fifth row are for $L = 12$ with $\chi = 150$, and the third and sixth row are for $L = 16$ with $\chi = 100$. The first column is for the compressed circuit, the second, third and fourth columns are for the first-, second- and fourth-order Trotter circuits. To have roughly equal gate counts, the used depths are $M = 8, 8, 7, 1$ for the chain and $M = 16, 4, 3, 1$ for the ladder, for the compressed circuit and first-, second- and fourth-order Trotter circuits, respectively.

Trotter circuits, before reaching some error threshold, even though we do not optimize based on OTOCs.

In Sec. C we show the OTOC values corresponding to the errors from Fig. 7, for compressed circuits and the targeted time-evolution operators. There we also show how the relative error of $C_{i=2,j=4}(t)$ propagates with stacking, for compressed and Trotter circuits that have similar optimized fidelity, indicating that we can maintain similar fidelity with compressed circuits that have a fraction of the amount of gates of the Trotter circuits.

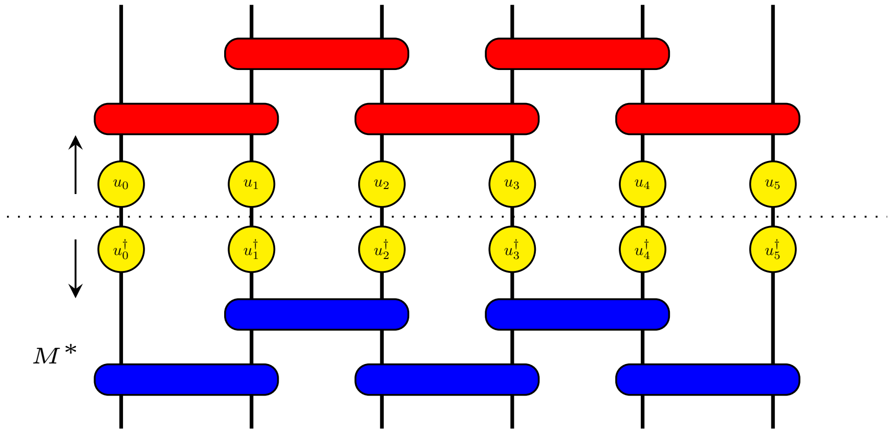


Figure 8: The gauge freedom that exists between the layers of a circuit. When we cut the circuit across the horizontal dashed line, and want to use the lowest $M^*$ layers to calculate an infidelity, we have to take into account the gauge freedom that is encoded by inserting a pair of conjugate one-qubit unitaries $u_i^\dagger u_i = I$ at each qubit, and absorbing one unitary upwards and the other downwards.

### 3.3 Analysis of the compressed circuit

In the previous Sections 3.1 and 3.2 we have seen that the compressed circuit outperforms the Trotter circuits. Here we investigate how this is achieved, by probing the structure of the layers and gates that make up the compressed and Trotter circuits.

Starting off, we take a compressed circuit and Trotter circuits with comparable gate counts, and consider the infidelity between a subset of layers $M^* < M$ (counting from the bottom layer) and the time evolution operator at a time $t^* < t$ that is smaller than the time-step $t$ at which the compressed circuit was optimized. Crucially, we must take into account the gauge freedom that exists between layers, where we are able to insert conjugate layers of one-qubit unitaries, and absorb one layer into the subset we are considering and the other layer into its complement. This process is illustrated in Fig. 8. Hence when calculating a subset infidelity for the compressed circuit, we add a layer of one-qubit unitaries between the subset and the time evolution operator at $t^*$, and minimize the infidelity with respect to these one-qubit unitaries. This way we account for the gauge freedom.

In Fig. 9 we show the results for the chain with $L = 8$ at $t = 1$, for a compressed circuit with $M = 8$ and Trotter circuits with $M = 8, 7, 1$ for first-, second- and fourth-order, which have gate counts close to that of the compressed circuit. Here we define a Trotter circuit with $M^*$ layers as having $M^*$ brickwall layers, and the largest shown $M^*$ is the full circuit, which e.g. for the second-order Trotter circuit involves adding half a brickwall layer to its largest subset. For the compressed circuit $M^* = 8$ corresponds to the full circuit. The dashed lines mark the times $t^* = tM^*/8$.

From Fig. 9 it is clear that at $t = 1$ there is significant overlap of the subsets with a time evolution operator at $t^* < t$ for both the compressed and Trotter circuits. However, in contrast to the first- and second-order Trotter circuits, where the infidelity dips are equidistant, and where for the first-order Trotter circuit the dip depth is decreasing with the number of stacked layers while for the second-order Trotter circuit it is constant, the dips of the compressed circuit are instead roughly symmetric and are smallest around $t^* \approx t/2$. A closer look reveals that

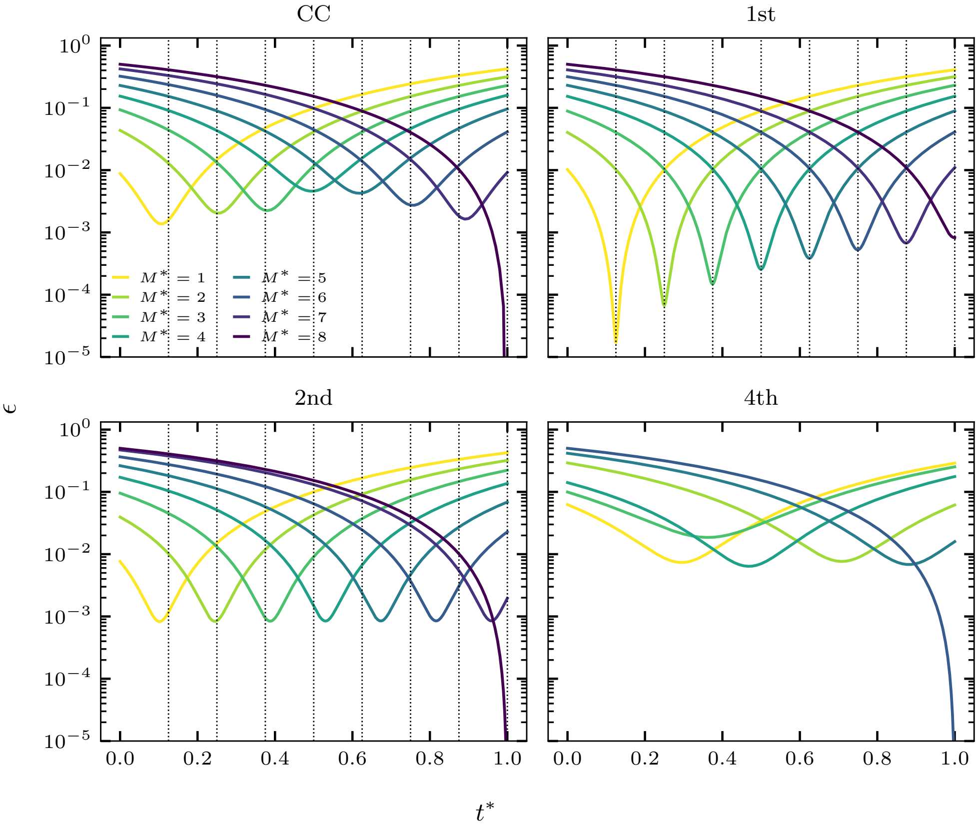


Figure 9: The infidelity between a subset of layers $M^* < M$, counting from the bottom layer, and the targeted time evolution operator at time $t^* < t$, where $t$ denotes the time-step at which the compressed circuit was optimized. The plots are for a Heisenberg chain with $L = 8$ at $t = 1$. In the top left panel we show the results for a compressed circuit with $M = 8$, in the top right for a first-order Trotter circuit with $M = 8$, in the bottom left for a second-order Trotter circuit with $M = 7$, and in the bottom right for a fourth-order Trotter circuit with $M = 1$. These depths were chosen such that the circuits have similar gate count. The curve with $M^* = M$ corresponds to the full circuit. The dashed lines mark times $tM^*/8$.

the infidelity at this point is roughly $10^{-2}$, which is more than one order of magnitude larger than for the first- and second-order Trotter circuit at similar $t^*$. This is even more remarkable when taking the final infidelity into account, which is $\epsilon = 1.8 \cdot 10^{-9}$ for the compressed circuit and therefore at least three orders of magnitudes better than the first-, second- and fourth-order Trotter circuits, which have $\epsilon = 8.2 \cdot 10^{-4}, 1.2 \cdot 10^{-6}, 2.1 \cdot 10^{-6}$. This indicates that the compressed circuit does not follow the target "trajectory" given by the unitary time evolution, but slightly deviates from it. However, it becomes "refocused" at $t^* = t$, which we sketch in Fig. 10. It is an interesting question for future research to understand the alternative trajectory, which might be beneficial for an optimal discretization of time evolution beyond the Trotter decomposition. In Sec. C we show that the refocussing also occurs for the OTOCs.

We note that we did not find these symmetric dips for all our compressed circuits, especially for larger $t$ and the ladder geometry. It remains an open question whether this is an artefact of the convergence of the optimization to a non-global minimum.

As a further comparison between compressed and Trotter circuits, we calculate the operator entanglement entropy (opEE) of their gates [47, 56]. Concretely, we take an optimized compressed circuit $\mathcal{C}$ and decompose each two-qubit gate $U_{ij} \in \mathcal{C}$ using a singular value de-

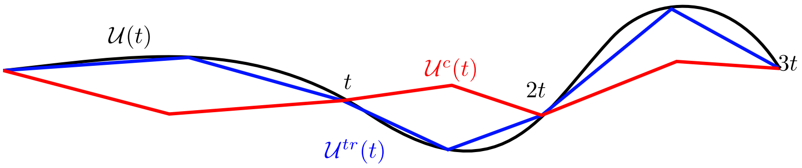


Figure 10: A sketch of the "refocussing" mechanism that potentially explains the structures observed in Fig. 9. Here the targeted time evolution $\mathcal{U}(t)$ is shown in black, the Trotter evolution $\mathcal{U}^{tr}(t)$ is shown in blue, and the compressed evolution $\mathcal{U}^c(t)$ is shown in red. While $\mathcal{U}^{tr}(t)$ follows the target trajectory quite closely, $\mathcal{U}^c(t)$ instead becomes "refocussed" at multiples of the optimization timestep $t$.

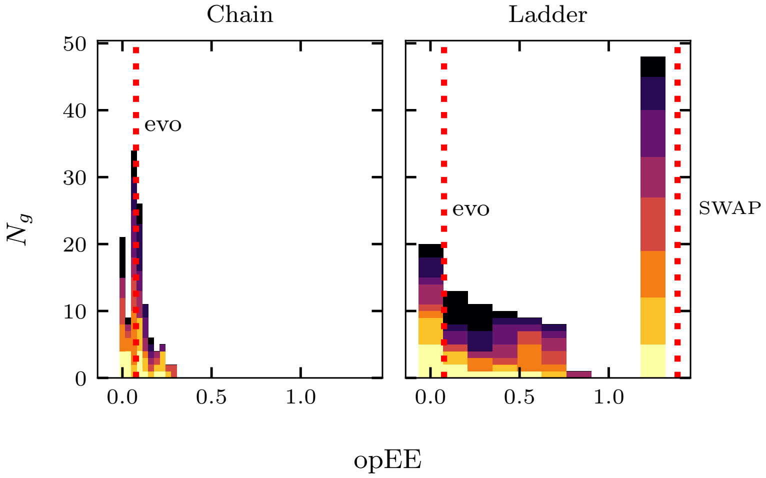


Figure 11: Stacked histograms for the opEE of the gates of a compressed circuit with depth $M = 8$, optimized at $t = 2$ for a $L = 16$ chain (left panels) and ladder (right panels). The colors denote the contents of each layer, with the lightest color for the bottom layer and the darkest for the top layer. The red vertical lines denote the values for the gates in a $M = 8$ first-order Trotter circuit, with the two lines in the ladder plots corresponding to the evolution and SWAP gates.

composition into

$$U_{ij} = \sum_{l=1}^{4} s_l v_i^l \otimes v_j^l \,, \tag{24}$$

where $v_i^l$ and $v_j^l$ are two sets of four one-qubit operators, acting on qubit $i$ and $j$ respectively, and where the four singular values $s_l$ encode the opEE of $U_{ij}$ as

$$\text{opEE} = -\sum_{l} s_l^2 \ln(s_l^2) \,. \tag{25}$$

In Fig. 11 we display the opEE of all gates in a $M = 8$ compressed circuit for the chain (left panel) and ladder (right panel) for $L = 16$ at $t = 2$. The histograms are stacked, with each color denoting the content of a layer, where the lightest color represents the bottom layer and the darkest color the top layer. The red vertical lines mark the values for the $M = 8$ first-order Trotter circuit, with the two lines in the ladder plots corresponding to the evolution and SWAP gates. These histograms show that the gates of the compressed circuit are more hetergenous compared to those of the Trotter circuits, since they have a relatively large spread in opEE instead of one or two values. Moreover, for the ladder it is seen that a several gates in the

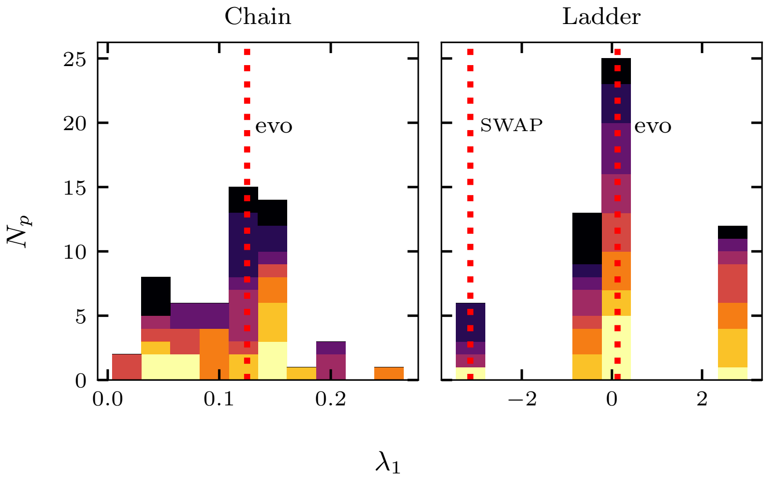


Figure 12: The distribution of the $\lambda_1$ parameter which enters the two-qubit unitary parameterization that was used in this work, shown for the chain (left panel) and ladder (right panel) with $L = 8$ at $t = 1$. The parameter count $N_p$ for a compressed circuit with $M = 8$ is shown as a stacked histogram, with the lightest color corresponding to the bottom layer and the darkest color to the top layer. The first-order Trotter evolution gate value $\lambda_1^{\text{evo}} = t/M$ and the SWAP gate value $\lambda_1^{\text{SWAP}} = -\pi$ are shown as dashed red lines. The other two-qubit parameters $\lambda_2$ and $\lambda_3$ are distributed similarly. Note the different scales of the x-axes.

compressed circuit assume an opEE that is near to that of the SWAP gate, which we view as an indication that the action of the SWAP gate is baked into our optimized circuits.

Finally we consider the distribution of the parameter $\lambda_1$ across the optimized two-qubit unitaries, which are parameterized as in (14). We found that $\lambda_2$ and $\lambda_3$ are distributed similarly. In Fig. 12 we show histograms for the parameter counts $N_p$ of $\lambda_1$ for the chain (left panel) and ladder (right panel) with $L = 8$ at $t = 1$, for a compressed circuit with $M = 8$. Note here the different scales of the x-axes. The histograms are again stacked, with the lightest color corresponding to the bottom layer and the darkest color to the top layer. The red dashed lines mark the values of the gates in the $M = 8$ first-order Trotter circuit, for which $\lambda_1^{\text{SWAP}} = -\pi$ and $\lambda_1^{\text{evo}} = t/M$, both having no one-qubit dressing (15). As in Fig. 11, we see that the gates of the compressed circuit have a larger spread than the gates of the Trotter circuit, which instead assume one or two values. Also, for the ladder we again observe an accumulation of gates near the SWAP value.

The gates appearing in the optimized circuits appear to encode more structure than gates from Trotter circuits and are generally speaking encoding a larger change of the wave function per gate compared to the case of Trotter circuits. This can be seen best in the limit of very small Trotter time steps, in which each appearing gate (except SWAP) is very close to identity, while in the opposite limit which we optimize for, each gate needs to be sufficiently different from identity in order to represent the same time evolution operator.

## 4 Conclusion and Outlook

In this work we have presented an approach which reduces the resource cost of digital quantum simulation compared to standard Trotter decompositions by globally optimizing a simple parameterized brickwall circuit in a way that is scalable to large systems. Crucially, the performance per gate is better even when the compressed circuit does not respect the connectivity of the simulated lattice, potentially allowing for high fidelity simulation of systems with a connectivity that is larger than that of the used quantum processor. To illustrate this we have compared the infidelity of the compressed and Trotter circuits with the targeted time evolution operators of Heisenberg chains and ladders, as well as the ability to reproduce their OTOCs.

We have shown that we can achieve similar accuracy of the time evolution operator with up to one order of magnitude less gates, depending on the desired accuracy and system. Moreover, we checked that this advantage persists when stacking the circuits many times, a central ingredient to simulating a quantum system over long times. This enables high fidelity propagation to times which are currently elusive with conventional Trotter decomposition methods.

Furthermore, we analyzed the structure of the compressed circuits. In the case of the chain, we observed a “refocussing” mechanism, which suppresses the infidelity at multiples of the optimized time step, while the evolution inside the optimized circuit appears to follow a trajectory which is further away from the targeted time evolution operator. It is an interesting question for further research to understand this trajectory and relate it also to recent studies of Trotter decompositions and its breakdown for large time steps [34, 35].

Our results open the door for many further directions. As a next step, one can for example take symmetries into account to further reduce the number of parameters. This might be especially favorable when exploiting translation symmetries. Furthermore, one can optimize the circuits with other cost functions than the fidelity, as was also done for example in [37]. Promising directions are using local observables or density matrices. While such an approach might simplify the convergence of the optimization, it is still an open question to what extent the accurate simulation of observables or other general quantities would be recovered.

We end by stressing that in this work we have used the simplest possible noise model, by assuming that each applied gate introduces the same amount of noise to the system and that therefore a minimization of the gate count reduces the overall noise. A refinement of this noise model will be the subject of future research.

## Acknowledgments

We thank Luis Colmenarez for useful comments on the manuscript. D.H. thanks Adam Smith, Frank Pollmann, and Hongzheng Zhao for useful discussions.

**Funding information** This project was supported by the Deutsche Forschungsgemeinschaft (DFG) through SFB 1143 (project-id 247310070) and the cluster of excellence ML4Q (EXC 2004, project-id 390534769). We also acknowledge support from the QuantERA II Programme that has received funding from the European Union’s Horizon 2020 research innovation programme (GA 101017733), and from the Deutsche Forschungsgemeinschaft through the project DQUANT (project-id 499347025).

## A Gate count equations

Here we state the equations for the NN two-qubit gate counts $N_g$ of the first-, second- and fourth-order Trotter circuits of depth $M$, which are used in Sec. 3. These are denoted by $N_{g\mathfrak{c}/\mathfrak{l}}^{\text{1st}}(M)$, $N_{g\mathfrak{c}/\mathfrak{l}}^{\text{2nd}}(M)$ and $N_{g\mathfrak{c}/\mathfrak{l}}^{\text{4th}}(M)$, respectively, where $\mathfrak{c}$ corresponds to the chain and $\mathfrak{l}$ to the triangular ladder. In deriving these equations we made maximal use of the ability to combine gates in subsequent Trotter steps. The compressed circuits have gate count $N_{g\mathfrak{c}}^{\text{1st}}$.

For the chain the equations are

$$N_{g\mathfrak{c}}^{1\mathrm{st}}(M) = M(L-1)\,, \tag{A.1}$$

$$N_{g\mathfrak{c}}^{2\mathrm{nd}}(M) = M(L-1) + \left\lfloor \frac{L}{2} \right\rfloor, \tag{A.2}$$

$$N_{g\mathfrak{c}}^{4\mathrm{th}}(M) = 5M(L-1) + \left\lfloor \frac{L}{2} \right\rfloor. \tag{A.3}$$

For the ladder, in which case we have to take into account the SWAP gates, the corresponding equations are

$$N_{g\mathfrak{l}}^{1\mathrm{st}}(M) = M\,(4L-7)\,, \tag{A.4}$$

$$N_{g\mathfrak{l}}^{2\mathrm{nd}}(M) = 2M\left(N_{g\mathfrak{l}}^{1\mathrm{st}}(1)+1\right) - (3M-1)\left\lfloor \frac{L}{2} \right\rfloor, \tag{A.5}$$

$$N_{g\mathfrak{l}}^{4\mathrm{th}}(M) = 5MN_{g\mathfrak{l}}^{2\mathrm{nd}}(1) - (5M-1)\left\lfloor \frac{L}{2} \right\rfloor. \tag{A.6}$$

## B Convergence of the optimization

In order to find the optimal compressed circuit using the gradient descent method outlined in Sec. 2.3, it is important to scan the hyperparameter space of the used optimizer. The reason is that there is no single set of hyperparameters which finds the best solution for all optimization problems. We find the best convergence by using the vanilla Adam optimizer [52], which is presented in Algorithm 1.

We scan the hyperparameter space $(\lambda, \delta, \beta_1, \beta_2)$ for the most favorable convergence properties. As mentioned in Sec. 2.3, it is crucial to continue iterating the algorithm when we

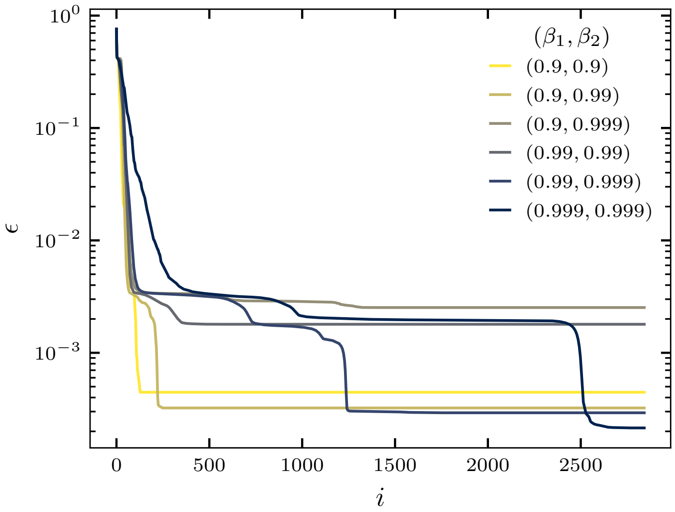


Figure 13: The infidelity $\epsilon$ as a function of the iteration step $i$ for an Adam optimizer with learning-rate $\tau = 0.01$, regularization $\delta = 10^{-4}$, and various decay rates $(\beta_1, \beta_2)$ with $\beta_1, \beta_2 \in \{0.9, 0.99, 0.999\}$. The optimization is performed for a size $L = 8$ ladder at time $t = 1$ with circuit depth $M = 8$. The lowest infidelity is reached with $(0.999, 0.999)$, but crucially this requires the optimizer to spend time in local minima without getting stopped by a convergence criterion when the infidelity has stagnated.

**Algorithm 1** *Vanilla Adam [52].* This gradient-descent optimizer updates the circuit parameters $\vec{\theta}$ to minimize the infidelity $\epsilon(\vec{\theta})$, by taking into account exponentially decaying running averages of the first moment $m$ and second moment $v$ of the infidelity gradient $g$ for each parameter separately. Instead of choosing the parameter updates to be proportional to $g$, as in vanilla gradient descent, here it is proportional to a memory of the previous gradients $m$. This results in a relatively stable minimization and to some extent prevents getting stuck in local minima. Moreover, since the optimization algorithm is first order, the magnitude of the parameter update is proportional to its uncertainty in decreasing the infidelity. For this reason, large updates are undesirable, whereas tiny updates are also undesirable since they halt the minimization and promote getting stuck in local minima. With this in mind, the update magnitude is forced to be desirable, by choosing it to be proportional to $m/\sqrt{v}$.

```
Hyperparameters:
    $\lambda$: Raw learning-rate
    $\beta_1$: First moment decay strength
    $\beta_2$: Second moment decay strength
    $\delta$: Regularization
    $N_{\text{iters}}$: Amount of iterations
Initial conditions:
    $m_0 \leftarrow 0$ (First moment initially zero)
    $v_0 \leftarrow 0$ (Second moment initially zero)
for ( $i = 0$; $i < N_{\text{iters}}$; i = i+1 ) do
    $g_i \leftarrow \nabla_{\vec{\theta}_{i-1}} \epsilon(\vec{\theta}_{i-1})$ (Calculate gradient at current parameters)
    $m_i \leftarrow \beta_1 m_{i-1} + (1-\beta_1) g_i$ (Extend running average of first moment)
    $m_i^* \leftarrow m_i/(1-\beta_1^i)$ (Bias correction)
    $v_i \leftarrow \beta_2 v_{i-1} + (1-\beta_2) g_i^2$ (Extend running average of second moment)
    $v_i^* \leftarrow v_i/(1-\beta_2^i)$ (Bias correction)
    $\vec{\theta}_i \leftarrow \vec{\theta}_{i-1} - \lambda m_i^*/(\sqrt{v_i^*} + \delta)$ (Update parameters)
end for
return $\vec{\theta}_i$ (Final circuit parameters)
```

reach a plateau in the fidelity. This is illustrated in Fig. 13, where we display the gradient descent of $\epsilon$ for a circuit with $M = 8$ layers on the time evolution operator of an $L = 8$ ladder at $t = 1$, and consider various $(\beta_1, \beta_2)$ with learning-rate $\tau = 0.01$ and regularization $\delta = 10^{-4}$. Here the largest fidelity is obtained with $\beta_1 = \beta_2 = 0.999$, but we have to overcome multiple plateaus, which would have been spoiled by using a convergence criterion.

## C OTOC details

First we display the OTOC values of the stacked compressed circuits and targeted time-evolution operators that were used to make Fig. 7. In the left two panel columns of Fig. 14 we show the OTOCs $C_{i=2,j}(t)$ for the chain and in the right two columns for the ladder. The first and third columns are for the compressed circuits, whereas the second and fourth columns are for the target unitaries. As already became apparent from Fig. 7, the agreement is excellent for all considered stacking times $t$.

Now we consider the analog of Fig. 6 for the relative error of the OTOC $C_{i=2,j=4}$. In particular, we consider the chain and ladder with $L = 12$ and take a couple compressed circuits for which the infidelities were optimized at $t = 2$ for the chain and $t = 1$ for the ladder, which we then stack up to $t = 20$. As in Fig. 6 we take compressed circuits with $M = 4, 8$ for the

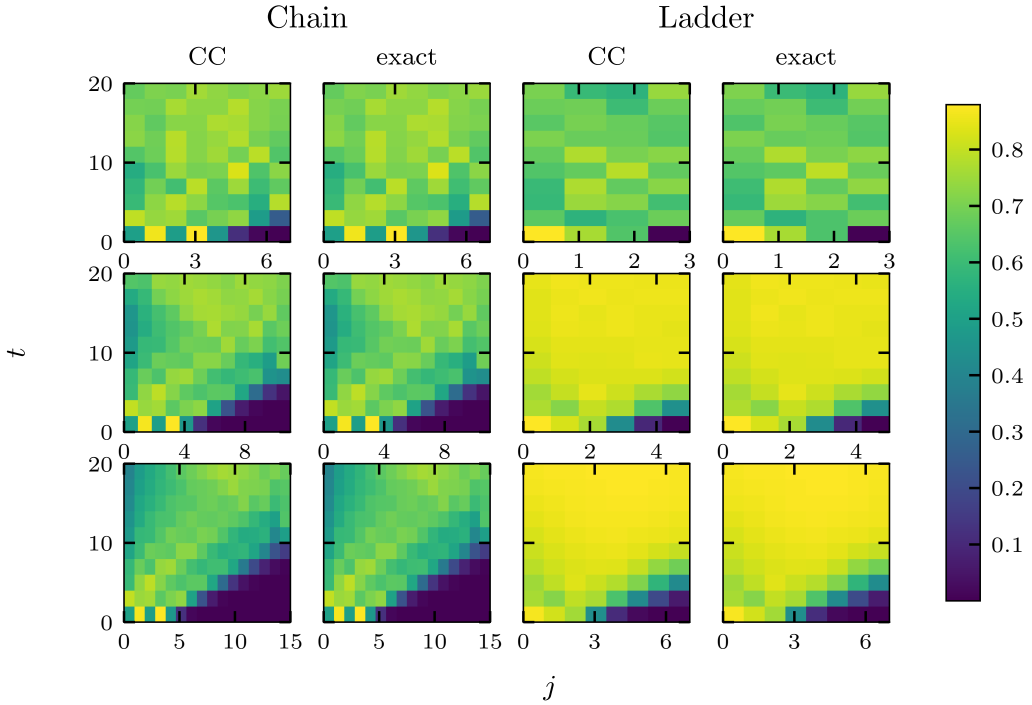


Figure 14: The OTOCs $C_{i=2,j}(t)$ as a function of site or rung $j$ and stacking time $t$ for the chain (left columns) and the ladder (right columns), for compressed circuits optimized at $t = 2$ and stacked up to ten times (first and third columns) and the corresponding target values (second and fourth columns). For the chain we take $M = 8$ and for the ladder $M = 16$. The top row is for $L = 8$ with $\chi = 256$, the middle row is for $L = 12$ with $\chi = 150$, and the bottom row is for $L = 16$ with $\chi = 100$.

chain and $M = 8, 16$ for the ladder, and we compare these with second-order Trotter circuits that have similar fidelity at the optimized time step, corresponding to $M = 5, 16$ for the chain and $M = 2, 4$ for the ladder. In Fig. 15 we show the results, with the left panel for the chain and the right panel for the ladder. The implications are the same as those derived from Fig. 6: With a smaller amount of gates we essentially get the same performance, in this case even for a quantity that does not appear in the objective function (19).

Finally, we check whether the refocussing that was observed for the infidelity $\epsilon$ in Fig. 9 also emerges for the OTOCs, which contrary to $\epsilon$ does not enter the cost function of the optimization scheme. In Fig. 16 we show the relative error of $C_{i=5,j=5}(t^*)$ between that of $M^*$ layers and that of the target unitary at time $t^*$. Before using the $M^*$ layers to calculate the OTOC at $t^*$, we minimize its infidelity with respect to the target unitary at $t^*$, taking into account the gauge invariance. As in Fig. 9, we perform the calculations for the Heisenberg chain with $L = 8$ and a $M = 8$ circuit optimized at $t = 1$, with the results shown in Fig. 16. We see that a similar refocussing takes place, with the minima for $M^* < M$ being elevated with respect to that at $M^* = M$ and with unequal spacing in time.

## D Emergence of lattice symmetries

The brickwall circuit ansatz (18) used in this work has the most general form, consisting of arbitrary two-body unitaries and not taking into account any symmetry of the targeted time-evolution operator, i.e. in our case those corresponding to the Hamiltonians (2). To restrict

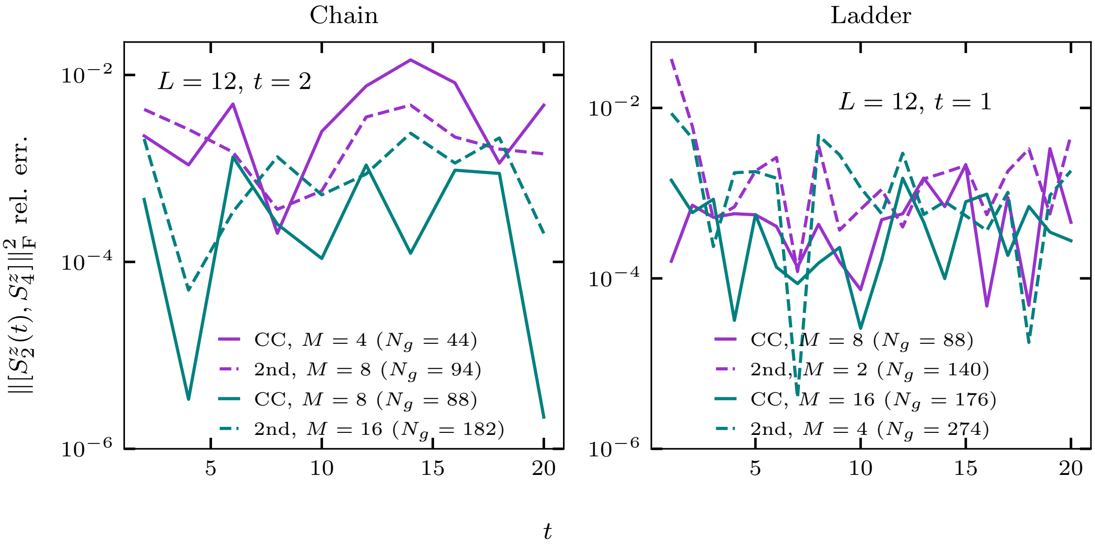


Figure 15: The relative error of the OTOC $C_{i=2,j=4}(t)$ versus stacking time $t$ for the chain (left panel) and ladder (right panel) with $L = 12$, for circuits optimized at $t = 2$ for the chain and $t = 1$ for the ladder, and stacked up to time $t = 20$. For the chain we consider $M = 4, 8$ and for the ladder $M = 8, 16$. For each $M$ we choose a second-order Trotter circuit with similar fidelity at the optimized $t$, i.e. $M = 5, 16$ for the chain and $M = 2, 4$ for the ladder. As a result the compressed circuits have significantly less gates than the corresponding Trotter circuits.

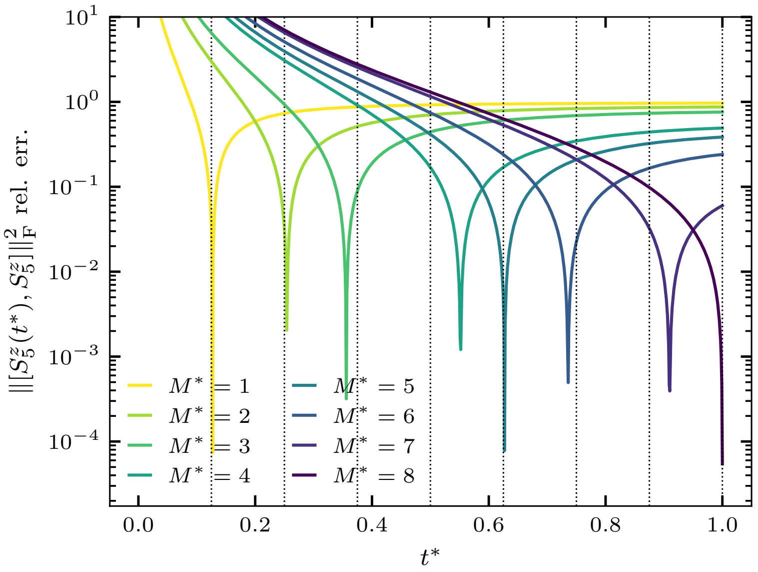


Figure 16: The relative error of the OTOC $C_{i=5,j=5}(t^*)$ between that of a subset $M^*$ of layers and that of the target unitary at time $t^*$, for the Heisenberg chain at $L = 8$ and a depth $M = 8$ circuit optimized at time $t = 1$. The curve with $M^* = M$ corresponds to the full circuit and the dashed lines mark times $tM^*/8$.

the ansatz space it could be useful to incorporate such symmetries into the circuit at the gate level.

Take for example the Heisenberg chain in (2), which possesses lattice inversion symmetry,

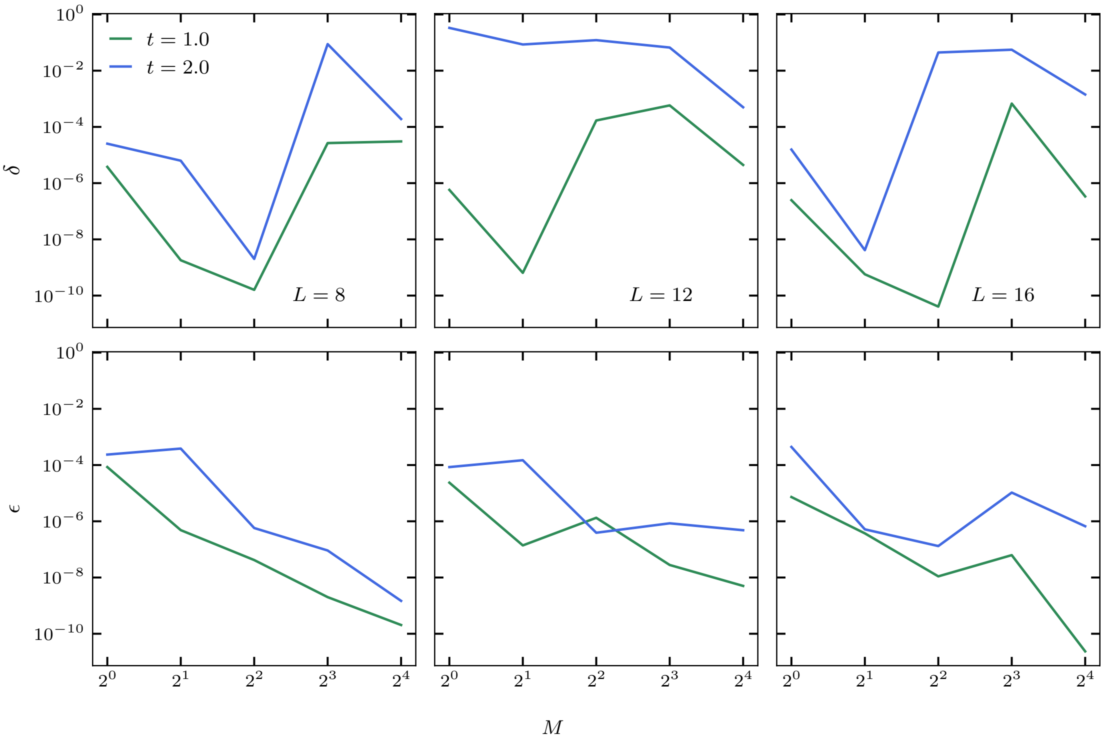


Figure 17: The average gate-wise infidelity $\delta$ of every gate with its mirrored counterpart (top panels), flipped across the middle bond, for all gates in compressed circuits which were optimized to approximate the lattice inversion symmetric Heisenberg chain time-evolution operator. For comparison, we also show the infidelity $\epsilon$ of the circuit as a whole with its mirrored counterpart (bottom panels). These quantities probe to which extent the inversion symmetry of the targeted unitary emerges in the compressed circuit. The infidelities are shown as a function of the circuit depth $M$, for times $t = 1$ and $t = 2$. The left panels are for system size $L = 8$, the middle panels are for $L = 12$, and the right panels are for $L = 16$.

being invariant under a flip of the lattice across the middle bond for even $L$. To incorporate this into the ansatz we let the gate acting on the bond between sites $i$ and $i+1$ also act on the mirrored bond between $L-2-i$ and $L-1-i$, albeit flipped across the time axis. Since this gate and its flipped counterpart should be equal for the inversion symmetry to be manifest, the gate parameterization (13) implies that the one-qubit unitary $u_i$ should be equal to $u_j$, and that $u_i'$ should be equal to $u_j'$, with the two-qubit unitary $v_{ij}$ being flip-symmetric by construction.

Since we did not incorporate this inversion symmetry into the circuits used for our simulations, it is an interesting question whether the chosen circuit ansatz in combination with the optimization procedure leads to its emergence. To probe this, we take an optimized circuit and for each of its gates we calculate the infidelity with its mirrored counterpart, and then average over all gates to get the average gate infidelity $\delta$. As for the subset infidelity from Fig. 9, here it is crucial to take into account the gauge symmetry. We also calculate the infidelity $\epsilon$ of the circuit as a whole with its mirrored counterpart, to determine if it is reasonable to expect the symmetry to emerge on the gate level. If this overall infidelity is high, it is unlikely that it is low at the gate level. The results are shown in Fig. 17.

## References

[1] A. Y. Kitaev, *Quantum measurements and the Abelian stabilizer problem*, (arXiv preprint) doi:10.48550/arXiv.quant-ph/9511026.

[2] L. K. Grover, *A fast quantum mechanical algorithm for database search*, in *STOC '96: Proceedings of the twenty-eighth annual ACM symposium on Theory of Computing*, Association for Computing Machinery, New York, USA, ISBN 9780897917858 (1996), doi:10.1145/237814.237866.

[3] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Comput. **26**, 1484 (1997), doi:10.1137/s0097539795293172.

[4] S. Ebadi et al., *Quantum optimization of maximum independent set using Rydberg atom arrays*, Science **376**, 1209 (2022), doi:10.1126/science.abo6587.

[5] M. Feld, B. Fröhlich, E. Vogt, M. Koschorreck and M. Köhl, *Observation of a pairing pseudogap in a two-dimensional Fermi gas*, Nature **480**, 75 (2011), doi:10.1038/nature10627.

[6] E. Cocchi, et al., *Equation of state of the two-dimensional Hubbard model*, Phys. Rev. Lett. **116**, 175301 (2016), doi:10.1103/physrevlett.116.175301.

[7] H. Bernien et al., *Probing many-body dynamics on a 51-atom quantum simulator*, Nature **551**, 579 (2017), doi:10.1038/nature24622.

[8] C. Gross and I. Bloch, *Quantum simulations with ultracold atoms in optical lattices*, Science **357**, 995 (2017), doi:10.1126/science.aal3837.

[9] P. Scholl et al., *Quantum simulation of 2D antiferromagnets with hundreds of Rydberg atoms*, Nature **595**, 233 (2021), doi:10.1038/s41586-021-03585-1.

[10] M. C. Bañuls et al., *Simulating lattice gauge theories within quantum technologies*, Eur. Phys. J. D **74**, 165 (2020), doi:10.1140/epjd/e2020-100571-8.

[11] A. Mil, et al., *A scalable realization of local $U(1)$ gauge invariance in cold atomic mixtures*, Science **367**, 1128 (2020), doi:10.1126/science.aaz5312.

[12] B. Yang et al., *Observation of gauge invariance in a 71-site Bose-Hubbard quantum simulator*, Nature **587**, 392 (2020), doi:10.1038/s41586-020-2910-8.

[13] S. Lloyd, *Universal quantum simulators*, Science **273**, 1073 (1996), doi:10.1126/science.273.5278.1073.

[14] C. Neill et al., *Ergodic dynamics and thermalization in an isolated quantum system*, Nat. Phys. **12**, 1037 (2016), doi:10.1038/nphys3830.

[15] Y. Salathé et al., *Digital quantum simulation of spin models with circuit quantum electrodynamics*, Phys. Rev. X **5**, 021027 (2015), doi:10.1103/physrevx.5.021027.

[16] U. L. Heras et al., *Digital quantum simulation of spin systems in superconducting circuits*, Phys. Rev. Lett. **112**, 200501 (2014), doi:10.1103/physrevlett.112.200501.

[17] B. P. Lanyon et al., *Universal digital quantum simulation with trapped ions*, Science **334**, 57 (2011), doi:10.1126/science.1208001.

[18] J. Richter and A. Pal, *Simulating hydrodynamics on noisy intermediate-scale quantum devices with random circuits*, Phys. Rev. Lett. **126**, 230501 (2021), doi:10.1103/PhysRevLett.126.230501.

[19] R. Mansuroglu, T. Eckstein, L. Nützel, S. A. Wilkinson and M. J. Hartmann, *Variational Hamiltonian simulation for translational invariant systems via classical pre-processing* in *2022 IEEE International conference on quantum computing and engineering (QCE)*, Broomfield, USA, ISBN 9781665491136 (2022), doi:10.1109/QCE53715.2022.00155.

[20] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love and M. Head-Gordon, *Simulated quantum computation of molecular energies*, Science **309**, 1704 (2005), doi:10.1126/science.1113479.

[21] A. Kandala, et al., *Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets*, Nature **549**, 242 (2017), doi:10.1038/nature23879.

[22] F. Arute et al., *Hartree-Fock on a superconducting qubit quantum computer*, Science **369**, 1084 (2020), doi:10.1126/science.abb9811.

[23] E. A. Martinez et al., *Real-time dynamics of lattice gauge theories with a few-qubit quantum computer*, Nature **534**, 516 (2016), doi:10.1038/nature18318.

[24] L. Funcke, et al., *Towards quantum simulations in particle physics and beyond on noisy intermediate-scale quantum devices*, Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. **380**, 2216 (2021), doi:10.1098/rsta.2021.0062.

[25] P. Frey and S. Rachel, *Realization of a discrete time crystal on 57 qubits of a quantum computer*, Sci. Adv. **8**, eabm7652 (2022), doi:10.1126/sciadv.abm7652.

[26] X. Mi et al., *Time-crystalline eigenstate order on a quantum processor*, Nature **601**, 531 (2021), doi:10.1038/s41586-021-04257-w.

[27] K. J. Satzinger et al., *Realizing topologically ordered states on a quantum processor*, Science **374**, 1237 (2021), doi:10.1126/science.abi8378.

[28] H. F. Trotter, *On the product of semi-groups of operators*, Proc. Am. Math. Soc. **10**, 545 (1959), doi:10.2307/2033649.

[29] M. Suzuki, *Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems*, Commun. Math. Phys. **51**, 183 (1976), doi:10.1007/BF01609348.

[30] M. Suzuki, *Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics*, J. Math. Phys. **26**, 601 (1985), doi:10.1063/1.526596.

[31] M. Suzuki, *Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations*, Phys. Lett. A **146**, 319 (1990), doi:10.1016/0375-9601(90)90962-N.

[32] M. Suzuki, *General theory of fractal path integrals with applications to many-body theories and statistical physics*, J. Math. Phys. **32**, 400 (1991), doi:10.1063/1.529425.

[33] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe and S. Zhu, *Theory of Trotter error with commutator scaling*, Phys. Rev. X **11**, 011020 (2021), doi:10.1103/PhysRevX.11.011020.

[34] M. Heyl, P. Hauke and P. Zoller, *Quantum localization bounds Trotter errors in digital quantum simulation*, Sci. Adv. **5**, eaau8342 (2019), doi:10.1126/sciadv.aau8342.

[35] C. Kargi, J. Pablo Dehollain, F. Henriques, L. M. Sieberer, T. Olsacher, P. Hauke, M. Heyl, P. Zoller and N. K. Langford, *Quantum chaos and universal trotterisation performance behaviours in digital quantum simulation*, in *Quantum information and measurement*, Optica Publishing Group, Washington DC, USA, ISBN 9781957171012 (2021), doi:10.1364/QIM.2021.W3A.1.

[36] M. Benedetti, M. Fiorentini and M. Lubasch, *Hardware-efficient variational quantum algorithms for time evolution*, Phys. Rev. Res. **3**, 033083 (2021), doi:10.1103/physrevresearch.3.033083.

[37] A. Bolens and M. Heyl, *Reinforcement learning for digital quantum simulation*, Phys. Rev. Lett. **127**, 110502 (2021), doi:10.1103/PhysRevLett.127.110502.

[38] S. Barison, F. Vicentini and G. Carleo, *An efficient quantum algorithm for the time evolution of parameterized circuits*, Quantum **5**, 512 (2021), doi:10.22331/q-2021-07-28-512.

[39] N. F. Berthusen, T. V. Trevisan, T. Iadecola and P. P. Orth, *Quantum dynamics simulations beyond the coherence time on noisy intermediate-scale quantum hardware by variational Trotter compression*, Phys. Rev. Res. **4**, 023097 (2022), doi:10.1103/PhysRevResearch.4.023097.

[40] S.-H. Lin, R. Dilip, A. G. Green, A. Smith and F. Pollmann, *Real- and imaginary-time evolution with compressed quantum circuits*, PRX Quantum **2**, 010342 (2021), doi:10.1103/PRXQuantum.2.010342.

[41] R. Haghshenas, J. Gray, A. C. Potter and G. K.-L. Chan, *Variational power of quantum circuit tensor networks*, Phys. Rev. X **12**, 011047 (2022), doi:10.1103/physrevx.12.011047.

[42] F. Arute et al., *Quantum supremacy using a programmable superconducting processor*, Nature **574**, 505 (2019), doi:10.1038/s41586-019-1666-5.

[43] IBM Quantum, *Compute resources*, https://quantum-computing.ibm.com/services?services=systems, accessed: 2022-04-12.

[44] *AWS bracket*, https://us-east-1.console.aws.amazon.com/braket/home?region=us-east-1#/devices/arn:aws:braket:us-west-1::device/qpu/rigetti/Aspen-M-1, accessed: 2022-04-12.

[45] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck and C. Hubig, *Time-evolution methods for matrix-product states*, Ann. Phys. **411**, 167998 (2019), doi:10.1016/j.aop.2019.167998.

[46] B. Kraus and J. I. Cirac, *Optimal creation of entanglement using a two-qubit gate*, Phys. Rev. A **63**, 062309 (2001), doi:10.1103/PhysRevA.63.062309.

[47] T. Zhou and D. J. Luitz, *Operator entanglement entropy of the time evolution operator in chaotic systems*, Phys. Rev. B **95**, 094206 (2017), doi:10.1103/PhysRevB.95.094206.

[48] G. Vidal, *Efficient classical simulation of slightly entangled quantum computations*, Phys. Rev. Lett. **91**, 147902 (2003), doi:10.1103/PhysRevLett.91.147902.

[49] H.-J. Liao, J.-G. Liu, L. Wang and T. Xiang, *Differentiable programming tensor networks*, Phys. Rev. X **9**, 031041 (2019), doi:10.1103/PhysRevX.9.031041.

[50] D. E. Rumelhart, G. E. Hinton and R. J. Williams, *Learning representations by back-propagating errors*, Nature **323**, 533 (1986), doi:10.1038/323533a0.

[51] S. H. Sack, R. A. Medina, R. Kueng and M. Serbyn, *Transition states and greedy exploration of the QAOA optimization landscape*, (arXiv preprint) doi:10.48550/arXiv.2209.01159.

[52] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, (arXiv preprint) doi:10.48550/arXiv.1412.6980.

[53] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush and H. Neven, *Barren plateaus in quantum neural network training landscapes*, Nat. Commun. **9**, 4812 (2018), doi:10.1038/s41467-018-07090-4.

[54] D. J. Luitz and Y. Bar Lev, *Information propagation in isolated quantum systems*, Phys. Rev. B **96**, 020406 (2017), doi:10.1103/physrevb.96.020406.

[55] K. Hémery, F. Pollmann and D. J. Luitz, *Matrix product states approaches to operator spreading in ergodic quantum systems*, Phys. Rev. B **100**, 104303 (2019), doi:10.1103/physrevb.100.104303.

[56] T. Prosen and I. Pižorn, *Operator space entanglement entropy in a transverse Ising chain*, Phys. Rev. A **76**, 032316 (2007), doi:10.1103/PhysRevA.76.032316.

APPENDIX D

# Article 4

PHYSICAL REVIEW RESEARCH **3**, 023236 (2021)

# Three-dimensional isometric tensor networks

Maurits S. J. Tepaske[*] and David J. Luitz[†]
*Max Planck Institute for the Physics of Complex Systems, Noethnitzer Strasse 38, 01167 Dresden, Germany*

(Received 2 June 2020; revised 3 June 2021; accepted 3 June 2021; published 22 June 2021)

Tensor network states are expected to be good representations of a large class of interesting quantum many-body wave functions. In higher dimensions, their utility is however severely limited by the difficulty of contracting the tensor network, an operation needed to calculate quantum expectation values. Here we introduce a method for the time evolution of three-dimensional isometric tensor networks which respects the isometric structure and therefore renders contraction simple through a special canonical form. Our method involves a tetrahedral site splitting which allows one to move the orthogonality center of an embedded tree tensor network in a simple cubic lattice to any position. Using imaginary time evolution to find an isometric tensor network representation of the ground state of the three-dimensional transverse field Ising model across the entire phase diagram, we perform a systematic benchmark study of this method in comparison with exact Lanczos and quantum Monte Carlo results. We show that the obtained energy matches the exact ground-state result accurately deep in the ferromagnetic and polarized phases, while the regime close to the critical point requires larger bond dimensions. This behavior is in close analogy with the two-dimensional case, which we also discuss for comparison.

DOI: 10.1103/PhysRevResearch.3.023236

## I. INTRODUCTION

The Hilbert space dimension of quantum many-body systems grows exponentially with the number of constituents, making the direct handling of many-body wave functions impractical for large systems. Tensor networks are an attempt to tame the many-body wave function, by expressing it in terms of local tensors, which are contracted according to the network structure. This reduces the complexity from an exponential to a polynomial number of variables. While in principle any wave function can be expressed as a tensor network, some particularly entangled states require exponentially large tensors. Fortunately, the manifold of wave functions expressible with small tensor networks includes wave functions with area law entanglement, which are expected to be relevant for the description of ground states of many local quantum many-body systems [1–3].

Tensor network states are particularly successful in one dimension, where they are known as "matrix-product states" (MPS) [4], which have become state-of-the-art machinery for the classical simulation of one-dimensional (1D) many-body systems. This popularity rests primarily on the existence of powerful algorithms to variationally optimize the energy of the state [e.g., the density matrix renormalization group (DMRG) [5]] and on the ability to compute matrix elements of local operators $\langle\phi|\hat{O}|\psi\rangle$ both exactly and efficiently. In particular the second property does not generalize to the higher-dimensional variants of MPS known as "projected-entangled pair states" (PEPS) [6]. It turns out that the exact calculation of a local correlator in an arbitrary PEPS state $\langle\text{PEPS}|\hat{O}|\text{PEPS}\rangle$—requiring the contraction of a higher-dimensional network—is generally inefficient for generic finite PEPS with open boundary conditions (OBC) already in two dimensions [6–10]. While PEPS are readily formulated in three dimensions (see Fig. 1), currently no efficient contraction method is known. So even though PEPS are efficient representations of area-law entangled quantum many-body wave functions, it is often difficult to extract useful information from them.

The central problem for the generalization of powerful 1D methods to higher dimensions is caused by the fact that cutting a bond in a higher-dimensional PEPS does not separate the network into two disconnected pieces, in contrast to 1D MPS. In MPS methods, the separation of the network into unique "left" and "right" parts by cutting any bond is exploited by using an orthonormal basis to represent the left/right states, and one can then decimate the basis to the dominant components by truncating to the largest singular values in an optimal way [11]. This property is the foundation of MPS evolution algorithms [12]. The absence of such separability in higher-dimensional PEPS diminishes the effectiveness of purely local evolution algorithms, where in the case of a nearest-neighbor interacting system the tensor network is optimized by iterating over the bonds and applying a two-body gate to each bond followed by a truncation of this bond according to the standard time-evolving block decimation (TEBD) [4,11]. Instead, optimal truncation and hence optimal evolution require each gate to be accompanied by a contraction of the full network (dubbed "full update"), which is inefficient as it

---

[*] mtepaske@pks.mpg.de
[†] dluitz@pks.mpg.de

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by the Max Planck Society.*

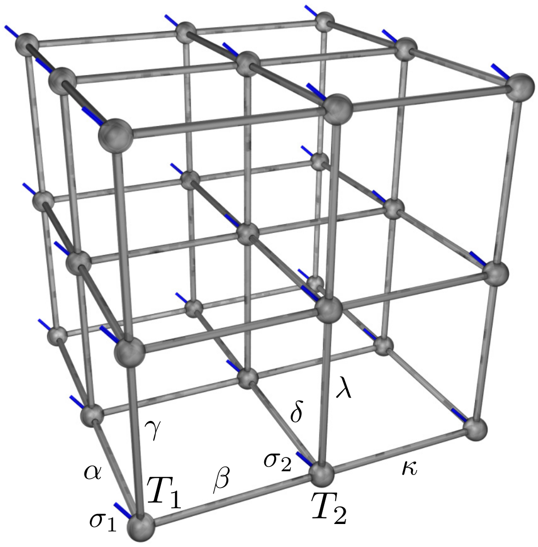


FIG. 1. A generic 3D PEPS ansatz for the cubic lattice, where the tensors $T^{\sigma_i}_{i\alpha\beta\gamma\delta\kappa\lambda}$ are represented by spheres. The blue legs denote the physical degrees of freedom $\sigma_i$ and the gray legs denote the virtual degrees of freedom. The connections depict contractions between the virtual legs of neighboring tensors.

generally scales exponentially with network size when performed exactly [8].

One way around the inefficiency of full contraction is to instead perform the contraction approximately [6,8–10], sacrificing precision for speed. Recently there appeared multiple works [13–16] which suggest an attractive alternative: to construct finite PEPS in an *explicit canonical form* in which it can be contracted both exactly and efficiently in a way that local truncation again becomes optimal just like for MPS, thereby circumventing the mentioned problems that occur when dealing with generic PEPS. This does induce a loss of generality, restricting its subspace in Hilbert space to a subspace of generic PEPS, thereby reducing the expressivity of the network [17]. While the effect of this restriction is not yet clear, it becomes irrelevant in the limit of large bond dimensions and therefore seems at least in principle controllable.

In [13] a class of finite two-dimensional (2D) PEPS called "isometric tensor network states" (isoTNS) was introduced for which $\langle\text{PEPS}\,|\,\text{PEPS}\rangle$ reduces to a canonical MPS norm, and which can be time evolved using an efficient local evolution algorithm called TEBD$^2$. Here we will generalize the isoTNS ansatz to three dimensions and develop an extension of TEBD$^2$ which we call TEBD$^3$. This upgrade to a higher spatial dimension is an important step in developing efficient techniques to simulate generic three-dimensional (3D) quantum many-body systems, especially for cases which are not accessible to quantum Monte Carlo methods due to a sign problem. This importance is stressed by the limited number of existing finite 3D PEPS algorithms [18,19] and generic simulation methods for 3D quantum many-body systems in general [20–22].

## II. METHOD

A generic finite 3D PEPS ansatz for a 3D many-body spin-$1/2$ system can be written in the local basis $\sigma_i = \pm 1$ as

$$|\text{PEPS}\rangle = \sum_{\sigma_1\ldots\sigma_N} \mathcal{C}\big(T_1^{\sigma_1}\ldots T_N^{\sigma_N}\big)|\sigma_1\ldots\sigma_N\rangle, \qquad (1)$$

where $T_i^{\sigma_i}$ represents the set of tensors which contain the complex-valued variational parameters and which are spatially arranged like the spins $\sigma_i$. Here $\mathcal{C}$ indicates that all tensors are contracted, giving complex scalar coefficients, which is usually done by choosing the amount of virtual degrees of freedom per $T_i^{\sigma_i}$ equal to the lattice connectivity and then contracting nearest neighbors. In Fig. 1 this is illustrated for a cubic lattice with open boundary conditions, where the pairs of virtual degrees of freedom are represented by the gray bonds and the physical (spin) degrees of freedom are represented by the blue free legs, i.e., in a particular basis we get the tensors $T^{\sigma_i}_{i\alpha\beta\gamma\delta\kappa\lambda}$.

In order to calculate $\langle\text{PEPS}\,|\,\text{PEPS}\rangle$ we would contract this tensor network with the physical legs of its conjugate, which we would then have to contract down to a scalar.

### A. General properties of isoTNS

The goal of this paper is to design a type of 3D tensor network that allows for the full network contraction to be done exactly and efficiently. To this end we impose an additional structure on the PEPS shown in Fig. 1, such that $\langle\text{PEPS}\,|\,\text{PEPS}\rangle = \langle\text{MPS}\,|\,\text{MPS}\rangle$ becomes manifest. In particular, we choose the majority of the $T_i$ to be isometric, meaning that these $T_i$ reduce to identities when contracted with their conjugate over a subset of the legs, e.g.,

$$\sum_{i\alpha\beta\gamma\delta\kappa} T^{\dagger\alpha\beta\gamma\delta\kappa\lambda}_{\sigma_i} T^{\alpha\beta\gamma\delta\kappa\eta}_{\sigma_i} = \mathbb{1}^{\lambda\eta} \qquad (2)$$

and

$$\sum_{i\alpha\beta\gamma\delta} T^{\dagger\alpha\beta\gamma\delta\kappa\lambda}_{\sigma_i} T^{\alpha\beta\gamma\delta\nu\eta}_{\sigma_i} = \mathbb{1}^{\kappa\nu}\mathbb{1}^{\lambda\eta}. \qquad (3)$$

If $T$ is unitary instead of isometric, these constraints also hold under $T \to T^\dagger$, which in the case of an isometry instead gives a projector.

In the language of tensor network diagrammatics we can represent these constraints by decorating the legs with arrows [4,13], where incoming arrows represent contracted indices in the isometry constraint and where outgoing arrows represent free indices. In Fig. 2 we illustrate this notation for tensors with four virtual legs. Here the diagrams in the upper panels encode the isometry constraints depicted in the lower panels. Specifically, Fig. 2(a) corresponds to Eq. (2) and Fig. 2(b) corresponds to Eq. (3). Note that for convenience we choose a single arrow direction after contracting the legs. In the same spirit we will omit the arrows on physical legs from here on, since we will always choose these to be incoming.

Using these isometric tensors we can construct tensor networks that identically reduce to a single pair of tensors (the "orthogonality center") upon contraction with its conjugate, which are called isoTNS [13]. In Fig. 3 we show a few snapshots of this reduction for a 2D isoTNS on a $3 \times 3$ square

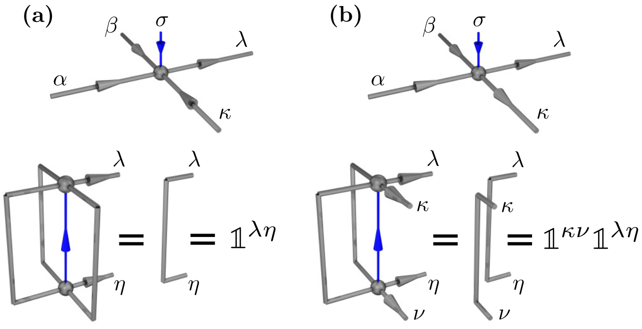


FIG. 2. The isometry constraints are encoded by decorating the tensor legs with arrows (top row). Here panel (a) corresponds to Eq. (2) and panel (b) corresponds to Eq. (3), where in the bottom figures the tensor shown in the top panel is contracted with its conjugate to yield identities.

lattice. Before employing the isometry constraints we have the network depicted in Fig. 3(a). Here we allow the red and gray legs to have distinct bond dimensions $\chi$ and $D$, which is motivated by the observation that for the chosen pattern of arrows we can first reduce the network to a canonicalized MPS amplitude, as shown in Fig. 3(b). Note that the gray legs which stem from the reduction effectively enlarge the local Hilbert space of the MPS.

The distinction between red MPS bonds and gray non-MPS bonds is an important aspect of the isoTNS ansatz, as it will allow us to increase the accuracy of correlators while increasing the bond dimension on only a small part of the tensor network. In Sec. III C we will see that in some situations the accuracy is indeed greatly increased when we increase $\chi$ while keeping $D$ constant. This is especially favorable because we will see in Sec. II B 4 that the computational cost of the time-evolution algorithm scales significantly more favorably in $\chi$ than in $D$.

Due to the remaining isometry structure we can further reduce the MPS amplitude down to a single site, as shown in Fig. 3(c). This final site, which is called the orthogonality center and which is colored red in Fig. 3, therefore fully encodes the norm, just like the orthogonality center of MPS in one dimension [4]. It also encodes the one-body correlators of that site, but if we want to calculate its two-body correlator with another site we can already no longer reduce the network to the orthogonality center. In terms of the arrow language this reduction requires that the arrows flow towards the MPS and within the MPS towards the orthogonality center, which can be characterized as having only incoming arrows.

For example, say we want to calculate a local correlator between the two nearest neighbors of the center in Fig. 3(b). In this case the MPS reduction that yields Fig. 3(c) from Fig. 3(b) would be halted at the operators, since there we can no longer utilize the isometry relations from Fig. 2. Consequently we also cannot utilize the isometric tensors that lie between the operators, meaning that we are left with a MPS correlator instead of a single-site correlator, the accuracy of which is controlled by $\chi$. This is illustrated for $\langle \text{isoTNS}\,|O_1O_2|\text{isoTNS}\rangle$ in Fig. 4.

Clearly the calculation of a MPS correlator is more costly than the single-site correlator, but crucially it still scales polynomially in its bond dimension and size [4]. When we instead consider local correlators with one or more operators located outside of the MPS, the isometry reductions yield a genuine 2D network, which would have to be contracted in order to get the correlator. As a result the correlator can no longer be computed efficiently, illustrating the importance of being able to move the embedded MPS through the network.

It should be noted that by calculating correlators as MPS correlators we do reduce the formal expressibility of the ansatz, as compared to generic PEPS, since it is known that MPS with finite $\chi$ can only encode exponentially decaying correlations [4] whereas generic PEPS can also encode algebraic correlations [23]. This means that the isoTNS ansatz can only encode exponential correlations. Nonetheless, for finite systems the bond dimension can always be chosen large enough to encode algebraic correlations which are cut off by the system size.

Before discussing the time-evolution algorithm, which we will do directly for 3D isoTNS, we consider how Fig. 3 generalizes to three dimensions. First we note that the suitable embedding even for two dimensions is actually a tree tensor network (TTN) with the geometry of a star [24,25], which turns into a MPS when the center occupies a corner as in Fig. 3. When it instead occupies the bulk there emerge four MPS strands from the center, which becomes six for a 3D cubic lattice. As long as there are no loops we can put the TTN in canonical form. The principles that underlie the isometry reduction generalize directly to three dimensions, and in Fig. 5 we show the 3D analog of Fig. 3.

### B. Evolving 3D isoTNS

We will now explain how the TTN is moved through the 3D isoTNS during its trotterized time evolution, specializing to nearest-neighbor interacting Hamiltonians. In order to take advantage of the isoTNS representation we always apply evolution gates when all sites occupy the TTN, with the orthogonality center at one of the sites. We will see that this gives rise to a threefold-nested TEBD, which we call TEBD$^3$ in analogy to TEBD$^2$ for 2D isoTNS.

The evolution takes place columnwise, starting with the middle slice in Fig. 5(a) that is shown separately in Fig. 6(a). To this end we first trotterize the evolution operator $\exp(-d\tau H)$, which is for an imaginary time step of size $d\tau$, in terms of columns $c_x$, $c_y$, and $c_z$. At first order in $d\tau$ we can trotterize as

$$e^{-d\tau H} \approx \prod_{c_x} e^{-d\tau H_{c_x}} \prod_{c_y} e^{-d\tau H_{c_y}} \prod_{c_z} e^{-d\tau H_{c_z}}, \qquad (4)$$

where the error is $\mathcal{O}(d\tau^2)$ and stems from the noncommutativity of columns that intersect. With only a bit more effort we can trotterize at second order:

$$e^{-d\tau H} \approx \prod_{c_x} e^{-\frac{d\tau}{2} H_{c_x}} \prod_{c_y} e^{-\frac{d\tau}{2} H_{c_y}} \prod_{c_z} e^{-d\tau H_{c_z}} \prod_{c_y} e^{-\frac{d\tau}{2} H_{c_y}} \prod_{c_x} e^{-\frac{d\tau}{2} H_{c_x}}, \qquad (5)$$

which has error $\mathcal{O}(d\tau^3)$. Here subsequent $c_x$ terms can be combined when performing multiple time steps. In

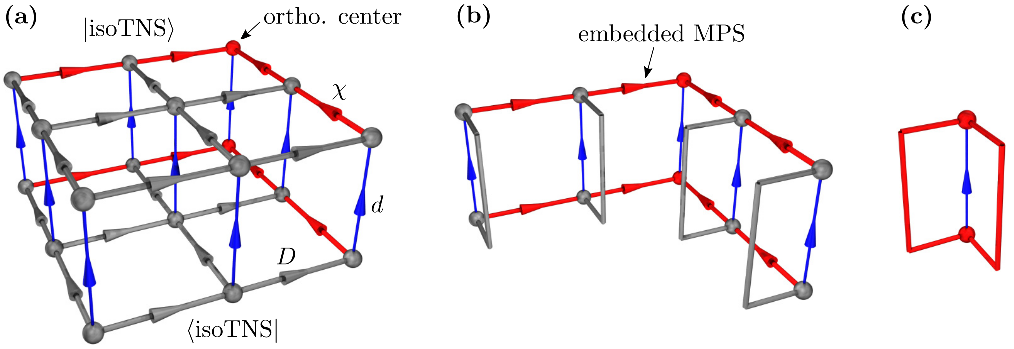


FIG. 3. Multiple stages of the reduction of ⟨isoTNS | isoTNS⟩ to a contraction of just two tensors, illustrated for a 2D isoTNS on a 3 × 3 square lattice. Panel (a) shows the amplitude before employing the isometry constraints, where the gray bonds have dimension $D$, the red bonds have dimension $\chi$, and the blue bonds have dimension $d$. Panel (b) shows the canonical MPS that is obtained after utilizing a large part of the isometry structure. Panel (c) shows the final pair of contracted tensors that remains after fully utilizing the isometry structure.

Appendix B we investigate the interplay between the trotterization error and the various truncation errors.

Since we will be dealing exclusively with nearest-neighbor interacting systems, each column is further trotterized in terms of two-body gates. At first order this gives us

$$e^{-d\tau H_{c_x}} = \prod_{b_i} e^{-d\tau h_{b_i}} + \mathcal{O}(d\tau^2), \tag{6}$$

where $b_i$ labels the bonds in the column and $h_{b_i} \in \mathbb{C}^{4\times 4}$ represents the Hamiltonian on $b_i$. At second order we get

$$e^{-d\tau H_{c_x}} = \prod_{i=1}^{L} e^{-\frac{d\tau}{2} h_{b_i}} \prod_{i=L}^{1} e^{-\frac{d\tau}{2} h_{b_i}} + \mathcal{O}(d\tau^3), \tag{7}$$

which combined with Eq. (5) yields an overall second-order trotterization.

### 1. *Evolving the columns of a slice*

To begin the evolution we apply a two-body gate at the initial orthogonality center, as illustrated in Fig. 6(b). For clarity we omit the transverse legs in most of the panels, only restoring them when crucial for interpretation. Before contracting the gate with the tensors in Fig. 7(a) we apply a QR decomposition into an orthogonal matrix Q and a triangular matrix R at each tensor to get a reduced bond [10], yielding the configuration in Fig. 7(b) where we now get to evolve a reduced space because the orange transient bonds are much smaller than the dark-red bonds (which consist of many virtual legs).

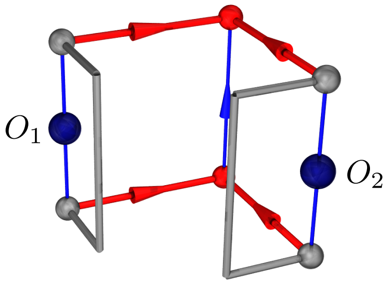


FIG. 4. The final product of the isometry reductions when calculating the local two-body correlator ⟨isoTNS $|O_1 O_2|$isoTNS⟩, leaving us with a MPS correlator that has to be fully computed.

After contracting the gate with the reduced tensors we perform a truncated singular value decomposition (SVD) $A \approx UsV^\dagger$, with $U$ an isometry and $sV^\dagger$ the new orthogonality center, to regain the reduced bond while shifting the orthogonality center, giving the configuration in Fig. 7(c). To ensure that the new bond is not larger than the bond in Fig. 7(a) we often need to truncate the singular values $s$, which can be done optimally since we are at the orthogonality center. After reabsorbing the reduced tensors we have now shifted the orthogonality center by one site while evolving the bond, giving us the configuration in Fig. 6(c).

We then evolve the next bond, so that we end up at the bottom as shown in Fig. 6(d). To evolve the next column we first need to transfer the TTN strand to this column, for which we use the column-splitting procedure introduced in [13] which is sequence of triangle splittings. This is illustrated in Figs. 6(e)–6(h).

A single triangle splitting is shown in Fig. 8. By performing two truncated SVDs on the orthogonality center in Fig. 8(a) we get the decomposition in Fig 8(b). To improve the quality of the column splitting, i.e., reduce the information loss in obtaining Fig. 8(g) from Fig. 8(d), we follow [13] and reduce the bipartite entanglement between the right and upper tensors of the triangle in Fig. 8(b). For this we insert a pair of unitary "disentanglers" $U^\dagger U = \mathbb{1}$ and optimize them such that the $\alpha$-Rényi entropy $S_\alpha$ between these tensors is minimized. If we put the orthogonality center $s$ on this bond we can write

$$S_\alpha = \frac{1}{1-\alpha} \ln \sum_i s_i^\alpha, \tag{8}$$

where $\alpha$ is the Rényi order. For $\alpha < 1$ this quantity is known to provide a bound on MPS precision [26]. After minimizing $S_\alpha$ we have a triangle with minimal entanglement across its red bond, and therefore we will end up with a split-off column that has minimal vertical entanglement. The disentangler can be easily optimized with gradient descent. In the case of $\alpha = 2$ there is also a cheaper optimization algorithm [27], but as in [13] we find that $\alpha < 1$ gives the best performance.

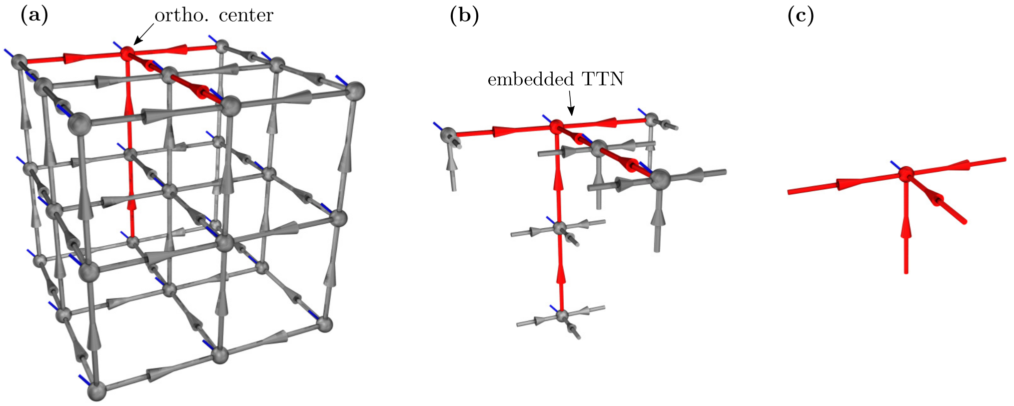


FIG. 5. The reduction of a particular 3D isoTNS configuration as it would occur upon contraction with its conjugate. In panel (a) we show the initial network. In panel (b) we show the canonical TTN that would remain after utilizing most of the isometry structure. In panel (c) we show how the canonical TTN further reduces to a single site, i.e., the orthogonality center.

After performing the triangle splitting on the orthogonality center in Fig. 6(d) we absorb the new center upwards as in Figs. 6(e) and 8(c). To continue the column splitting we then temporarily combine the legs as designated in Fig. 8(c) and repeat the triangle splitting, giving the configuration in Fig. 6(f), which becomes Fig. 6(g) after a truncated SVD. This completes the column splitting, and after absorbing the transient column into the next column we have successfully transferred the TTN strand, as shown in Fig. 6(h). This absorption increases the vertical bond dimension from $\chi$ to $D\chi$, which we truncate back to $\chi$ before proceeding with the evolution. After repeating the steps in Fig. 6 on the middle column in Fig. 6(h), and subsequently evolving the final column, we have moved the TTN strand from the first to the last column while evolving all columns. The final configuration is shown in Fig. 10(a).

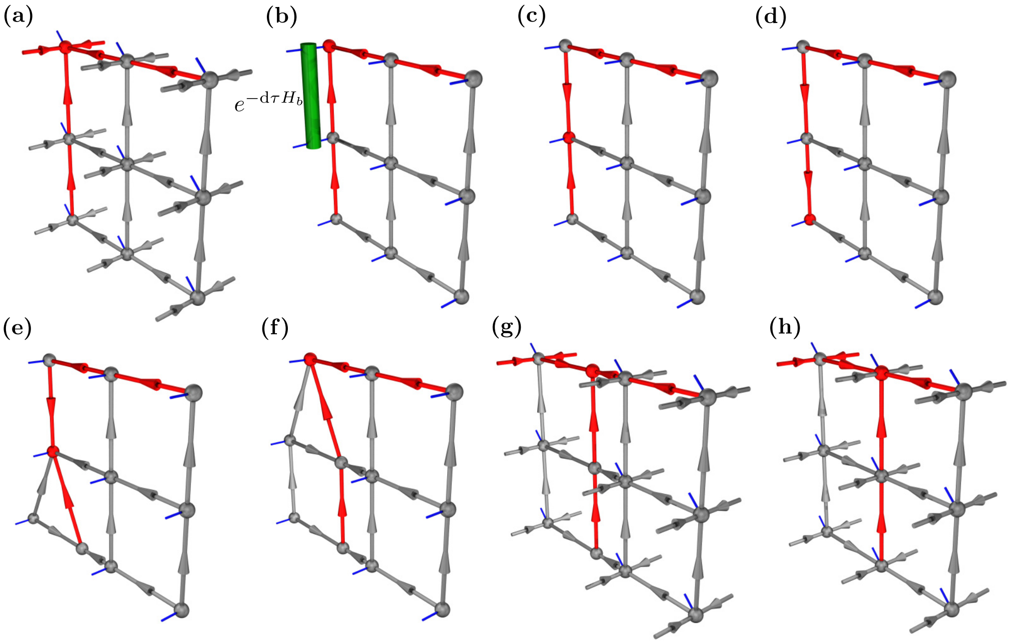


FIG. 6. The evolution of a slice in the 3D isoTNS, where for clarity we omit the transverse legs in most of the panels. In panel (a) we show the initial slice, which is the middle slice from Fig. 5. In panel (b) we apply a two-body gate at the orthogonality center, after which in panel (c) we shift the orthogonality center downwards using a SVD. In panel (d) we repeat this after which the orthogonality center is at the bottom of the first column. In panel (e) we perform the triangle splitting from Fig. 8 on the orthogonality center, which we repeat in panel (f). Now that we have reached the top of the column we perform a truncated SVD to finalize the column splitting, after which we have transferred the TTN strand to a temporary column that has only virtual legs, as shown in panel (g) where we restored the transverse legs to emphasize that the virtual column has no transverse legs. By absorbing the virtual column into the neighboring isometry column, yielding the configuration in panel (h), we have finally moved the TTN strand to the middle column. The increased vertical bond dimension $D\chi$ is truncated back to $\chi$ before evolving the new column. Repeating the column splitting and evolution once more we will have evolved all columns in the slice.

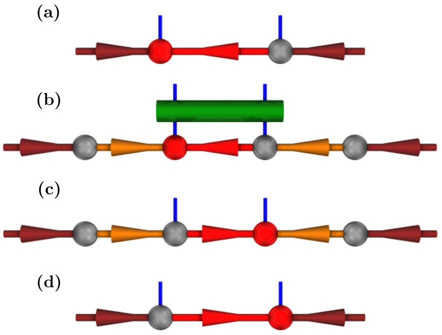


FIG. 7. The reduced time evolution used in TEBD$^3$. The bond depicted in panel (a) is to be evolved, to which end we first create a reduced bond by applying a QR decomposition at each tensor. This yields a pair of transient bonds with dimension $d\chi$ that are colored orange. Because these orange bonds are much smaller than the dark-red bonds, the gate application shown in panel (b) and the subsequent truncated SVD that yields panel (c) are much cheaper than when working in the gauge from panel (a). Having evolved the reduced bond we reabsorb the reduced tensors to get the original bond shown in panel (d), for which the orthogonality center has been shifted relative to panel (a).

### 2. *Splitting a slice*

Having evolved the columns in the first slice we now want to transfer the TTN strands to the next slice and repeat the evolution. To achieve this we use a tetrahedron splitting on the orthogonality center, which is illustrated in Fig. 9. By using a series of such splittings we can split off a transient slice containing the TTN strands, converting Fig. 10 into Fig. 10(e). Starting at the orthogonality center in Fig. 10(a), shown separately in Fig. 9(a), we perform a sequence of SVDs to get the tetrahedron in Fig. 9(b). The face constituted by the tensors $A$, $B$, and $C$ is part of the new slice.

To improve the fidelity of the slice splitting we now need to consider the tripartite entanglement between $A$, $B$, and $C$. Hence we insert a pair of tripartite disentanglers $U^\dagger U = \mathbb{1}$ that minimizes a tripartite extension of the $\alpha$-Rényi entropy $S_\alpha$ (similar to the tripartite information I3 [28]):

$$S3_\alpha(A|B|C) = S_\alpha(A|BC) + S_\alpha(B|AC) + S_\alpha(C|AB), \qquad (9)$$

where, e.g., $S_\alpha(A|BC)$ is the bipartite $\alpha$-Rényi entropy between tensor $A$ and the contraction of $B$ and $C$. We minimize $S3_\alpha(A|B|C)$ by iterating over its terms and performing a single step of bipartite disentangling each time. It should be noted that each bipartite disentangler is here a three-body operator. After minimizing $S3_\alpha(A|B|C)$ we get a triangle $ABC$ with minimal entanglement, and hence we will end up with a split-off slice that has minimal internal entanglement.

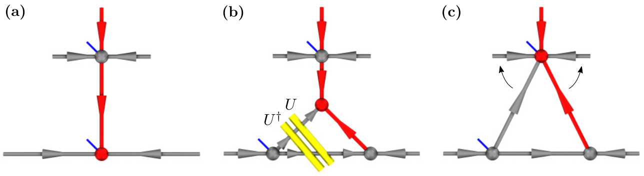


FIG. 8. The triangle splitting of the orthogonality center that is used multiple times in succession to shift a TTN strand to a neighboring column. By performing two truncated SVDs we split the orthogonality center shown in panel (a) into the triangle shown in panel (b), where the upper tensor is the new orthogonality center and where the two lower tensors are isometries. Note that the top and bottom-right tensors have only virtual legs. The amount of truncation during the SVDs is based on the usual color coding, and we have inserted a pair of disentanglers $U^\dagger U = \mathbb{1}$, depicted as yellow cylinders, in order to minimize the entanglement across the red bond. After the splitting we absorb the top tensor upwards, as shown in panel (c). To now perform the splitting from panel (b) on the new orthogonality center we first temporarily combine the tilted legs as indicated by the arrows, so that the orthogonality center again looks as in panel (a).

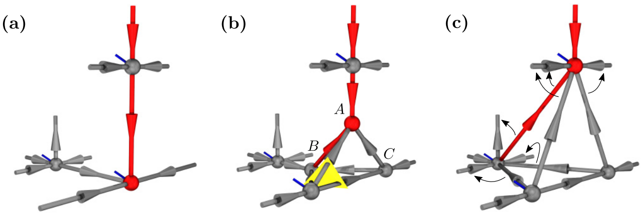


FIG. 9. The tetrahedron splitting used in the slice splitting that transfers the TTN strands to the next slice. In panel (a) we show the initial orthogonality center. In panel (b) we show the tetrahedron that results from three consecutive SVDs, where the pair of tripartite disentanglers is depicted as a yellow triangle, which reduces the tripartite entanglement $S3_\alpha(A|B|C)$. Note that only one of the four produced tensors has a physical leg. In panel (c) we have absorbed $A$ upwards and $B$ backwards. To repeat the tetrahedron splitting on the new orthogonality center we temporarily combine the tilted legs as indicated by the arrows.

In Appendix A we compare this tripartite disentangler with a direct 3D extension of the bipartite disentangler from Fig. 8, where we replace each side of the yellow triangle in Fig. 9 by a pair of bipartite disentanglers.

To continue the slice splitting we absorb $A$ upwards and $B$ backwards, which is illustrated in Figs. 9(c) and 10(b). We then temporarily combine the legs as indicated in Fig. 9(c) and repeat the tetrahedron splitting, yielding Fig. 10(c). Note that the new vertical red bond is enlarged to $D\chi$. With the orthogonality center at the top we perform a triangle splitting to get Fig. 10(d), where we also truncated the enlarged vertical bonds back to $\chi$. This produces the first column of the new slice. By combining the tilted legs as in Fig. 9 we can repeat the splitting to move the TTN strand to the final column. We complete the slice splitting by performing the column splitting from Figs. 6(e)–6(h), resulting in Fig. 10(e). The transient slice is then absorbed into the next slice, after which the TTN strands have been successfully transferred, as illustrated in Fig. 10(f). Before continuing with the evolution we truncate the enlarged bonds.

### 3. *Evolving the full network*

With the machinery developed in the previous sections we can now perform a trotterized time evolution where all truncation occurs at the orthogonality center, so that the local evolution is globally optimal. To illustrate this we consider the details of a single TEBD$^3$ iteration, starting from the network in Fig. 11(a).

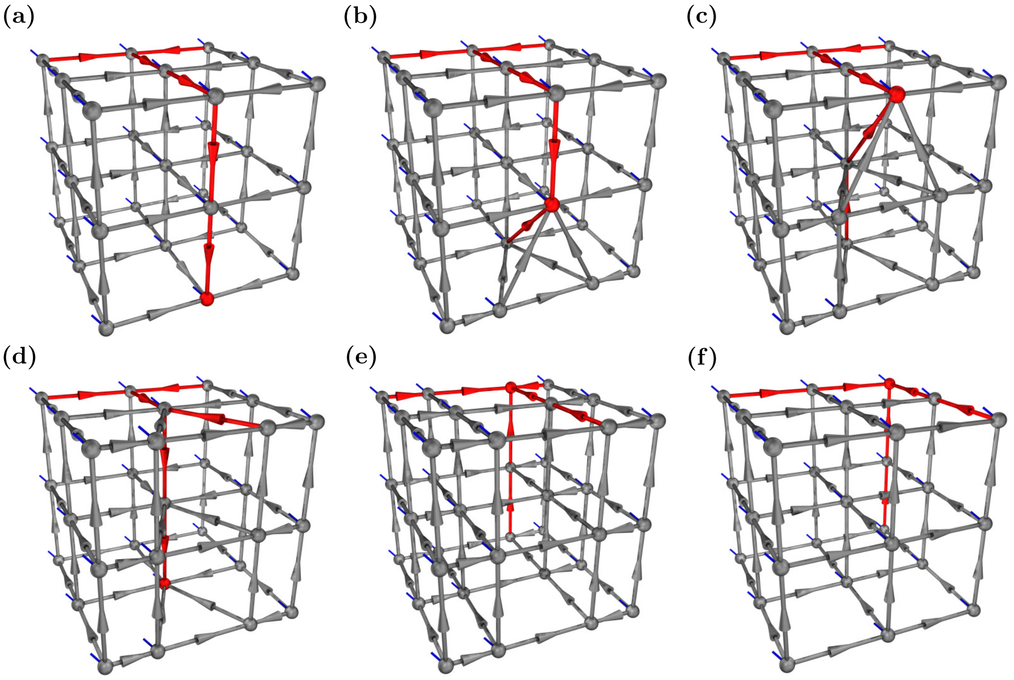


FIG. 10. The slice splitting by which we transfer the two TTN strands to the next slice, after evolving the vertical bonds of a slice. In panel (a) we show the initial isoTNS, which is obtained from Fig. 5 by evolving the middle slice as in Fig. 6. In panel (b) we show the first step of the slice splitting, where we employ the tetrahedron splitting from Fig. 9 on the orthogonality center, thereby moving it upwards by one site. After temporarily combining the tilted legs with the horizontal legs as indicated by the arrows in Fig. 9(c) we repeat the tetrahedron splitting to get panel (c). Finally we apply the triangle splitting from Fig. 8 to the uppermost tensor in panel (d). We then move the orthogonality center to the bottom of the column, resulting in panel (d). We repeat this combination of evolving and column splitting until we reach the final column, which is instead split via the procedure described in Figs. 6(e)–6(h), yielding the split-off slice in panel (e). To finalize the transfer we absorb the transient slice into the neighboring slice, yielding panel (f). The increased bond dimensions are truncated before continuing the evolution.

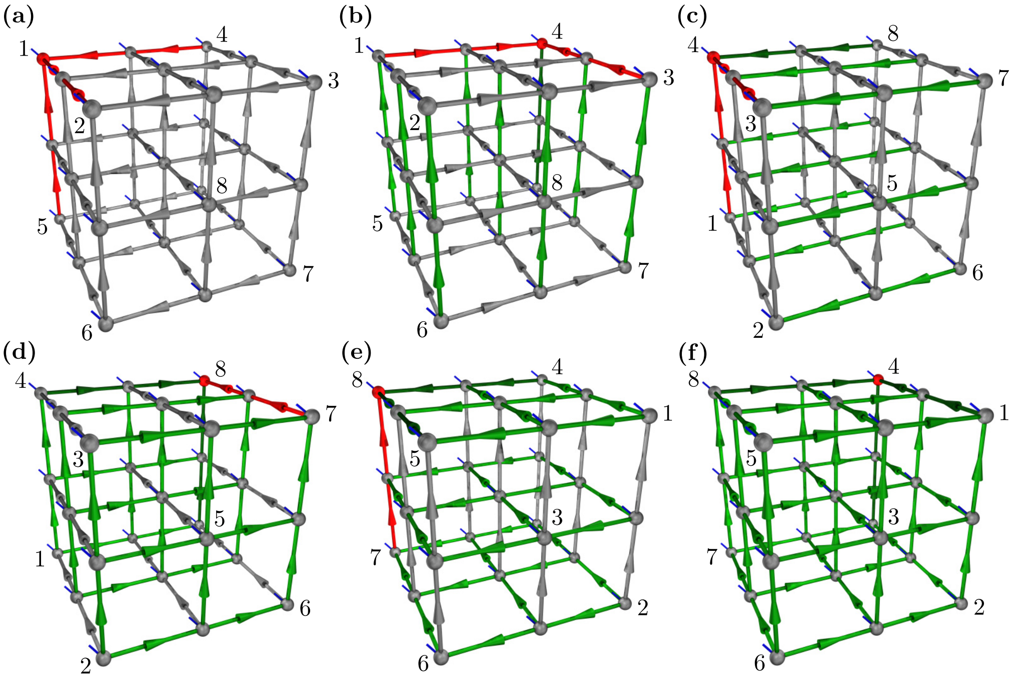


FIG. 11. A single iteration of TEBD$^3$. Starting from the initial configuration in panel (a) we evolve all columns to get panel (b). Here the evolved gray bonds are colored green and the evolved red bonds are colored dark green. To evolve the next set of columns we rotate the network as indicated by the numbering of the corners, yielding panel (c), which becomes panel (d) after evolving the new columns. To evolve the final set of columns we rotate the network to panel (e), which turns into panel (f) upon evolving the columns. With all bonds evolved this concludes a single iteration of TEBD$^3$.

We evolve the columns of the first slice, following the procedure in Fig. 6, and subsequently move the TTN strands to the neighboring slice using the slice splitting of Fig. 10. Repeating these operations on the middle slice and afterwards evolving the final slice we end up with Fig. 11(b), where all columns have now been evolved as indicated by the green bonds.

To evolve the rest of the bonds we rotate the network in Fig. 11(b) such that the TTN again has the position from Fig. 11(a) while the columns consist of nonevolved bonds. We have numbered the corners in Fig. 11 to show a possible way of doing this. After evolving the columns, yielding Fig. 11(d), we rotate the network to get Fig. 11(e), which becomes Fig. 11(f) upon evolving the final columns. Since all bonds have now been evolved we have completed the iteration of TEBD$^3$.

The TEBD$^3$ algorithm has two main sources of error: the error due to the multitude of truncated SVDs and the error due to the trotterization. In [13] it was found that for TEBD$^2$ the column-splitting truncation error and trotterization error conspire to yield an energy minimum in $d\tau$ space. In Appendix B we show that the same occurs for TEBD$^3$, now resulting from the interplay between the column- and slice-splitting truncation errors and the trotterization error.

#### *4. Computational complexity*

The computational complexity of the TEBD$^3$ algorithm can be attributed to various operations, in particular to the SVDs that occur during these operations. Here we include only terms that are potentially leading in either $d$, $D$, or $\chi$. We moreover assumed that we perform full SVDs instead of partial SVDs in obtaining all estimates.

Starting with the slice splitting and subsequent absorption, we find that the tetrahedron splitting has cost $\mathcal{O}[dD^{10}\chi^2 \min(d, D\chi^2)]$ when it is performed on a tensor with the maximum amount of combined legs (see Fig. 9). The truncation of the enlarged bonds after absorbing the split-off slice has cost $\mathcal{O}(dD\chi^6)$.

Other contributions arise from the evolution of the bonds in a column, where the pairs of QR decompositions that precede gate application have cost $\mathcal{O}[d\chi^5 \min(d, \chi^2)]$. These decompositions yield pairs of intermediate bonds with sizes $\eta$ and $\zeta$, which depend on the external bonds involved in the decompositions, so that the subsequent SVD has cost $\mathcal{O}[d^3 \min(\eta\zeta^2, \eta^2\zeta)]$. The pair that potentially yields the leading order in $D$ occurs in the bulk and has $\eta = \zeta = \chi \min(d, D^4)$.

When simultaneously $d \leqslant \chi^2$ and $D^4$ the total cost reduces to $\mathcal{O}(d^2D^{10}\chi^2) + \mathcal{O}(dD\chi^6) + \mathcal{O}(d^6\chi^3)$. Because the costliest operations are performed roughly $N$ times we have the linear scaling $\mathcal{O}(N)$ in system size.

## III. BENCHMARKING

### A. The benchmarking system

As a proof of principle we probe the accuracy of TEBD$^3$ for imaginary time evolution to find an isoTNS approximation for the many-body ground state of a simple 3D quantum many-body system: the ferromagnetic transverse field Ising model (TFIM) on a cubic lattice with OBC:

$$H = -\sum_{\langle ij\rangle} \sigma_i^z \sigma_j^z - h \sum_i \sigma_i^x. \tag{10}$$

Here $\vec{\sigma}_i \equiv (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ corresponds to a spin 1/2 on site $i$, with $\sigma^{x,y,z}$ the usual Pauli matrices, and the two-body sum runs over nearest-neighbor pairs $\langle i, j\rangle$.

Working in the $\sigma^z$ basis, we see from Eq. (10) that the TFIM Hamiltonian becomes classical for $h \to 0$ where it has a twofold degenerate ferromagnetic ground state $|\uparrow\uparrow \dots \uparrow\rangle$ and $|\downarrow\downarrow \dots \downarrow\rangle$. Quantum fluctuations are generated by the uniform magnetic transverse (i.e., along the $x$ direction) field of strength $h$. In the limit of strong fields $h \to \infty$, the ground state is unique and aligned with the field and a simple product state $|\to\to \cdots \to\rangle$ in the $\sigma_x$ basis. Between these two limits, the competition between the $x$ and $z$ basis leads to a much more complex ground state [29] and to a quantum phase transition as a function of $h$ between the ferromagnetic phase at $h < h_c$ and the polarized phase at $h > h_c$. The value of the critical field on the 3D simple cubic lattice was numerically estimated to be $h_c \approx 5.158\,13(6)$ [30].

Our main reason for choosing this model as a benchmark is its simplicity and the fact that it can be solved exactly using quantum Monte Carlo using the stochastic series expansion [31,32]. Our two benchmark observables are the energy density $E/N = \langle H\rangle/N$ and the $x$ magnetization $m_x = \sum_i \langle\sigma_i^x\rangle/N$, which we compare to exact values in order to assess the accuracy of TEBD$^3$.

Quantum phase transitions are associated with a divergent correlation length and are therefore extremely challenging to study using tensor network methods. While a phase transition occurs at a singular point in the thermodynamic limit, a critical region is expected for finite systems. In the case of OBC the region gets shifted towards $h < h_c$, which is especially pronounced for the smaller $L$, where the ratio of one- to two-body couplings is significantly larger. We expect the 3D isoTNS ansatz to be particularly challenged in this critical region, whereas its performance is likely to improve when progressing deeper into both phases, since the ground state in both phases is a dressed product state.

### B. The benchmarking procedure

We perform imaginary time TEBD$^3$ propagation of the wave function represented by an isoTNS with bond dimensions $(D, \chi)$ for the TFIM (10) with $L = 3, 4, 10$ at various fields $h = 0.5, 1.0, \dots, 8$ (across the critical point). For $L = 3$ we use bond dimensions $D = 2, 4, 6$ with $\chi = 2D, 4D, 6D$, which allows us to observe the convergence of both $E/N$ and $m_x$ towards the exact values. For $L > 3$ we are more constrained in our choice of $(D, \chi)$ due to the large cost as derived in Sec. II B 4. Hence for $L = 4$ we use $D = 2, 3$ with $\chi = 2D, 4D, 6D$, and to illustrate the capacity of TEBD$^3$ to reach large system sizes we consider a (2,4) isoTNS for $L = 10$.

For the initial state we take a $\sigma_z$ product state $|\psi_{\text{ini}}\rangle = |\uparrow\uparrow \dots \uparrow\rangle$, which we evolve in imaginary time until the energy density is converged. This means that we perform the operation

$$|\psi_0\rangle \approx e^{-\beta \hat{H}} |\uparrow\uparrow \dots \uparrow\rangle, \tag{11}$$

which for TEBD algorithms is done by trotterizing it into small time steps $d\tau$ as explained in Sec. II B. We typically evolve to $\beta \gtrsim 40$.

As mentioned Sec. II B 3, TEBD$^3$ has an energy minimum in $d\tau$ space, so for each simulation we choose $d\tau$ such that it coincides with the minimum. In Fig. 17 of Appendix B we show the $d\tau$ minima for various $(D, \chi)$ on a $L = 3$ lattice in the critical region. Similarly, we consider various Rényi entropy orders $\alpha$ for the disentangling and pick the one with lowest energy. It is usually between $\alpha = 1/2$ and 1, although it varies across the phase diagram.

The exact values for the energy density were obtained with the stochastic series expansion, for which we used the ALPS library [33,34]. Here convergence to the ground state was checked by a $\beta$-doubling scheme. For $L = 3$ we also performed Lanczos exact diagonalization, providing exact energy density and additionally the $x$ magnetization of the ground state at all $h$.

In order to provide a reference for the quality of our TEBD$^3$ results, we furthermore performed TEBD$^2$ calculations for the 2D TFIM using the algorithm detailed in [13]. Here we chose a comparable system size of $5 \times 5$, for which exact Lanczos values are easily obtained.

### C. Results

In this section we present the results of our TEBD$^3$ benchmarks for the 3D TFIM ground state.

We start off with $L = 3$, for which the performance can be comprehensively probed. In Fig. 12 we have plotted $E/N$ and $m_x$ across a range of $h$. Here the thermodynamic critical point is denoted by a gray dotted line, but as mentioned in Sec. III A this point is spread into an extended region for finite systems, and furthermore shifted to smaller $h < h_c$ due to the use of OBC.

The top panel of Fig. 12 shows $E/N$ across a range of $h$ for various $(D, \chi)$, along with the exact Lanczos values as a red dotted line. The inset shows the relative error in $E/N$. Here we see that deep in the polarized and ferromagnetic phases the accuracy is high, whereas in the critical region around $h \approx 3$ the performance is clearly worse.

For $D = 2$ and 4 we see that the accuracy is converged in $\chi$, meaning that $D$ must be increased (i.e., the column and slice splitting need higher fidelity) to further improve the performance. For $D = 6$ it is not clear whether convergence is reached, meaning that convergence becomes slower as $D$ grows. On a related note, it can be seen that (4,16) performs better than (6,12), illustrating why large $D$ is only useful when combined with a few multiples larger $\chi$.

The bottom panel of Fig. 12 shows $m_x$ for various $h$, again together with the exact Lanczos values as a red dashed line. The inset contains the absolute error $\delta = m_x^e - m_x$ relative to the exact value $m_x^e$, which is chosen over the relative error since $m_x^e$ vanishes as $h \to 0$. Deep in the polarized and ferromagnetic phases we see excellent agreement already for small bond dimensions, with the critical region clearly requiring larger tensors to get near the exact line.

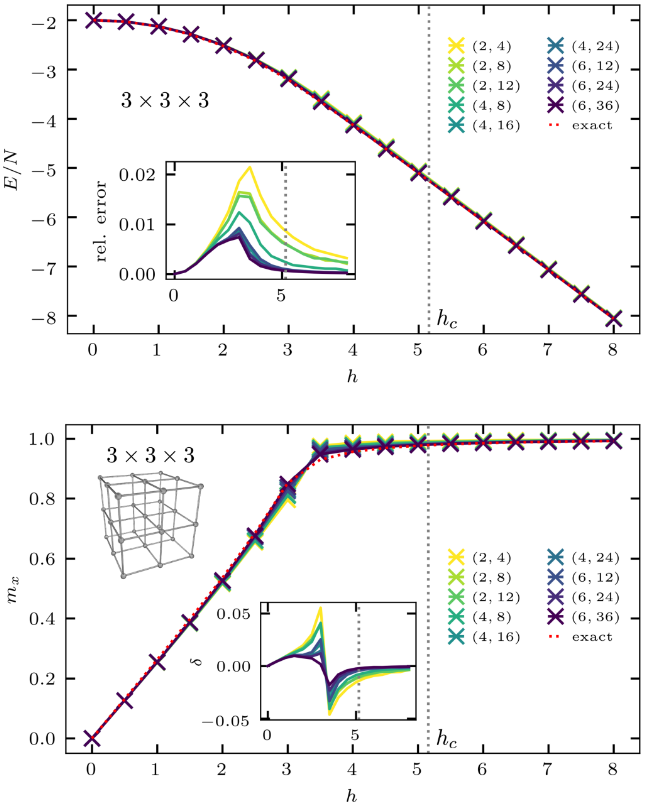


FIG. 12. Top: The energy density $E/N$ and its accuracy relative to the exact Lanczos value for the $3 \times 3 \times 3$ cubic TFIM with OBC at various field strengths $h$ across the critical point. Each curve belongs to a different set of bond dimensions $(D, \chi)$, and in the main plot we show the exact values with a red dotted line. Bottom: The $x$ magnetization $m_x$ and its error $\delta = m_x^e - m_x$ relative to the exact Lanczos value $m_x^e$ for various field strengths around the critical point, for the same TFIM system as in the top panel.

Next we consider $L = 4$, for which we show the results in Fig. 13. The curves display similar behavior as for $L = 3$, but due to the high computational cost we did not probe its performance beyond $D = 3$. A noticeable difference is the sharper peak in energy accuracy, which is likely due to the shrinking critical region (that has moreover shifted to larger $h$). For $L = 4$ the exact values were obtained via Monte Carlo, with the accompanying errors falling inside the linewidth.

Now we consider $L = 10$ (i.e., $N = 10^3$) in Fig. 14. Here we were not able to go beyond $D = 2$, which embodies the large difference in computational complexity between the bond dimensions and system size, as found in Sec. II B 4. We again observe a peak in relative error of $E/N$ at the critical region, which is now centered on the thermodynamic $h_c$. The $x$ magnetization is now also seen to saturate around $h_c$.

Overall, the comparison of our results across different system sizes reveals an excellent representation of the ground-state wave function deep in the ferromagnetic phase, even for the relatively small $(D, \chi)$ which are reachable at large system sizes on current computers. The accuracy is slightly worse in the polarized phase, especially for the smaller $(D, \chi)$, but nonetheless high accuracies can be reached on relatively

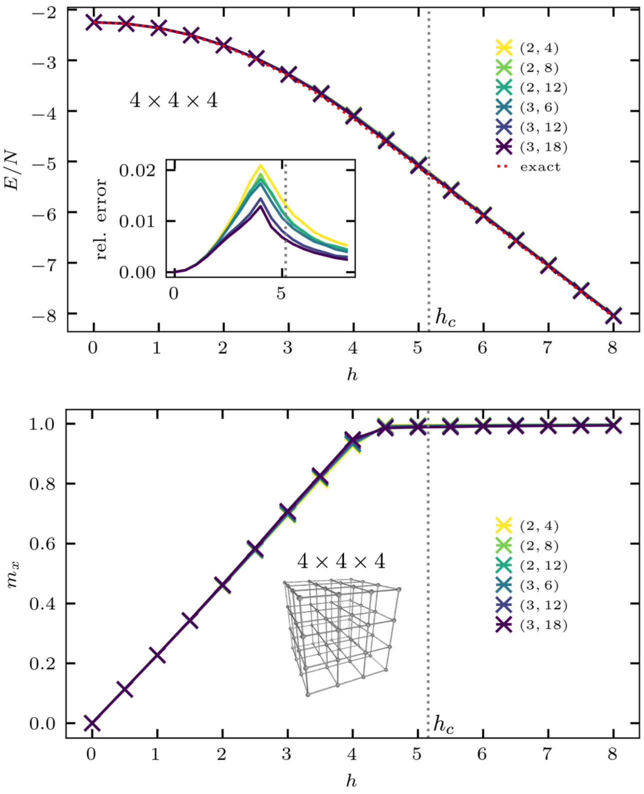


FIG. 13. Top: The energy density $E/N$ and its accuracy relative to the exact QMC value for the $4 \times 4 \times 4$ cubic TFIM with OBC at various field strengths $h$ across the critical point. Each curve belongs to a different set of bond dimensions $(D, \chi)$, and in the main plot we show the exact values with a red dotted line. Bottom: The $x$ magnetization $m_x$ for various field strengths around the critical point, for the same TFIM system as in the top panel.

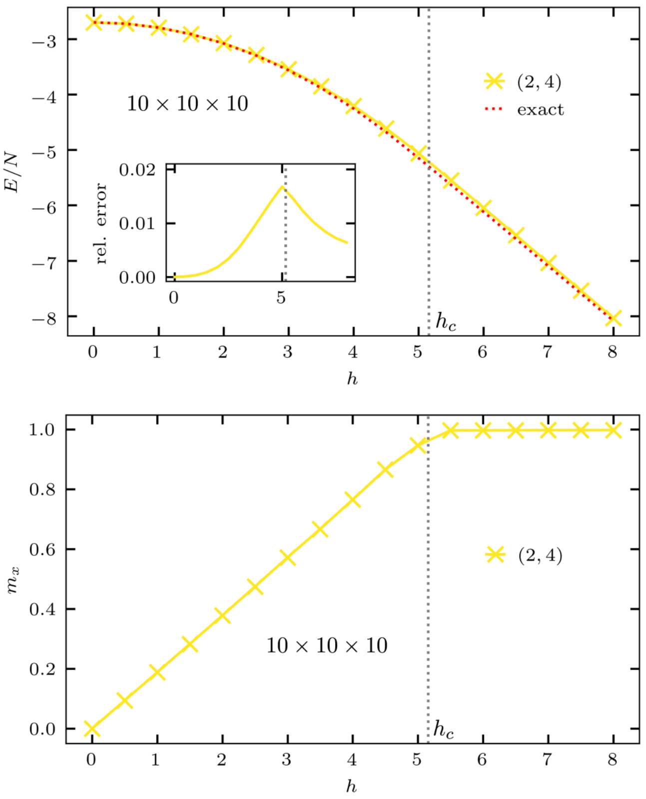


FIG. 14. Top: The energy density $E/N$ of the (2,4) isoTNS and its accuracy relative to the exact QMC value for the $10 \times 10 \times 10$ cubic TFIM with OBC at various field strengths $h$ across the critical point. The exact values are shown as a red dotted line. Bottom: The $x$ magnetization $m_x$ for various field strengths around the critical point, for the same TFIM system as in the top panel.

small lattices. As expected, the accuracy is worst in the critical region, but also here the performance is significantly improved upon increasing the bond dimensions, showing a clear trend toward the exact values both for the energy density and $x$ magnetization.

In order to put the TEBD$^3$ benchmarks into perspective we have also performed a TEBD$^2$ benchmark for the 2D TFIM on a $5 \times 5$ square lattice. It is clear that the simulation of a 2D system with TEBD$^2$ is easier than that of a 3D system with TEBD$^3$, since the 3D version involves more truncation. We therefore expect TEBD$^2$ to have higher accuracy for similar bond dimensions.

In Fig. 15 we show $E/N$ and $m_x$ for multiple $(D, \chi)$ at various $h$ for the $5 \times 5$ TFIM. In the top panel we see that the relative error of $E/N$ is again smallest when deep in the ferromagnetic and polarized phases, with the critical region around $h_c \approx 3$ again posing the biggest challenge. As expected, we see that (4,16) performs better in two dimensions than in three dimensions, reaching just below 0.5% compared to just below 1% in Fig. 12. The same is apparent from the bottom panel, where we see that (4,16) already closely matches the exact $m_x$ in the critical region, which is reached only with (6,36) in Fig. 12.

## IV. CONCLUSION

We have introduced a method for the simulation of 3D quantum lattice models using a representation of the wave function as a 3D isometric tensor network state (isoTNS). Generalizing the method for two dimensions presented in [13] we introduced a tetrahedral splitting and accompanying tripartite disentangling, such that optimal time-evolving block decimation can be carried out in cubic 3D networks. We call the resulting evolution algorithm TEBD$^3$.

Our systematic benchmark for the 3D transverse field Ising model in the full range of transverse fields across the critical point reveals that our method yields accurate results, and that the systematic error incurring from finite bond dimensions can be controlled systematically by increasing $(D, \chi)$. This behavior is identical to what is observed in TEBD$^2$. The regime close to the critical point is particularly challenging and requires larger bond dimensions, beyond the capacity of our computers for large systems.

While imaginary time evolution using TEBD$^3$ is arguably the simplest method for finding the ground state of a quantum many-body system, it is known even in one dimension that it is not optimal and that local variational energy minimization (e.g., DMRG in one dimension) is far more efficient. We expect a similar behavior for isoTNS in higher dimensions and it is possible that our results can be further improved by

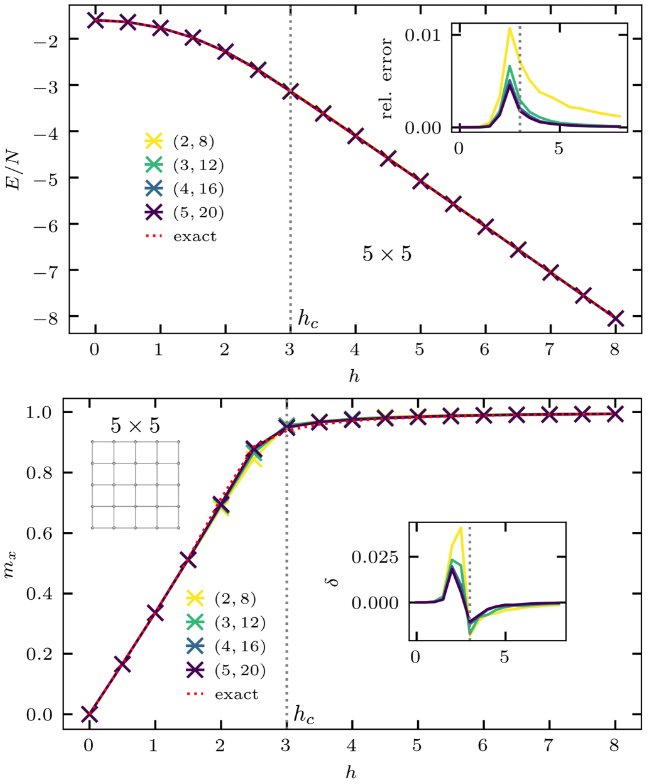


FIG. 15. Top: The energy density $E/N$ and its accuracy relative to the exact Lanczos value for the $5 \times 5$ square TFIM with OBC at various field strengths $h$ across the critical point. Each curve belongs to a different set of bond dimensions $(D, \chi)$, and in the main plot we show the exact values with a red dotted line. Bottom: The $x$ magnetization $m_x$ and its error $\delta = m_x^e - m_x$ relative to the exact Lanczos value $m_x^e$ for various field strengths around the critical point, for the same TFIM system as in the top panel.

the formulation of a DMRG analog for 3D isoTNS, a direction which we leave for future study.

## ACKNOWLEDGMENTS

We are grateful to M. Zaletel and F. Pollmann for useful suggestions to further improve isoTNS. We acknowledge financial support from the Deutsche Forschungsgemeinschaft through SFB1143 (Project number 247310070).

## APPENDIX A: BIPARTITE VERSUS TRIPARTITE DISENTANGLING

To quantify the difference in performance between the bipartite and tripartite 3D disentanglers from Sec. II B 2 we use each to calculate the ground-state energy density of the $4 \times 4 \times 4$ TFIM at various $h$. In Fig. 16 we show the performance of the isoTNS configurations (2,12) and (3,15). Here it is clear that the improvement in accuracy when using tripartite over bipartite disentanglers is largest in the critical region $h \in [3, 5]$, whereas the improvement becomes progressively smaller outside of this region.

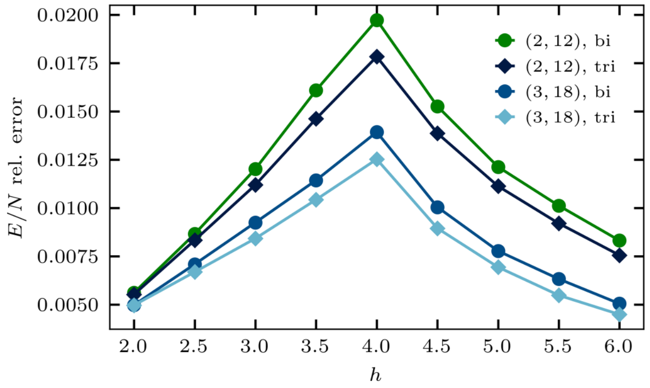


FIG. 16. A comparison of the bipartite and tripartite disentanglers, showing the energy accuracy for the $4 \times 4 \times 4$ TFIM ground state at various $h$ and two $(D, \chi)$. We see that the improvement in using a tripartite over a bipartite disentangler is most significant in the critical region.

Because the chosen $(D, \chi)$ are converged in $\chi$ we cannot improve the bipartite curves by further increasing the TTN's bond dimension. This illustrates that the quality of the tetrahedron splitting, and hence of the slice splitting, cannot be compensated by only improving the TTN. This is easily understood if we recognize that the slice splitting is a major component of TEBD$^3$, since it serves to transfer the TTN to the next slice with minimal information loss, and that this part of the algorithm is mainly controlled by $D$ and not $\chi$ (see Fig. 9). In particular, even though a larger $\chi$ might improve the time evolution on a fixed TTN, this gain is lost when transferring the TTN.

## APPENDIX B: MINIMA IN $d\tau$ SPACE

As noted in [13], the error due to the multitude of truncated SVDs and the error due to the trotterization combine to yield an energy minimum in $d\tau$ space for TEBD$^2$. Here we show that this also occurs for TEBD$^3$ and we will furthermore illustrate the difference between first- and second-order trotterization.

In Fig. 17 we show the relative error in energy density for the ground state of the $3 \times 3 \times 3$ TFIM at $h = 3.5$, at multiple points in $d\tau$ space. For each $(D, \chi)$ we plot both

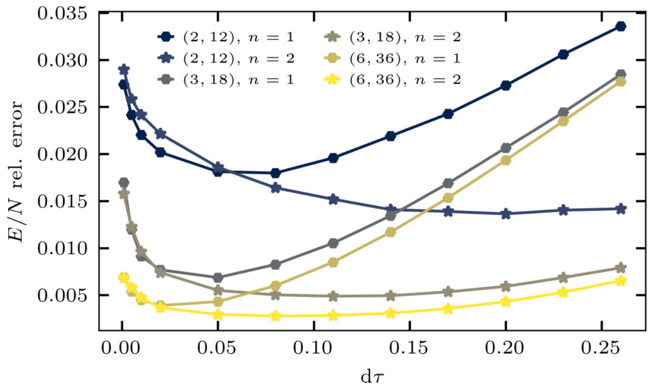


FIG. 17. The $d\tau$ dependency of the relative error in energy density of the $3 \times 3 \times 3$ TFIM at $h = 3.5$, for various isoTNS configurations $(D, \chi)$.

the first-order ($n = 1$) and second-order ($n = 2$) trotterization results, from which we see that for all considered cases the $n = 2$ minimum lies below the $n = 1$ minimum. We can also see that the minima shift to lower $d\tau$ as we increase $D$, for both $n = 1$ and 2, which is in accordance with the findings in [13] for TEBD$^2$.

---

[1] M. B. Hastings, An area law for one-dimensional quantum systems, J. Stat. Mech.: Theory Exp. (2007) P08024.

[2] M. B. Hastings, Locality in Quantum and Markov Dynamics on Lattices and Networks, Phys. Rev. Lett. **93**, 140402 (2004).

[3] J. Eisert, M. Cramer, and M. B. Plenio, Colloquium: Area laws for the entanglement entropy, Rev. Mod. Phys. **82**, 277 (2010).

[4] U. Schollwoeck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. (NY) **326**, 96 (2011).

[5] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, Phys. Rev. Lett. **69**, 2863 (1992).

[6] F. Verstraete and J. I. Cirac, Renormalization algorithms for quantum-many body systems in two and higher dimensions, arXiv:cond-mat/0407066.

[7] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Computational Complexity of Projected Entangled Pair States, Phys. Rev. Lett. **98**, 140506 (2007).

[8] M. Lubasch, J. I. Cirac, and M.-C. Bañuls, Unifying projected entangled pair state contractions, New J. Phys. **16**, 033014 (2014).

[9] F. Verstraete, V. Murg, and J. I. Cirac, Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems, Adv. Phys. **57**, 143 (2008).

[10] M. Lubasch, J. I. Cirac, and M.-C. Banuls, Algorithms for finite projected entangled pair states, Phys. Rev. B **90**, 064425 (2014).

[11] G. Vidal, Efficient Classical Simulation of Slightly Entangled Quantum Computations, Phys. Rev. Lett. **91**, 147902 (2003).

[12] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwoeck, and C. Hubig, Time-evolution methods for matrix-product states, Ann. Phys. (NY) **411**, 167998 (2019).

[13] M. P. Zaletel and F. Pollmann, Isometric Tensor Network States in Two Dimensions, Phys. Rev. Lett. **124**, 037201 (2020).

[14] M. Zaletel, Isometric Tensor Network States in Two Dimensions, lecture delivered at Perimeter Institute, Waterloo, ON, 2019, https://www2.perimeterinstitute.ca/videos/isometric-tensor-network-states-two-dimensions.

[15] R. Haghshenas, M. J. O'Rourke, and G. K.-L. Chan, Conversion of projected entangled pair states into a canonical form, Phys. Rev. B **100**, 054404 (2019).

[16] K. Hyatt and E. M. Stoudenmire, DMRG Approach to Optimizing Two-Dimensional Tensor Networks, arXiv:1908.08833.

[17] T. Soejima, K. Siva, N. Bultinck, S. Chatterjee, F. Pollmann, and M. P. Zaletel, Isometric tensor network representation of string-net liquids, Phys. Rev. B **101**, 085117 (2020).

[18] P. Teng, Generalization of the tensor renormalization group approach to 3-d or higher dimensions, Physica A **472**, 117 (2017).

[19] A. García-Sáez and J. I. Latorre, Renormalization group contraction of tensor networks in three dimensions, Phys. Rev. B **87**, 085130 (2013).

[20] S. Anders, M. B. Plenio, W. Dür, F. Verstraete, and H.-J. Briegel, Ground-State Approximation for Strongly Interacting Spin Systems in Arbitrary Spatial Dimension, Phys. Rev. Lett. **97**, 107206 (2006).

[21] A. Sfondrini, J. Cerrillo, N. Schuch, and J. I. Cirac, Simulating two- and three-dimensional frustrated quantum systems with string-bond states, Phys. Rev. B **81**, 214426 (2010).

[22] M. Schmitt and M. Heyl, Quantum dynamics in transverse-field Ising models from classical networks, SciPost Phys. **4**, 013 (2018).

[23] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Criticality, the Area Law, and the Computational Power of Projected Entangled Pair States, Phys. Rev. Lett. **96**, 220601 (2006).

[24] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Classical simulation of quantum many-body systems with a tree tensor network, Phys. Rev. A **74**, 022320 (2006).

[25] F. A. Y. N. Schroeder, D. H. P. Turban, A. J. Musser, N. D. M. Hine, and A. W. Chin, Tensor network simulation of multi-environmental open quantum dynamics via machine learning and entanglement renormalisation, Nat. Commun. **10**, 1062 (2019).

[26] F. Verstraete and J. I. Cirac, Matrix product states represent ground states faithfully, Phys. Rev. B **73**, 094423 (2006).

[27] J. Hauschild, E. Leviatan, J. H. Bardarson, E. Altman, M. P. Zaletel, and F. Pollmann, Finding purifications with minimal entanglement, Phys. Rev. B **98**, 235163 (2018).

[28] M. Rangamani and M. Rota, Entanglement structures in qubit systems, J. Phys. A: Math. Theor. **48**, 385301 (2015).

[29] D. J. Luitz, F. Alet, and N. Laflorencie, Universal Behavior Beyond Multifractality in Quantum Many-Body Systems, Phys. Rev. Lett. **112**, 057203 (2014).

[30] H. W. J. Blöte and Y. Deng, Cluster Monte Carlo simulation of the transverse Ising model, Phys. Rev. E **66**, 066110 (2002).

[31] A. W. Sandvik and J. Kurkijärvi, Quantum Monte Carlo simulation method for spin systems, Phys. Rev. B **43**, 5950 (1991).

[32] A. W. Sandvik, A generalization of Handscomb's quantum monte carlo scheme-application to the 1D Hubbard model, J. Phys. A: Math. Gen. **25**, 3667 (1992).

[33] B. Bauer, L. D. Carr, H. G. Evertz, A. Feiguin, J. Freire, S. Fuchs, L. Gamper, J. Gukelberger, E. Gull, S. Guertler, A. Hehn, R. Igarashi, S. V. Isakov, D. Koop, P. N. Ma, P. Mates, H. Matsuo, O. Parcollet, G. Pawłowski, J. D. Picon *et al.*, The ALPS project release 2.0: Open source software for strongly correlated systems, J. Stat. Mech.: Theory Exp. (2011) P05001.

[34] A. F. Albuquerque, F. Alet, P. Corboz, P. Dayal, A. Feiguin, S. Fuchs, L. Gamper, E. Gull, S. Gürtler, A. Honecker, R. Igarashi, M. Körner, A. Kozhevnikov, A. Läuchli, S. R. Manmana, M. Matsumoto, I. P. McCulloch, F. Michel, R. M. Noack, G. Pawłowski *et al.*, The ALPS project release 1.3: Open-source software for strongly correlated systems, J. Magn. Magn. Mater. **310**, 1187 (2007).

# APPENDIX E

# Article 5

PHYSICAL REVIEW B **107**, L201114 (2023)

**Letter**

# Compressed quantum error mitigation

Maurits S. J. Tepaske and David J. Luitz [*]
*Physikalisches Institut, Universität Bonn, Nußallee 12, 53115 Bonn, Germany*

(Received 21 February 2023; accepted 12 May 2023; published 22 May 2023)

We introduce a quantum error mitigation technique based on probabilistic error cancellation to eliminate errors which have accumulated during the application of a quantum circuit. Our approach is based on applying an optimal "denoiser" after the action of a noisy circuit and can be performed with an arbitrary number of extra gates. The denoiser is given by an ensemble of circuits distributed with a quasiprobability distribution. For a simple noise model, we show that efficient, local denoisers can be found and we demonstrate their effectiveness for the digital quantum simulation of the time evolution of simple spin chains.

DOI: 10.1103/PhysRevB.107.L201114

*Introduction.* Quantum information processing has been theoretically shown to hold great promise and quantum algorithms were developed which can in principle achieve an exponential speedup over their classical counterparts, both for general purpose computing [1–4] and quantum simulation [5–9]. However, present day quantum computing prototypes still suffer from significant noise processes which hinder the execution of many potentially groundbreaking quantum algorithms [10]. Nontrivial quantum algorithms typically require large sequences of quantum gates, each of which introduces dissipation and hence an overall loss of coherence, eventually rendering the results useless.

Until quantum error correction [11,12] becomes practical, *quantum error mitigation* seems to be more feasible to increase the accuracy of expectation values. Here the goal is to induce the (partial) cancellation of errors that stem from noisy quantum gates by extending the circuit corresponding to the desired algorithm with an ensemble of gates [13,14], sampled from a quasiprobability distribution.

The traditional way to accomplish this is with the gate-wise method from [13,14], where noise is mitigated by inverting the noise channel of each gate separately, i.e., the cancellation of errors is performed for each gate on its own. Here the local noise channel is approximated in a way such that it can be easily inverted analytically, e.g., using Pauli twirling [14]. Gates are then sampled from the inverted noise channel by interpreting it as a quasiprobability distribution. Because in this gatewise approach every noisy gate has to be modified separately, the sign problem is exponentially large in the number of gates, limiting the practicality of the mitigation. The success of the gatewise approach resulted in a large body of work concerning these methods [15–23], including extensions for simultaneous mitigation of multiple gates by Pauli-twirling entire layers [24,25] or variationally constructing a mitigating matrix product operator [26].

In principle, errors during the execution of a circuit can propagate and accumulate. These propagated errors can potentially blow up and lead to large errors for the circuit as a whole [27,28]. Here we introduce a mitigation technique that takes into account the propagation of errors, can be performed with a tunable number of extra gates, and works for non-Clifford local noise channels since the inversion of the accumulated global noise channel is implicit. We first execute the targeted noisy circuit completely, letting the noise propagate and accumulate, and only afterwards we apply an extra random circuit sampled from a quasiprobability distribution. We call the corresponding ensemble of random circuits a *denoiser* and we construct it such that upon averaging the accumulated errors cancel. Essentially, the denoiser inverts a global noise channel. Since we will construct it as a local brickwall circuit, following the classical preprocessing approach from [29], we call this *compressed* quantum error mitigation.

*Method.* Due to the inevitable coupling of a quantum processor to its environment, every qubit operation is affected by noise. Therefore, the simplest technique to minimize the impact of the resulting noise is to minimize the number of operations when performing a quantum algorithm. In [29] we showed that many-body time evolution operators can be efficiently compressed into brickwall circuits with high fidelity per gate.

In this Letter, we consider the noise explicitly by treating quantum operations as (generally nonunitary) quantum channels, corresponding to completely positive and trace preserving (CPTP) maps [30]. For example, instead of a noiseless two-qubit gate $G$, which acts on a quantum state $|\rho\rangle\rangle$ in superoperator form as $\mathcal{G}|\rho\rangle\rangle = G \otimes G^{*}|\rho\rangle\rangle$, we get the noisy channel $\tilde{\mathcal{G}} = \mathcal{N}\mathcal{G}$, where the noise channel $\mathcal{N}$ implements the two-qubit noise [31]. These channels are used to construct a "supercircuit" $\tilde{\mathcal{C}} = \prod_{i=1}^{N_{\mathcal{G}}} \tilde{\mathcal{G}}_i$, consisting of $N_{\mathcal{G}}$ channels, which is affected by multiqubit accumulated noise. This supercircuit encodes an ensemble of circuits [31]. For simplicity, we assume that the noisy channels $\tilde{\mathcal{G}}_i$ in each half brickwall layer are lattice inversion and translation invariant, such that we can construct a denoiser with these properties, limiting the number of variational parameters.

---

[*] david.luitz@uni-bonn.de

©2023 American Physical Society

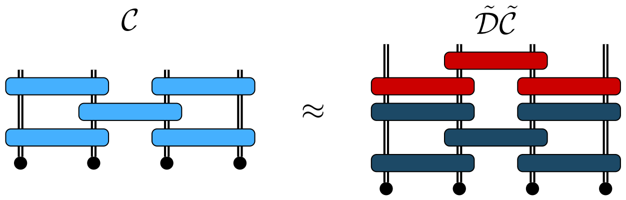


FIG. 1. Example of the quantum error mitigation procedure used in this work for the time evolution of the wave function of a spin chain. The ideal second-order Trotter supercircuit $\mathcal{C}$ of depth $M_{\mathrm{trot}} = 1$ (light blue) is approximated by applying a denoiser $\tilde{\mathcal{D}}$ of depth $M = 1$ (red) to the noisy Trotter supercircuit $\tilde{\mathcal{C}}$ (dark blue). Because the denoiser is applied after fully executing the noisy Trotter supercircuit, it represents an approximate inverse of the global noise channel with a precision tunable by the depth of the denoiser.

The purpose of quantum error mitigation is to modify the ensemble of circuits described by $\tilde{\mathcal{C}}$ in a way that we can use it to obtain the noiseless expectation values. In superoperator language, we do this by following the supercircuit $\tilde{\mathcal{C}}$ with a denoiser supercircuit $\tilde{\mathcal{D}}$, such that $\tilde{\mathcal{D}}\tilde{\mathcal{C}}$ is as close to the noiseless supercircuit $\mathcal{C} = C \otimes C^*$ as possible. Here $C$ is the target unitary circuit. Because the noise channel $\mathcal{N}$ is nonunitary, hence making the supercircuit $\tilde{\mathcal{C}}$ nonunitary, we need to use a nonunitary denoiser to retrieve the unitary $\mathcal{C}$.

We illustrate the mitigation procedure in Fig. 1, where a denoiser with one layer is used to mitigate errors for a second-order Trotter supercircuit with one layer. This circuit architecture is commonly used to simulate the time evolution of a quantum many-body system, until some time $t$, with controllable precision [29,32–42], and we will use it to benchmark the denoiser. In practice, we cannot directly implement a supercircuit, and so we have to utilize its interpretation as an ensemble of circuits. Essentially, after executing a shot of the noisy circuit we sample the denoiser and apply it. The goal is to construct the denoiser in a way that averaging over many of its samples cancels the accumulated errors and gives us a good approximation of the noiseless expectation values.

It should be noted that our approach requires more gate applications on the quantum processor than with the gatewise scheme, since there each sample from the mitigation quasiprobability distribution can be absorbed into the original circuit, whereas our approach increases the circuit depth. We take this into account by imposing the same noise on the denoiser. Furthermore, within our scheme, the dimensionality of the quasiprobabilistic mitigating ensemble can be controlled, in contrast to the gatewise approach where it is equal to the gate count.

To facilitate the stochastic interpretation we parametrize each two-qubit denoiser channel $\mathcal{G}_i$ as a sum of CPTP maps, such that we can sample the terms in this sum and execute the sampled gate on the quantum processor. Concretely, we use a trace preserving sum of a unitary and a nonunitary channel. For the unitary part we take a two-qubit unitary channel $\mathcal{U}(\vec{\phi}_i) = U(\vec{\phi}_i) \otimes U^*(\vec{\phi}_i)$, with $U(\vec{\phi}_i)$ a two-qubit unitary gate parametrized by $\vec{\phi}_i$. For this we take the two-qubit ZZ rotation $\exp[-i\alpha(\sigma_z \otimes \sigma_z)]$ with angle $\alpha$, which can be obtained from native gates on current hardware [43], and dress it with four general one-qubit unitaries, only two of which are independent if we want a circuit that is space inversion symmetric around every bond. The resulting gate has seven real parameters $\vec{\phi}_i$.

For the nonunitary part, which is essential because $\tilde{\mathcal{D}}$ has to cancel the nonunitary accumulated noise to obtain the noiseless unitary circuit, we use a general one-qubit measurement followed by conditional preparation channel $\mathcal{M}(\vec{\zeta}_i)|\rho\rangle\rangle = \sum_l K_l \otimes K_l^*|\rho\rangle\rangle$. It has Kraus operators $K_1 = |\psi_1\rangle\langle\,\psi\,|$ and $K_2 = |\psi_2\rangle\langle\,\bar{\psi}\,|$ if we measure in the orthonormal basis $\{|\psi\rangle, |\bar{\psi}\rangle\}$, where $|\bar{\psi}\rangle$ is uniquely defined by $|\psi\rangle$ as they are antipodal points on the Bloch sphere. If the measurement yields $|\psi\rangle$ we prepare $|\psi_1\rangle$ and if we measure $|\bar{\psi}\rangle$ we prepare $|\psi_2\rangle$. These states can be arbitrary points on the Bloch sphere, i.e., $|\psi_1\rangle = V(\vec{\kappa}_1)|0\rangle$, $|\psi_2\rangle = V(\vec{\kappa}_2)|0\rangle$, and $|\psi\rangle = V(\vec{\kappa}_3)|0\rangle$, with $V$ a general one-qubit unitary and each $\vec{\kappa}_i$ a three-dimensional vector, resulting in a real nine-dimensional $\vec{\zeta}_i$. This yields the two-qubit correlated measurement $\mathcal{M}(\vec{\zeta}_i) \otimes \mathcal{M}(\vec{\zeta}_i)$.

With these parts we construct the parametrization

$$\mathcal{G}_i = \eta_0 \mathcal{U}(\vec{\phi}_i) + \eta_1 \mathcal{M}(\vec{\zeta}_i) \otimes \mathcal{M}(\vec{\zeta}_i), \tag{1}$$

with coefficients $\eta_i \in \mathbb{R}$ that satisfy $\eta_0 + \eta_1 = 1$ because $\mathcal{G}_i$ is trace preserving. Note that here the tensor product symbol corresponds to combining two one-qubit channels to make a two-qubit channel, whereas in most of the paper it is used to link the column and row indices of a density matrix. We construct the denoiser from the noisy channels $\tilde{\mathcal{G}}_i = \mathcal{N}\mathcal{G}_i$. With this parametrization one denoiser channel has 17 independent real parameters, such that a denoiser of depth $M$, i.e., consisting of $M$ brickwall layers, has $34M$ real parameters (we use one unique channel per half brickwall layer). For reference, a general channel has $544M$ parameters.

To determine the mitigated expectation values we use the full expression

$$\langle\hat{O}\rangle_{p=0} = \langle\langle\mathbb{1}|(\hat{O}\otimes\mathbb{1})\mathcal{C}|\rho_0\rangle\rangle \approx \langle\langle\mathbb{1}|(\hat{O}\otimes\mathbb{1})\tilde{\mathcal{D}}\tilde{\mathcal{C}}|\rho_0\rangle\rangle, \tag{2}$$

where $|\rho_0\rangle\rangle$ is the initial state and $|\mathbb{1}\rangle\rangle$ is the vectorized identity operator on the full Hilbert space. To evaluate this on a quantum processor, we use the stochastic interpretation of (1) to resample (2). In particular, from each channel (1) we get a unitary with probability $p_0 = |\eta_0|/\gamma$ and a measurement followed by conditional preparation with probability $p_1 = |\eta_1|/\gamma$. Here $\gamma = |\eta_0| + |\eta_1|$ is the sampling overhead, which characterizes the magnitude of the sign problem from negative $\eta_i$ [13,14,18,20,44,45]. For quasiprobability distributions, i.e., with $\gamma > 1$, every denoiser sample has an extra sign $\mathrm{sgn}(\eta) = \prod_{g=1}^{N_{\mathcal{G}}} \mathrm{sgn}(\eta_g)$, where $\mathrm{sgn}(\eta_g)$ is the sign of the sampled coefficient of the $g$th channel. $\gamma = 1$ means that all signs are positive. Observables $\langle\hat{O}\rangle_{p=0}$ for the noiseless circuit are then approximated by resampling the observables from the denoiser ensemble [13]

$$\langle\hat{O}\rangle_{p=0} \approx \gamma\langle\mathrm{sgn}(\eta)\hat{O}\rangle_p, \tag{3}$$

where $\gamma = \prod_{g=1}^{N_{\mathcal{G}}} \gamma_g$ is the overall sampling overhead, with $\gamma_g$ the overhead of the $g$th gate. Clearly, a large $\gamma$ implies a large variance of $\langle\hat{O}\rangle_{p=0}$ for a given number of samples, with accurate estimation requiring the cancellation of large signed terms.

The number of samples required to resolve this cancellation of signs is bounded by Hoeffding's inequality, which states that a sufficient number of samples to estimate $\langle \hat{O} \rangle_{p=0}$ with error $\delta$ at probability $1 - \omega$ is bounded by $(2\gamma^2/\delta^2)\ln(2/\omega)$ [44]. Since $\gamma$ scales exponentially in $\gamma_g$, it is clear that a denoiser with large $M$ and $\gamma \gg 1$ will require many samples.

We observed that decompositions with $\gamma > 1$ are crucial for an accurate denoiser. Restricting to $\gamma = 1$ leads to large infidelity and no improvement upon increasing the number of terms in (1) or the depth $M$ of the denoiser. Simply put, probabilistic error cancellation of gate noise introduces a sign problem and it is crucial to find optimal parametrizations (1) which minimize $\gamma$ to make the approach scalable. This issue arises in all high performance error mitigation schemes [13,20,24,44], because the inverse of a physical noise channel is unphysical and cannot be represented as a positive sum over CPTP maps. This is clearly visible in the spectra of the denoiser, which lies outside the unit circle (cf. Fig. 4). This makes the tunability of the number of gates in each denoiser sample a crucial ingredient, which allows control over the sign problem, because we can freely choose the $\eta_i$ in (1).

For the parametrization (1) of denoiser channels, we try to find a set of parameters for error mitigation by minimizing the normalized Frobenius distance between the noiseless and denoised supercircuits [29]

$$\epsilon = ||\mathcal{C} - \tilde{\mathcal{D}}\tilde{\mathcal{C}}||_F^2/4^L, \tag{4}$$

which bounds the distance of output density matrices and becomes zero for perfect denoising.

We carry out the minimization of $\epsilon$ on a classical processor, using gradient descent with the differential programming algorithm from [29]. Instead of explicitly calculating the accumulated global noise channel and subsequently inverting it, we approximate the noiseless supercircuit $\mathcal{C}$ with the denoised supercircuit $\tilde{\mathcal{D}}\tilde{\mathcal{C}}$, effectively yielding a circuit representation $\mathcal{D}$ of the inverse noise channel.

*Results.* To benchmark the denoiser we apply it to the second-order Trotter circuits of the spin-1/2 Heisenberg chain with periodic boundary conditions (PBC)

$$H = \sum_{i=1}^{L} \left(\sigma_1^i \sigma_1^{i+1} + \sigma_2^i \sigma_2^{i+1} + \sigma_3^i \sigma_3^{i+1}\right), \tag{5}$$

where $\sigma_\alpha^i = (\mathbb{1}^i, \sigma_x^i, \sigma_y^i, \sigma_z^i)$ is the Pauli algebra acting on the local Hilbert space of site $i$. A second-order Trotter circuit for evolution time $t$ with depth $M_{\mathrm{trot}}$ consists of $M_{\mathrm{trot}} - 1$ half brickwall layers with time step $t/M_{\mathrm{trot}}$ and two layers with half time step [29,34]. We consider circuits that are affected by uniform depolarizing noise with probability $p$ for simplicity, but our approach can be used for any non-Clifford noise. The two-qubit noise channel is

$$\mathcal{N} = \left(1 - \frac{16p}{15}\right)\mathbb{1} + \frac{p}{15} \bigotimes_{j=i}^{i+1} \left(\sum_{\alpha=0}^{3} \sigma_\alpha^j \otimes \sigma_\alpha^{j*}\right), \tag{6}$$

which acts on neighboring qubits $i$ and $i+1$ and is applied to each Trotter and denoiser gate and $p = 0.01$ unless stated otherwise. We study circuits with depths $M_{\mathrm{trot}} = 16, 32, 64$ for evolution times $t = 0.5, 1, \ldots, 5$ and denoisers $\tilde{\mathcal{D}}$ with depths $M = 1, 2, 4, 6, 8$.

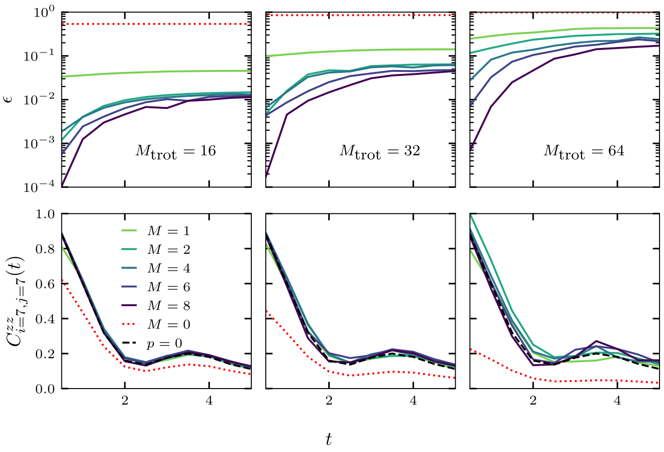


FIG. 2. Normalized distance $\epsilon$ between the denoised Trotter supercircuit $\tilde{\mathcal{D}}\tilde{\mathcal{C}}$ and the noiseless Trotter supercircuit $\mathcal{C}$ (top panels), at evolution times $t = 0.5, 1, \ldots, 5$, and the two-point $z$-spin correlator $C^{zz}_{i=L/2,j=L/2}(t)$ of a spin on the middle site at times 0 and $t$ (bottom panels), for the infinite temperature initial state. We consider denoisers with depths $M = 1, 2, 4, 6, 8$ and second-order Trotter circuits with depths $M_{\mathrm{trot}} = 16, 32, 64$. In the top panels we use a Heisenberg chain with $L = 8$ and in the bottom panels with $L = 14$, both with periodic boundary conditions. All gates are affected by two-qubit depolarizing noise with $p = 0.01$. The nondenoised results are labeled with $M = 0$ and the noiseless values with $p = 0$.

In the top panels of Fig. 2 we show $\epsilon$ (4) for a chain of size $L = 8$ as a function of time $t$. Here it can be seen that even for $M_{\mathrm{trot}} = 32$ a denoiser with $M = 1$ already improves $\epsilon$ by roughly an order of magnitude at all considered $t$. Depending on $M_{\mathrm{trot}}$ and $t$, further increasing $M$ lowers $\epsilon$, with the biggest improvements occurring for high precision Trotter circuits with large depth $M_{\mathrm{trot}} = 64$ and short time $t = 0.5$, where the Trotter gates are closer to the identity than in the other cases. At the other extreme, for $M_{\mathrm{trot}} = 16$ the improvements are relatively small upon increasing $M > 2$. In all cases the denoiser works better at early times than at late times, again indicating that it is easier to denoise Trotter gates that are relatively close to the identity.

To probe the accuracy of the denoiser on quantities that do not enter the optimization, as a first test we consider the two-point correlator between spins at different times [46]

$$C_{ij}^{zz}(t) = \langle\langle \mathbb{1} | (\sigma_i^z \otimes \mathbb{1}) \tilde{\mathcal{D}}\tilde{\mathcal{C}}(t) (\sigma_j^z \otimes \mathbb{1}) | \mathbb{1} \rangle\rangle / 2^L, \tag{7}$$

where we have chosen the infinite temperature initial state and $\tilde{\mathcal{C}}(t)$ is the Trotter supercircuit for time $t$. In the bottom panels of Fig. 2 we show $C^{zz}_{i=L/2,j=L/2}(t)$ for the supercircuits from the upper panels, now for a $L = 14$ chain. Here we see that at $M_{\mathrm{trot}} = 16$ we can retrieve the noiseless values already with $M = 1$, but that increasing $M_{\mathrm{trot}}$ makes this more difficult. At $M_{\mathrm{trot}} = 64$ we see larger deviations and improvement upon increasing $M$ is less stable, but nonetheless we are able to mitigate errors to a large extent.

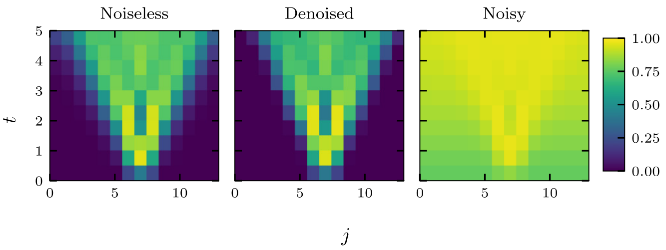


FIG. 3. Out-of-time-ordered correlator $C^{\text{otoc}}_{i=L/2,j}(t)$ as a function of the operator position $j$ and time $t$, for the infinite temperature initial state, for a denoised second-order Trotter supercircuit with Trotter depth $M_{\text{trot}} = 32$ and denoiser depth $M = 2$. We consider evolution times $t = 0.5, 1, \dots, 5$, for the periodic $L = 14$ Heisenberg chain that is affected by two-qubit depolarizing noise with $p = 0.01$.

As a further test, we compute the out-of-time-ordered correlator (OTOC) [29,39,47–50]

$$C^{\text{otoc}}_{ij}(t) = \mathrm{Re}\langle\langle \mathbb{1} | (\sigma_j^{z\dagger} \otimes \mathbb{1}) \tilde{\mathcal{D}}\tilde{\mathcal{C}}(-t) \times (\sigma_i^z \otimes \sigma_i^{z*}) \tilde{\mathcal{D}}\tilde{\mathcal{C}}(t) (\sigma_j^z \otimes \mathbb{1}) | \mathbb{1}\rangle\rangle / 2^L. \quad (8)$$

In Fig. 3 we show the results for $i = L/2$, for a Trotter circuit with depth $M_{\text{trot}} = 32$ and a denoiser with depth $M = 2$. Here we see that a denoiser with $M \ll M_{\text{trot}}$ is able to recover the light cone of correlations, which are otherwise buried by the noise. In the Supplemental Material [51] we consider how the denoiser performs at different noise levels $p$ and how the denoised supercircuits perform under stacking. There we also calculate domain wall magnetization dynamics and show the distribution of the optimized denoiser parameters and the sampling overhead associated to the denoiser as a whole.

In Fig. 4 we show the eigenvalues of the noisy supercircuits for a noisy second-order Trotter supercircuit with $M_{\text{trot}} = 16$ at $t = 1$ (left), the corresponding optimized denoiser with $M = 4$ (center), and the denoised supercircuit (right). The eigenvalues $\lambda$ of a unitary supercircuit lie on the unit circle and in the presence of dissipation they are pushed to the center. We see that the spectrum of the denoiser lies outside the unit circle, making it an unphysical channel which cures the effect of the noise on the circuit, such that the spectrum of the denoised circuit is pushed back to the unit circle. The noiseless eigenvalues are shown as blue bars, making it clear that the denoiser is able to recover the noiseless eigenvalues from the noisy circuit. In the Supplemental Material [51] we show the spectra for a $p = 0.036$ denoiser, where we observe a clustering of eigenvalues reminiscent of Refs. [52–54]. There we also investigate the channel entropy of the various supercircuits [55,56].

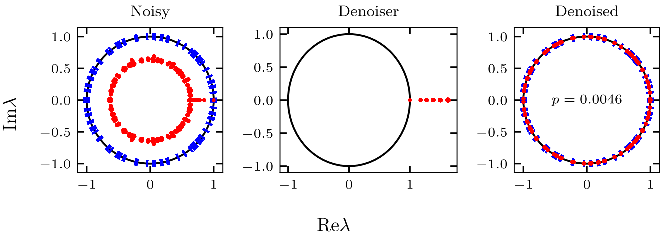


FIG. 4. Complex eigenvalues $\lambda$ of the noisy second-order Trotter supercircuit with $M_{\text{trot}} = 16$ at time $t = 1$ (left), the corresponding optimized denoiser with $M = 4$ (center), and the denoised Trotter supercircuit (right). The Trotter circuit is for a $L = 6$ Heisenberg model with PBC and all two-qubit channels are affected by depolarizing noise with $p = 0.0046$. The unit circle, on which unitary eigenvalues must lie, is shown in black and the noiseless eigenvalues are shown as blue bars. It is evident that the denoiser recovers all the noiseless eigenvalues from the noisy circuit.

*Conclusion.* We have introduced a probabilistic error cancellation scheme, where a classically determined denoiser mitigates the accumulated noise of a (generally non-Clifford) local noise channel. The required number of mitigation gates, i.e., the dimensionality of the corresponding quasiprobability distribution, is tunable and the parametrization of the corresponding channels provides control over the sign problem that is inherent to probabilistic error cancellation. We have shown that a denoiser with one layer can already significantly mitigate errors for second-order Trotter circuits with up to 64 layers.

This effectiveness of low-depth compressed circuits for denoising, in contrast with the noiseless time evolution operator compression from [29], can be understood from the nonunitarity of the denoiser channels. In particular, measurements can have nonlocal effects, since the measurement of a single qubit can reduce some highly entangled state (e.g., a GHZ state) to a product state, whereas in unitary circuits the spreading of correlations forms a light cone.

To optimize a denoiser with convenience at $L > 8$, the optimization can be formulated in terms of matrix product operators [26,29] or channels [16], which is convenient because the circuit calculations leading to the normalized distance $\epsilon$ and its gradient are easily formulated in terms of tensor contractions and singular value decompositions [29,57]. This provides one route to a practical denoiser, which is relevant because the targeted noiseless circuit and the accompanying noisy variant in (4) need to be simulated classically, confining the optimization procedure to limited system sizes with an exact treatment or limited entanglement with tensor networks. Nonetheless, we can use, e.g., matrix product operators to calculate (4) for some relatively small $t$, such that the noiseless and denoised supercircuits in (4) have relatively small entanglement, and then stack the final denoised supercircuit on a quantum processor to generate classically intractable states. Analogously, we can optimize the channels exactly at some classically tractable size and then execute them on a quantum processor with larger size. Both approaches are limited by the light cone of many-body correlations, as visualized in Fig. 3, because finite-size effects appear when the light-cone width becomes comparable with system size.

*Acknowledgments.* We are grateful for extensive discussions with D. Hahn. This project was supported by the Deutsche Forschungsgemeinschaft (DFG) through the cluster of excellence ML4Q (EXC 2004, Project-id No. 390534769). We also acknowledge support from the QuantERA II Programme that has received funding from the European Union's Horizon 2020 research innovation programme (GA 101017733) and from the DFG through the project DQUANT (Project-id No. 499347025).

[1] A. Y. Kitaev, Quantum measurements and the Abelian stabilizer problem, Electron. Colloquium Comput. Complex. **TR96-003** (1996).

[2] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM J. Comput. **26**, 1484 (1997).

[3] S. Ebadi *et al.*, Quantum optimization of maximum independent set using Rydberg atom arrays, Science **376**, 1209 (2022).

[4] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[5] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys. **21**, 467 (1982).

[6] S. Lloyd, Universal quantum simulators, Science **273**, 1073 (1996).

[7] B. Stefano, V. Filippo, and C. Giuseppe, An efficient quantum algorithm for the time evolution of parameterized circuits, Quantum **5**, 512 (2021).

[8] M. C. Bañuls *et al.*, Simulating lattice gauge theories within quantum technologies, Eur. Phys. J. D **74**, 165 (2020).

[9] P. Scholl *et al.*, Quantum simulation of 2d antiferromagnets with hundreds of Rydberg atoms, Nature (London) **595**, 233 (2021).

[10] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[11] A. R. Calderbank and P. W. Shor, Good quantum error-correcting codes exist, Phys. Rev. A **54**, 1098 (1996).

[12] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, Phys. Rev. A **52**, R2493 (1995).

[13] K. Temme, S. Bravyi, and J. M. Gambetta, Error Mitigation for Short-Depth Quantum Circuits, Phys. Rev. Lett. **119**, 180509 (2017).

[14] S. Endo, S. C. Benjamin, and Y. Li, Practical Quantum Error Mitigation for Near-Future Applications, Phys. Rev. X **8**, 031027 (2018).

[15] J. Vovrosh, K. E. Khosla, S. Greenaway, C. Self, M. S. Kim, and J. Knolle, Simple mitigation of global depolarizing errors in quantum simulations, Phys. Rev. E **104**, 035309 (2021).

[16] S. Filippov, B. Sokolov, M. A. C. Rossi, J. Malmi, E.-M. Borrelli, D. Cavalcanti, S. Maniscalco, and G. García-Pérez, Matrix product channel: Variationally optimized quantum tensor network to mitigate noise and reduce errors for the variational quantum eigensolver, arXiv:2212.10225.

[17] N. Cao, J. Lin, D. Kribs, Y.-T. Poon, B. Zeng, and R. Laflamme, Nisq: Error correction, mitigation, and noise simulation, arXiv:2111.02345.

[18] C. Piveteau, D. Sutter, and S. Woerner, Quasiprobability decompositions with reduced sampling overhead, npj Quantum Inf. **8**, 12 (2022).

[19] M. Gutiérrez, C. Smith, L. Lulushi, S. Janardan, and K. R. Brown, Errors and pseudothresholds for incoherent and coherent noise, Phys. Rev. A **94**, 042338 (2016).

[20] J. Jiaqing, W. Kun, and W. Xin, Physical implementability of linear maps and its application in error mitigation, Quantum **5**, 600 (2021).

[21] E. Magesan, D. Puzzuoli, C. E. Granade, and D. G. Cory, Modeling quantum noise for efficient testing of fault-tolerant circuits, Phys. Rev. A **87**, 012324 (2013).

[22] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, arXiv:2210.00921.

[23] S. Ferracin, A. Hashim, J.-L. Ville, R. Naik, A. Carignan-Dugas, H. Qassim, A. Morvan, D. I. Santiago, I. Siddiqi, and J. J. Wallman, Efficiently improving the performance of noisy quantum computers, arXiv:2201.10672.

[24] E. van den Berg, Z. K. Minev, A. Kandala, and K. Temme, Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors, Nat. Phys. (2023), doi:10.1038/s41567-023-02042-2.

[25] B. McDonough, A. Mari, N. Shammah, N. T. Stemen, M. Wahl, W. J. Zeng, and P. P. Orth, Automated quantum error mitigation based on probabilistic error reduction, in *2022 IEEE/ACM Third International Workshop on Quantum Computing Software (QCS)* (IEEE, New York, 2022).

[26] Y. Guo and S. Yang, Quantum error mitigation via matrix product operators, PRX Quantum **3**, 040313 (2022).

[27] S. Flannigan, N. Pearson, G. H. Low, A. Buyskikh, I. Bloch, P. Zoller, M. Troyer, and A. J. Daley, Propagation of errors and quantitative quantum simulation with quantum advantage, Quantum Sci. Technol. **7**, 045025 (2022).

[28] P. M. Poggi, N. K. Lysne, K. W. Kuper, I. H. Deutsch, and P. S. Jessen, Quantifying the sensitivity to errors in analog quantum simulation, PRX Quantum **1**, 020308 (2020).

[29] M. S. J. Tepaske, D. Hahn, and D. J. Luitz, Optimal compression of quantum many-body time evolution operators into brickwall circuits, SciPost Phys. **14**, 073 (2023).

[30] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, UK, 2011).

[31] D. Aharonov, A. Kitaev, and N. Nisan, Quantum Circuits with Mixed States, in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing, STOC '98* (Association for Computing Machinery, New York, NY, 1998), pp. 20–30.

[32] H. F. Trotter, On the product of semi-groups of operators, Proc. Am. Math. Soc. **10**, 545 (1959).

[33] M. Suzuki, Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems, Commun. Math. Phys. **51**, 183 (1976).

[34] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, Time-evolution methods for matrix-product states, Ann. Phys. (NY) **411**, 167998 (2019).

[35] J. Ostmeyer, Optimised Trotter decompositions for classical and quantum computing, arXiv:2211.02691.

[36] C. Kargi, J. P. Dehollain, F. Henriques, L. M. Sieberer, T. Olsacher, P. Hauke, M. Heyl, P. Zoller, and N. K. Langford, Quantum chaos and universal Trotterisation behaviours in digital quantum simulations, *Quantum Information and Measurement VI 2021* (Optica Publishing Group, 2021).

[37] M. Heyl, P. Hauke, and P. Zoller, Quantum localization bounds trotter errors in digital quantum simulation, Sci. Adv. **5**, eaau8342 (2019).

[38] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Theory of Trotter Error with Commutator Scaling, Phys. Rev. X **11**, 011020 (2021).

[39] K. Hémery, F. Pollmann, and D. J. Luitz, Matrix product states approaches to operator spreading in ergodic quantum systems, Phys. Rev. B **100**, 104303 (2019).

[40] R. Mansuroglu, T. Eckstein, L. Nützel, S. A. Wilkinson, and M. J. Hartmann, Variational Hamiltonian simulation for trans-

lational invariant systems via classical pre-processing, Quantum Sci. Technol. **8**, 025006 (2023).
[41] H. Zhao, M. Bukov, M. Heyl, and R. Moessner, Making Trotterization adaptive for NISQ devices and beyond, arXiv:2209.12653.
[42] N. F. Berthusen, T. V. Trevisan, T. Iadecola, and P. P. Orth, Quantum dynamics simulations beyond the coherence time on NISQ hardware by variational Trotter compression, Phys. Rev. Res. **4**, 023097 (2022).
[43] I-C. Chen, B. Burdick, Y. Yao, P. P. Orth, and T. Iadecola, Error-mitigated simulation of quantum many-body scars on quantum computers with pulse-level control, Phys. Rev. Res. **4**, 043027 (2022).
[44] R. Takagi, Optimal resource cost for error mitigation, Phys. Rev. Res. **3**, 033178 (2021).
[45] Y. Guo and S. Yang, Noise effects on purity and quantum entanglement in terms of physical implementability, npj Quantum Inf. **9**, 11 (2023).
[46] M. Dupont and J. E. Moore, Universal spin dynamics in infinite-temperature one-dimensional quantum magnets, Phys. Rev. B **101**, 121106(R) (2020).
[47] Y.-L. Zhang, Y. Huang, and X. Chen, Information scrambling in chaotic systems with dissipation, Phys. Rev. B **99**, 014303 (2019).
[48] D. J. Luitz and Y. Bar Lev, Information propagation in isolated quantum systems, Phys. Rev. B **96**, 020406(R) (2017).
[49] J. Maldacena, S. H. Shenker, and D. Stanford, A bound on chaos, J. High Energy Phys. 08 (2016) 106.
[50] A. I. Larkin and Yu. N. Ovchinnikov, Quasiclassical method in the theory of superconductivity, Sov. JETP **28**, 1200 (1969).
[51] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.107.L201114 for the denoiser performance at various noise strengths; the denoiser spectra at high noise strength; the entropy of the various supercircuits that appeared in the main text; the performance of the denoised supercircuits upon stacking; histograms of the optimized denoiser parameters; the sampling overhead of the optimized denoisers; and simulations of domain wall dynamics as another test of the denoiser performance.
[52] K. Wang, F. Piazza, and D. J. Luitz, Hierarchy of Relaxation Timescales in Local Random Liouvillians, Phys. Rev. Lett. **124**, 100604 (2020).
[53] J. L. Li, D. C. Rose, J. P. Garrahan, and D. J. Luitz, Random matrix theory for quantum and classical metastability in local Liouvillians, Phys. Rev. B **105**, L180201 (2022).
[54] O. E. Sommer, F. Piazza, and D. J. Luitz, Many-body hierarchy of dissipative timescales in a quantum computer, Phys. Rev. Res. **3**, 023190 (2021).
[55] T. Zhou and D. J. Luitz, Operator entanglement entropy of the time evolution operator in chaotic systems, Phys. Rev. B **95**, 094206 (2017).
[56] W. Roga, Z. Puchała, Ł. Rudnicki, and K. Życzkowski, Entropic trade-off relations for quantum operations, Phys. Rev. A **87**, 032308 (2013).
[57] K. Noh, L. Jiang, and B. Fefferman, Efficient classical simulation of noisy random quantum circuits in one dimension, Quantum **4**, 318 (2020).

# Supplementary material – Compressed quantum error mitigation

Maurits S. J. Tepaske and David J. Luitz[*]
*Physikalisches Institut, Universität Bonn, Nußallee 12, 53115 Bonn, Germany*
(Dated: May 11, 2023)

In this supplementary material we further investigate the denoiser. We probe the denoiser performance at various noise strengths, consider the denoiser spectra at high noise strength, calculate the entropy of the various supercircuits that appeared in the main text, check how the denoised circuits perform under stacking, plot the optimized denoiser parameters in histograms, determine the sampling overhead of the optimized denoisers, and finally consider the simulation of domain wall dynamics as another test of the denoiser performance.

## DENOISER PERFORMANCE AT VARIOUS NOISE LEVELS

To probe how the denoiser performs at different noise strengths $p$, we take a second-order Trotter supercircuit of depth $M_{\text{trot}} = 16$ for the time evolution of the wave function to time $t = 1$, and optimize the denoiser at various noise strengths in the interval $p \in \left[10^{-3}, 10^{-1}\right]$. In Fig. 1 we show the normalized distance (left panel) and the $z$ spin correlator (right), for denoiser depths $M = 1, 2, 4, 6, 8$. For comparison, we show the results for the noisy limit, i.e. without a denoiser ($M = 0$, red dashed), and for the exact limit without noise ($p = 0$, black dashed).

The error of the entire circuit $\epsilon$ improves with denoiser depth $M$ for the full range of $p$, and depends roughly quadratically on $p$. This is illustrated with the purple dashed line in the left panel of Fig. 1. It is interesting to observe that even for larger noise strength $p$, the local observable $C^{zz}$ improves significantly even with denoisers of depth $M = 1$. For large noise strengths, we generally see that the optimization of the denoiser becomes difficult, leading to nonmonotonic behavior as a function of $p$, presumably because we do not find the global optimum of the denoiser.

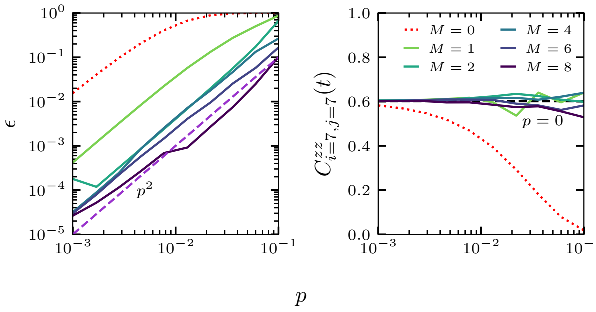


FIG. 1. The normalized distance $\epsilon$ (left) and $z$ spin correlator $C^{zz}_{i=L/2,j=L/2}$ (right), for a second-order Trotter supercircuit of depth $M_{\text{trot}} = 16$ for time $t = 1$, affected by various two-qubit depolarizing errors $p$. We compare the values obtained with and without a denoiser, i.e. $M > 0$ and $M = 0$, to the noiseless values ($p = 0$). The denoiser is affected by the same noise as the Trotter circuit. We consider denoisers with depths $M = 1, 2, 4, 6, 8$, and we use a $L = 8$ Heisenberg chain with PBC for the normalized distance, while for the correlator we use $L = 14$.

## SUPERCIRCUIT SPECTRA

It is interesting to analyze the spectra of the supercircuits considered in this work. As mentioned in the main text, the spectrum of the ideal, unitary supercircuit $\mathcal{C}$ lies on the unit circle. The comparison to this case is therefore instructive. In the main text, we showed an example of the spectra in Fig. 4 for moderate noise strength. Here, we show additional data for stronger noise $p = 0.036$ in Fig. 2 for a denoiser with $M = 4$ layers, optimized to mitigate errors for a second-order Trotter supercircuit with $M_{\text{trot}} = 16$ layers at time $t = 1$.

The eigenvalues $\lambda$ of the noisy supercircuit $\tilde{\mathcal{C}}$ are clustered close to zero, far away from the unit circle (except for $\lambda = 1$), showing that the circuit is strongly affected by the noise. To mitigate the impact of the noise, the

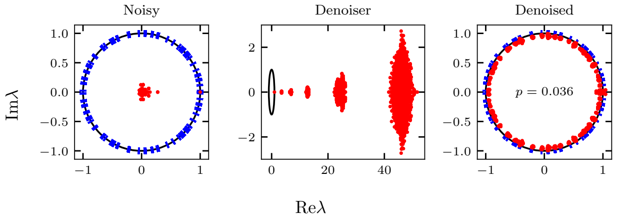


FIG. 2. The complex eigenvalues $\lambda$ of the noisy second-order Trotter supercircuit with $M_{\text{trot}} = 16$ at time $t = 1$ (left), the corresponding optimized denoiser with $M = 4$ (center), and the denoised Trotter supercircuit (right). The Trotter circuit is for a $L = 6$ Heisenberg model with PBC, and all two-qubit channels are affected by depolarizing noise with $p = 0.036$. The unit circle, on which unitary eigenvalues must lie, is shown in black, and the noiseless eigenvalues are shown as blue bars. It is clear that the denoiser recovers with high accuracy the noiseless eigenvalues from the noisy circuit.

---

[*] david.luitz@uni-bonn.de

denoiser consequently has to renormalize the spectrum strongly. If it accurately represents the inverse of the global noise channel, its spectrum has to lie far outside the unit circle, which is the case. Interestingly, we observe a clustering of eigenvalues which is reminiscent to the spectra found in [1–3]. By comparison to these works, we suspect that this is due to the local nature of the denoiser, and warrants further investigation.

The right panel of Fig. 2 shows the result of the denoiser, pushing the eigenvalues back to the unit circle, nearly with the exact same distribution along the circle as the noiseless eigenvalues (blue bars). Due to the strong noise, this is not achieved perfectly, and it is clear that this cannot work in principle if the global noise channel has a zero eigenvalue.

## SUPERCIRCUIT ENTROPIES

The complexity of an operator can be quantified by its operator entanglement entropy [4]. Here we calculate the half-chain channel entanglement entropy $S$ [5] of the noiseless $\mathcal{C}$, noisy $\tilde{\mathcal{C}}$, denoiser $\tilde{\mathcal{D}}$, and denoised $\tilde{\mathcal{D}}\tilde{\mathcal{C}}$ supercircuits. We define $S$ as the entanglement entropy of the state that is related to a supercircuit $\mathcal{C}$ via the Choi-Jamiołkowski isomorphism, i.e. $\psi_{\mathcal{C}} = \chi_{\mathcal{C}}/N$, where the process matrix $\chi_{\mathcal{C}}^{ab,cd} = \mathcal{C}^{ac,bd}$ is simply a reshaped supercircuit and $N$ ensures normalization. Then we have $S = -\mathrm{Tr}\left[\psi_{\mathcal{C}} \ln \psi_{\mathcal{C}}\right]$. This entropy measure is a particular instance of the "exchange entropy", which characterizes the information exchange between a quantum system and its environment [5].

In Fig. 3 we plot the various $S$ for a second-order Trotter circuit with $M_{\mathrm{trot}} = 16$ at $t = 2$, for a denoiser with $M = 4$, both affected by two-qubit depolarizing noise with $p \in [10^{-3}, 10^{-1}]$. The Trotter circuit is for a Heisenberg model with $L = 6$ and PBC. We see that at large $p$, the noise destroys entanglement in the noisy supercircuit, and that the denoiser $S$ increases to correct for this, such that the denoised supercircuit recovers the noiseless $S$.

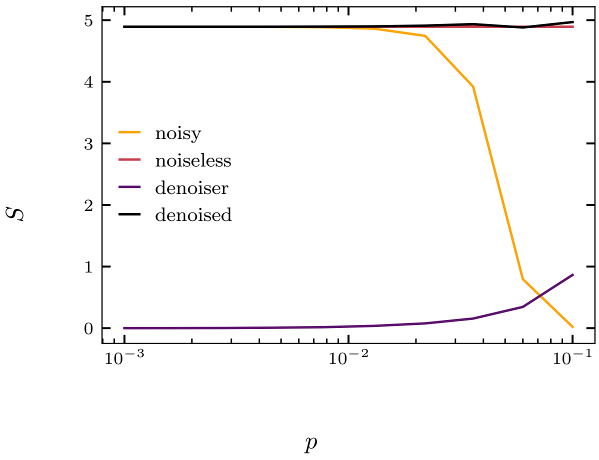


FIG. 3. The half-chain channel entanglement entropy $S$ at different two-qubit depolarizing noise strengths $p$, for a second-order Trotter supercircuit with $M_{\mathrm{trot}} = 16$ and $t = 2$, for a $M = 4$ denoiser. The Trotter circuit is for a Heisenberg model with PBC of size $L = 6$. The different curves correspond to the different supercircuits, i.e. the noisy supercircuit, the denoiser, the corresponding denoised supercircuit, and the noiseless variant.

## STACKING DENOISED SUPERCIRCUITS

Here we investigate how denoised supercircuits perform upon repeated application. We optimize the denoiser for a Trotter supercircuit for a fixed evolution time $t$. Then, to reach later times, we stack the denoised supercircuit $n$ times to approximate the evolution up to time $nt$:

$$\mathcal{C}(nt) \approx \left(\tilde{\mathcal{D}}(t)\tilde{\mathcal{C}}(t)\right)^n \tag{1}$$

In Fig. 5 we stack a denoised $t = 1$ supercircuit up to $n = 20$ times and calculate the correlation function, defined in the main text, for the middle site. We consider Trotter depths $M_{\mathrm{trot}} = 8, 16, 32, 64$ and denoiser depths $M = 1, 2, 4, 6, 8$, for a $L = 14$ Heisenberg chain with $p = 0.01$ depolarizing two-qubit noise. The noisy results correspond to $M = 0$ and the noiseless results to $p = 0$. In Fig. 4 we calculate the OTOC, defined in the main text, with stacked time evolution for a denoised $t = 2$ supercircuit with $M_{\mathrm{trot}} = 32$ and $M = 2$, stacked up to ten times. We see that the stacked supercircuit performs very well, and the additional precision obtained by using deep denoisers ($M = 8$) pays off for long evolution times, where we see convergence to the exact result (black dashed lines in Fig. 5) as a function of $M$.

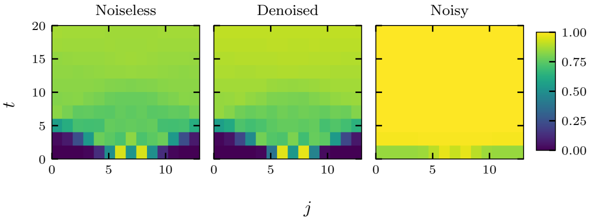


FIG. 4. The out-of-time-ordered correlator $C^{\mathrm{otoc}}_{i=L/2,j}(t)$ as a function of the operator position $j$ and stacked time $t$, for the infinite temperature initial state, for a denoised second-order Trotter supercircuit with Trotter depth $M_{\mathrm{trot}} = 32$ and denoiser depth $M = 2$. It is optimized at $t = 2$ and stacked up to ten times. The calculations are for the periodic $L = 14$ Heisenberg chain that is affected by two-qubit depolarization with $p = 0.01$. The denoiser is affected by the same noise.

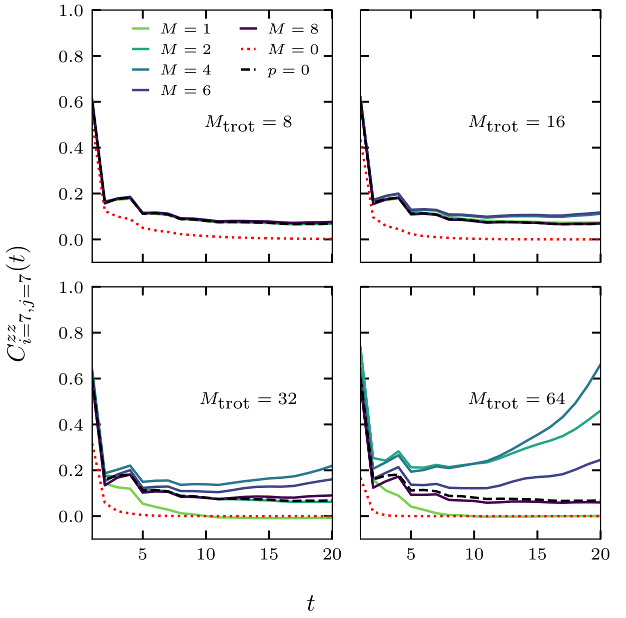


FIG. 5. The two-point $z$-spin correlator $C^{zz}_{i=L/2,j=L/2}(t)$ of a spin on the middle site at times 0 and $t$, for the infinite temperature initial state, for denoised second-order Trotter supercircuits that are optimized at evolution time $t = 1$ and then stacked up to twenty times. We use Trotter depths $M_{\text{trot}} = 8, 16, 32, 64$ and denoiser depths $M = 1, 2, 4, 6, 8$. The calculations were performed for a periodic Heisenberg model with $L = 14$ and PBC, affected by two-qubit depolarizing noise with strength $p = 0.01$, which also affects the denoiser. The non-denoised results are labelled with $M = 0$, and the noiseless results with $p = 0$. The panels are arranged as $M_{\text{trot}} = 8, 16, 32, 64$ for top left, top right, bottom left, bottom right, respectively.

## DISTRIBUTION OF OPTIMIZED ZZ CHANNELS

The costliest and most noise-susceptible operation is the two-qubit $ZZ$ rotation with angle $\alpha$, which is the foundation of the unitary piece in our channel parameterization, defined in the main text. For completeness, we here present the $\alpha$ angles of the optimized denoisers. The results are shown in Fig. 6, which contains histograms for the channel count $N_{\mathcal{G}}$ versus $\alpha$. The histograms are stacked, with the lightest color corresponding to the angles of the denoiser at $t = 0.5$ and the darkest at $t = 5$. The top four panels are for a denoiser with $M = 2$ and the bottom four with $M = 8$. We consider $M_{\text{trot}} = 8, 16, 32, 64$. We see that in both cases the distribution widens upon increasing $M_{\text{trot}}$, indicating that the unitary channels start deviating more from the identity. Moreover, while the $M = 2$ denoisers in all cases except $M_{\text{trot}} = 64$ have ZZ contributions close to the identity, this is clearly not the case for $M = 8$.

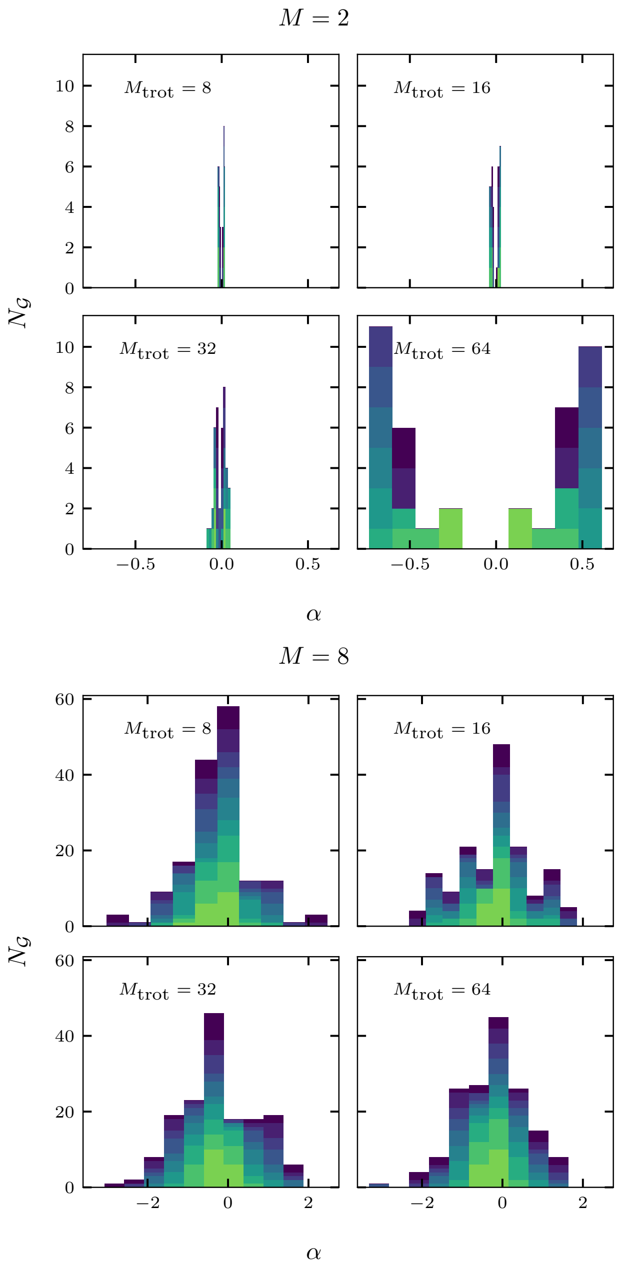


FIG. 6. The distribution of the ZZ angle $\alpha$ of $M = 2$ denoisers (top panels) and $M = 8$ denoisers (bottom panels), with the lightest color corresponding to the denoiser for the Trotter supercircuit with $t = 0.5$, and the darkest color with $t = 5$. As usual, we consider the Heisenberg model on a periodic chain, and second-order Trotter supercircuits with depths $M_{\text{trot}} = 8, 16, 32, 64$, which together with the denoiser is affected by a two-qubit depolarizing noise with $p = 0.01$. The panels are arranged as $M_{\text{trot}} = 8, 16, 32, 64$ for top left, top right, bottom left, bottom right, respectively.

## SAMPLING OVERHEAD OF OPTIMIZED DENOISERS

For simplicity, we did not focus on obtaining denoisers with the smallest sampling overhead $\gamma$, which is required to minimize the sign problem and hence ease the sampling of mitigated quantities. Instead, we let the optimization freely choose the $\eta_i$ in the denoiser parameterization, as defined in the main text. In Fig. 7 we show the sampling overhead of the denoisers from Fig. 2 of the main text. We see that for $M = 1$ and $M = 2$ the sampling overhead is relatively small and uniform across the different $t$, whereas for $M > 2$ the optimization sometimes yields a denoiser with large $\gamma$ and other times with small $\gamma$. This could be related to the difference in $\alpha$ distributions from Fig. 6. The large fluctuations of $\gamma$ appears to stem from the difficulty in finding optimal deep denoisers, and our optimization procedure likely only finds a local minimum in these cases.

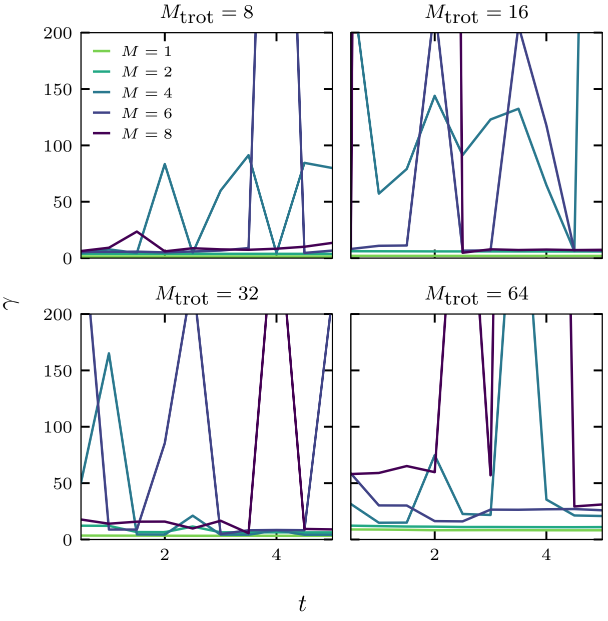


FIG. 7. The sampling overhead $\gamma$ of the optimized denoisers from Fig. 2 of the main text, with denoiser depths $M = 1, 2, 4, 6, 8$ and Trotter depths $M_{\mathrm{trot}} = 8, 16, 32, 64$ at times $t = 0.5, 1, ..., 5$, for the Heisenberg model on a chain with PBC affected by two-qubit depolarizing noise with $p = 0.01$. The panels are arranged as $M_{\mathrm{trot}} = 8, 16, 32, 64$ for top left, top right, bottom left, bottom right, respectively.

## DOMAIN WALL MAGNETIZATION

As another test of the denoiser performance we evolve the periodic $z$-spin domain wall $|\mathrm{dw}\rangle = \bigotimes_{i=1}^{L/2} |1\rangle \otimes \bigotimes_{i=L/2}^{L} |0\rangle$ and consider the domain wall magnetization

$$Z^{\mathrm{dw}}(t) = \sum_{i=1}^{L} (-1)^{\lfloor 2i/L \rfloor} \langle\langle \mathbb{1}|(\sigma_i^z \otimes \mathbb{1})\tilde{\mathcal{D}}\tilde{\mathcal{C}}(t)|\mathrm{dw}\rangle|\mathrm{dw}^*\rangle. \quad (2)$$

Here $\tilde{\mathcal{C}}(t)$ is the Trotter supercircuit for time $t$. In Fig. 8 we show $Z^{\mathrm{dw}}$ for the circuits from Fig. 2 of the main text, calculated for a $L = 14$ chain. As in our other tests, we see that at $M_{\mathrm{trot}} = 8$ we can recover the noiseless values already with $M = 1$, and that increasing $M_{\mathrm{trot}}$ makes this more difficult. At $M_{\mathrm{trot}} = 64$ we see larger deviations, and improvement upon increasing $M$ is less stable, but nevertheless we are able to mitigate errors to a large extent.

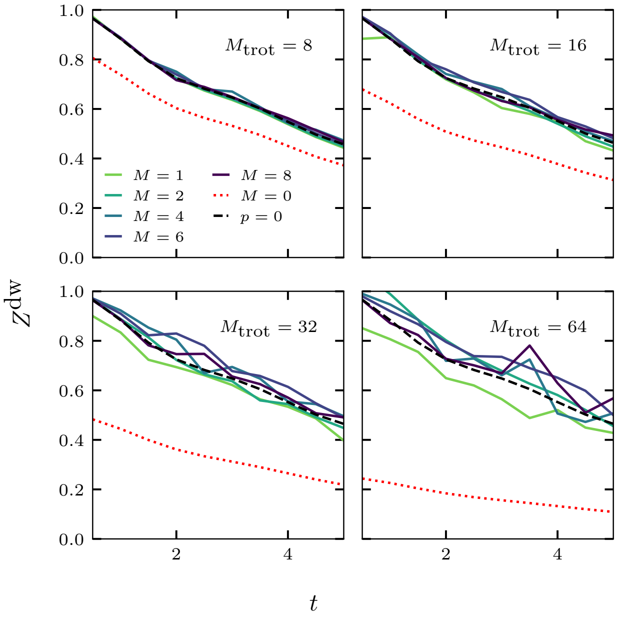


FIG. 8. The domain wall magnetization $Z^{\mathrm{dw}}$ after evolving a periodic density wall $|\mathrm{dw}\rangle|\mathrm{dw}^*\rangle$ with the denoised second-order Trotter supercircuits $\tilde{\mathcal{D}}\tilde{\mathcal{C}}$ from Fig. 2 of the main text. These supercircuits have various Trotter depths $M_{\mathrm{trot}} = 8, 16, 32, 64$, denoiser depths $M = 1, 2, 4, 6, 8$, and evolution times $t = 0.5, 1, ..., 5$, for the periodic $L = 14$ Heisenberg chain that is affected by two-qubit depolarizing noise of strength $p = 0.01$. The denoiser is affected by the same noise. The non-denoised results are labelled with $M = 0$ and the noiseless results with $p = 0$. The panels are arranged as $M_{\mathrm{trot}} = 8, 16, 32, 64$ for top left, top right, bottom left, bottom right, respectively. We see that the denoiser allows us to recover the noiseless behavior.

[1] K. Wang, F. Piazza, and D. J. Luitz, "Hierarchy of relaxation timescales in local random liouvillians," Phys. Rev. Lett. **124**, 100604 (2020).

[2] O. E. Sommer, F. Piazza, and D. J. Luitz, "Many-body hierarchy of dissipative timescales in a quantum computer," Physical Review Research **3**, 023190 (2021).

[3] J. L. Li, D. C. Rose, J. P. Garrahan, and D. J. Luitz, "Random matrix theory for quantum and classical metastability in local liouvillians," Phys. Rev. B **105**, L180201 (2022).

[4] T. Zhou and D. J. Luitz, "Operator entanglement entropy of the time evolution operator in chaotic systems," Physical Review B **95**, 094206 (2017).

[5] W. Roga, Z. Puchała, Ł. Rudnicki, and K. Życzkowski, "Entropic trade-off relations for quantum operations," Phys. Rev. A **87**, 032308 (2013).