A Mathematical Analysis of Coarsening Processes Driven by Vanishing

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

vorgelegt von
Constantin Eichenberg
geboren in
Henstedt-Ulzburg

Bonn, Mai 2021

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

Erste Gutachterin: Prof. Dr. Barbara Niethammer
Zweiter Gutachter: Prof. Dr. Juan J. L. Velázquez

Tag der Promotion: 27.09.2021
Erscheinungsjahr: 2022

Good bye then. Be safe, friend. Don't you dare go Hollow.

(Laurentius of the Great Swamp)

Danksagung

Die Fertigstellung dieser Dissertation wäre ohne die großartige Unterstützung von institutioneller und privater Seite sehr wahrscheinlich nicht möglich gewesen. Deshalb möchte im Folgenden meine Dankbarkeit zum Ausdruck bringen.

Zunächst einmal bedanke ich mich beim Institut für Angewandte Mathematik der Universität Bonn und der Deutschen Forschungsgemeinschaft für die Bereitstellung und Finanzierung meiner Stelle im Rahmen des *SFB 1060 - Die Mathematik der emergenten Effekte*. An dieser Stelle soll auch die Bonn International Graduate School nicht unerwähnt bleiben, die meine Teilnahme an internationalen Events wie Konferenzen, Workshops und Sommerschulen ermöglicht, sowie mit diversen Veranstaltungen vor Ort stets für ein freundliches und anregendes Klima unter den Doktoranden gesorgt hat.

Als nächstes möchte ich mich herzlich bei allen bedanken, die mich mich fachlich unterstützt haben. An erster Stelle natürlich bei meiner Doktormutter Prof. Dr. Barbara Niethammer, die mir nicht nicht nur in mathematischen Fragen zur Seite gestanden, sondern mir auch stets Zuspruch in Momenten gegeben hat, in denen ich an meinen Fähigkeiten gezweifelt habe. Ich bedanke mich bei Prof. Dr. Juan J. L. Velázquez mit seinem gefühlt unerschöpflichen Maß an Wissen, Kreativität und Optimismus, der insbesondere maßgebliche Ideen zu meiner ersten Veröffentlichung beigesteuert hat. Und last but not least möchte ich mich bei meinem ehemaligen Büronachbar und Co-Autor Prof. Dr. André Schlichting für die großartige Zusammenarbeit bedanken, die sowohl fachlich als auch persönlich sehr bereichernd war.

Zuletzt gibt es eine große Anzahl von Personen, die mir im Privaten den Rücken gestärkt haben. Meine Frau Sonja, die mich schon so lange begleitet und mir immer ein Licht in dunkler Stunde ist. Meine Eltern, die mich in allen Lebenslagen unterstützen. David, mit dem ich bei unzähligen Spaziergängen und Kaffeepausen über Mathematik, Philosophie und Politik diskutiert und viel Freude und Leid geteilt habe. Meine Bandkollegen, die mir durch das gemeinsame Musizieren stets einen wunderbaren Ausgleich zur Mathematik beschert haben. Es gibt viele weitere Freunde und Weggefährten, die mir im Kleinen und Großen geholfen haben und deren vollständige Aufzählung den Rahmen dieses Textes sprengen würde. Euch allen ist diese Arbeit gewidmet.

Abstract

In this thesis, we investigate several models of coarsening processes with the property that an agent, commonly referred to as particle, vanishes from the system once it reaches size zero and hence the average particle size tends to increase over time. We show how different types of interaction (local vs. mean-field interaction, deterministic vs. stochastic modeling) affect the general coarsening behavior and require different mathematical tools and strategies to analyze.

In Chapter 1 we investigate a class of mass transfer models on a one-dimensional lattice with nearest-neighbour interactions. The evolution is given by the discrete backward parabolic equation $\partial_t x = -\frac{\beta}{|\beta|}\Delta x^\beta$, with β in the fast diffusion regime $(-\infty, 0) \cup (0, 1]$. Particles with mass zero are deleted from the system, which leads to a coarsening of the mass distribution. The rate of coarsening suggested by scaling is $t^{\frac{1}{1-\beta}}$ if $\beta \neq 1$ and exponential if $\beta = 1$. We prove that such solutions actually exist by an analysis of the time-reversed evolution. In particular we establish positivity estimates and long-time equilibrium properties for discrete parabolic equations with initial data in $\ell^\infty_+(\mathbb{Z})$. The contents of this chapter were published in [16].

In Chapter 2 we consider a class of nonlocal coarsening models after Lifshitz, Slyozov and Wagner with singular particle interaction, which is the mean-field version of the model investigated in Chapter 1. For these equations we establish existence of general measure valued solutions by approximation with empirical measures, a result that extends the existing well-posedness theory for LSW equations. We use the size-ranking formulation of the equation to establish convergence of the approximate solutions in the L^1-Wasserstein distance. Furthermore, we show that there exists a one-parameter family of self-similar solutions, all of which have compact support but only one of them being smooth, a phenomenon that is typical for LSW models.

In Chapter 3 we study the exchange-driven growth model that arises as mean-field limit of a stochastic particle system and describes a process in which pairs of clusters exchange atomic particles. The rate of exchange between clusters here is given by the interaction kernel $K(k, l) = (k\,l)^\lambda$ for $\lambda \in [0, 2)$. We rigorously establish the coarsening rates and convergence to the self-similar profile found by Ben-Naim and Krapivsky [7] by linking the evolution to a discrete weighted heat equation on the positive integers by a nonlinear time-change. For this equation, we establish a new weighted Nash inequality that yields scaling-invariant decay and continuity estimates. Together with a replacement identity that links the discrete operator to its continuous analog, we derive a discrete-to-continuum scaling limit for the weighted heat equation and deduce coarsening rates and self-similar convergence of the exchange-driven growth model. The contents of this chapter are joint work with A. Schlichting and were published in [17].

Contents

Introduction

The ambition of an applied mathematician is to capture aspects of the observable reality of our universe in mathematical models that help to explain and predict future observations. The scope of this endevour is not limited to the classical natural sciences but is applicable to fields such as computer science and social sciences as well. It often turns out that there are fundamental mathematical ideas and concepts that can describe the essence of a variety of phenomena in a rough qualitative way. A prime example of this is the tendency of many dynamical systems to attain an equilibrium state in the limit of large times: Mixing hot and cold water results in the mixture attaining the lukewarm mean-temperature which is the stable equilibrium state, while biological ecosystems tend to approach equilibria that are infamous for being unstable.

In this thesis we want to explore dynamical systems that fall into the category of coarsening processes. As an abstract concept, coarsening describes the competition between certain agents, where an agent j is completely specified by a single non-negative number $x_j(t)$ (e.g physical mass or size, population, wealth, etc.) which can be an integer or non-integer quantity and evolves in time by interaction with other agents. The interaction between agents can be purely deterministic but may also involve a stochastic process. For convenience and in reference to classical physical models we call the agents particles from now on and interpret x_j as mass or size. The phenomenon of coarsening then refers to the growth of the average particle size or, more generally, a suitable length-scale of the system, as mass concentrates in larger and fewer particles (see Figure 0.1).

Coarsening processes are ubiquitous in nature and social environments. Classical examples from physics and chemistry include phase separation in mixtures [47, 59] and grain growth in polycrystals [34], while social phenomena like population dynamics [46] and wealth exchange [36] also exhibit coarsening dynamics.

The driving mechanisms for coarsening in the models that we consider in this thesis are relatively simple. We consider systems where the quantitiy $\sum_j x_j$, which represents the total mass of the system is constant in time. Then the interaction between particles should favor large particles over small ones, i.e particles with a large amount of mass tend to grow further, while small particles tend to shrink. And finally, particles that eventually reach size $x_j = 0$ cannot regain mass or interact with other particles and are practically deleted from the system. The vanishing of particles together with the

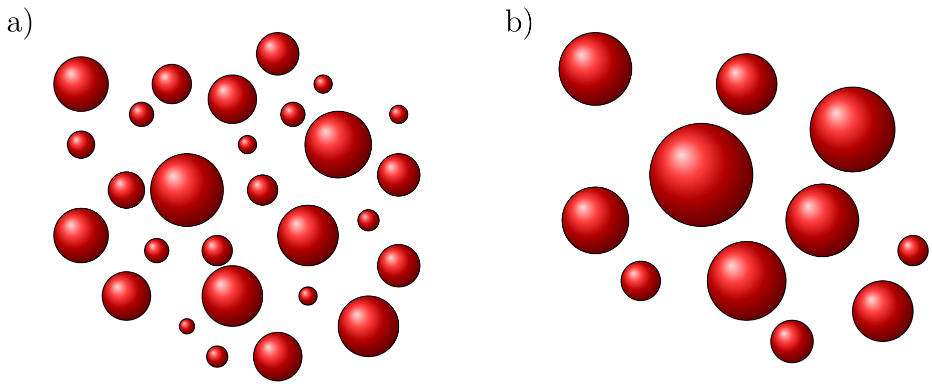

Figure 0.1.: Coarsening at an a) early and b) later stage.

conservation of mass automatically implies that the average particle size among living particles, i.e particles with positive size, is non-decreasing, hence we observe coarsening.

In applications, the number of particles is typically very large, hence for modeling purpose it is reasonable to consider systems with infinitely many particles. In this context it is often appropriate to express the statistics of the system in the size distribution function $f(t,x)$, i.e for every interval $I \subset [0,\infty)$ we have

$$\int_I f(t,x)\,\mathrm{d}x = \text{fraction of particles with size } x \in I.$$

The average particle size among living particles, which is a measure for the rate of coarsening, is then given by

$$\langle x \rangle = \frac{\int_{(0,\infty)} x f(t,x)\,\mathrm{d}x}{\int_{(0,\infty)} f(t,x)\,\mathrm{d}x},$$

and since the first moment of f is conserved, the growth of the average particle size is equivalent to the decay of the fraction of living particles of the system. The natural question is whether it is possible to predict the growth of $\langle x \rangle$ for large times. In this regard, all the models we discuss in this thesis have a scaling invariance that suggests an explicit coarsening rate of the form $\langle x \rangle \sim t^{\alpha}$ for some exponent α that is immediate from scaling. Even more remarkable, simulations and experimental observations of coarsening processes often show that the long-time dynamics attain a self-similar form, i.e there exists a profile Φ with first moment equal to one such that

$$f(t,x) \sim \rho t^{-2\alpha} \Phi(t^{-\alpha} x) \quad \text{for } t \gg 1,$$

where ρ is the first moment of f. This is a much stronger statement than merely estimating the coarsening rate, since it completely characterizes the dynamics for large times independent of the specific initial shape of the distribution.

As promising as the scaling heuristics often look, it turns out that estimating the coarsening rate is usually a difficult task. One reason for this is the existence of stationary states. In many deterministic coarsening models, configurations where all particles have the same size are stationary points of the evolution and can be attained from non-constant initial configurations. This means that for these “bad” initial data, the system stops coarsening at all after a finite time. Unfortunately, up to this point there exists no systematic approach to deal with this problem. The general opinion is that initial data that become stationary are somewhat rare, although it is not clear how to precisely formulate this. Another observation is that the stationary points are usually unstable, hence it is reasonable to believe that they might not be relevant if one adds stochastic noise to the evolution. Even if solutions do not become stationary, it can still happen that the coarsening is slower than the predicted rate, (see Chapter 1). On the other other hand, for upper bounds on the coarsening rate one can sometimes apply the robust approach of Kohn and Otto [40] which gives a time-integrated bound on a suitably chosen length-scale. In general, upper coarsening bounds seem to be more universal, although they can also be violated, depending on the choice of length scale (e.g the sequential vanishing example in [33]).

To address the even harder question of self-similarity, we have to consider the locality of the particle interaction. To even make an ansatz for a self-similar profile Φ, one usually needs an evolution equation for the size distribution $f(t,x)$. This is only possible in so called mean-field models, where every particle interacts symmetrically with each other particle. In this case the whole information of the system is already in the size distribution, since there is no underlying topology of the particle configuration. On the other hand, if the particle interaction is more local, (e.g nearest neighbor interaction, see Chapter 1), then there is no closed equation for the size distribution and no way of predicting the self-similar form beforehand. A natural question is whether the statistics of systems with local interactions behave like their mean-field counterparts for suitable initial data. The next difficulty lies in the fact that, even if one can make an ansatz for a self-similar profile from the evolution of the size distribution, several self-similar solutions might exist (see Chapter 2). In fact, this is a prominent phenomenon in the classical LSW theory of coarsening [47, 59, 50], where a one-parameter family of self-similar solutions emerge. Notably, all the profiles have compact support but only one of them is smooth. Then there is also a dense set of initial data that do not converge to any self-similar solution [53]. It is natural to conjecture that a suitably regularized evolution converges to the unique smooth self-similar solution. Introducing a regularization then can also be interpreted as adding stochastic noise to the underlying microscopic model.

In the scope of this thesis we investigate some of the aforementioned aspects of coarsening in three relatively basic models and demonstrate how the behavior of coarsening models change when considering local interactions vs. mean-field interactions and deterministic vs. stochastic dynamics. In Chapter 1 we consider a one-dimensional model as in [33] with nearest-neighbor interaction that takes the form of the discrete PDE

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}x_k &= F_\beta(x_{k-1}(t)) - 2F_\beta(x_k(t)) + F_\beta(x_{k+1}(t)), \quad k \in \mathbb{Z},\\ F_\beta(x) &= -\frac{\beta}{|\beta|}x^\beta, \quad \beta \in (-\infty,0)\cup(0,1].\end{aligned}$$

The particle interaction is governed by the flux function F_β that is chosen such that large particles tend to grow and small particles tend to shrink. If the left neighbor of a particle k dies, the next non-zero particle to the left becomes the new left neighbor (and the same holds for the right neighbor) so that a particle always has a left and a right neighbor. This is only a simple toy model for coarsening, but a lot of the difficulties and pathologies mentioned above are already present on this level. Our main contribution is the existence of certain initial data that coarsen with the predicted rate $\langle x \rangle \sim t^{\frac{1}{1-\beta}}$, and as far as we know, this is the first example of solutions in a model with local interactions that coarsen with the optimal rate. Since the argument is not fully constructive we can not characterize this class of initial data explicitly, but show that the constant stationary configuration can be approximated arbitrarily well by data from this class. Although this is a positive result, it also demonstrates the pathologies of the model. The solutions we construct coarsen in a very non-generic organized way that one would not expect to be physically relevant. As a byproduct we also show the existence of initial data that become stationary after finite time or coarsen with an arbitrarily slow rate.

In Chapter 2 we consider the mean-field version of the model from Chapter 1, which can be described by the evolution of the size distribution $f(t,x)$ that evolves according to a singular LSW equation:

$$\partial_t f = \partial_x \left((x^{-\beta} - \theta) f \right), \quad \theta = \frac{\int_0^\infty x^{-\beta} f \, \mathrm{d}x}{\int_0^\infty f \, \mathrm{d}x}.$$

We show that solutions for general measure valued initial data exist, a result that uses techniques from the classical LSW theory [52] but requires some new ideas and estimates to deal with the singular terms. Then the mean-field nature of the model also enables us to find self-similar solutions. Here we get the typical result of a one-parameter family of self-similar solutions, each with compact support but only one smooth profile. It is very likely that one can also prove the existence of initial data that do not converge to any of the self-similar solutions as in [53], which is possible future work. It is worth to mention that, as in the local case, configurations where all particles have the same size are stationary for the evolution.

We conclude the thesis with an analysis of the exchange-driven growth model (EDG) in Chapter 3. In Chapter 2 we already showed how the mean-field interaction in coarsening processes makes it possible to find self-similar solutions, however there is no universal long-time behavior of solutions because of the existence of stationary states and a continuous family of self-similar profiles. As we mentioned before, the general idea to induce more universal behavior is through stochastic interaction. Naively one could simply add some noise term to the deterministic interaction and hope this would eliminate unwanted behavior. While this might work, it is not clear how to prove such results rigorously if the deterministic case is already not very well understood. Instead we consider the coarsening dynamics in EDG which have an intrinsically stochastic element. As before, the equation for the dynamics is stated for the size distribution, however this time the particles (which we call clusters in this context) consist of "elementary particles" and have a discrete amount of mass. Hence the size distribution is a probability measure on the natural numbers $c_k(t), k \in \mathbb{N}_0$ and the evolution is given by

$$\begin{aligned} \dot{c}_k &= \sum_{l\geq 1} K(l,k-1)c_l c_{k-1} - \sum_{l\geq 1} K(k,l-1)c_k c_{l-1} \\ &\quad - \sum_{l\geq 1} K(l,k)c_l c_k + \sum_{l\geq 1} K(k+1,l-1)c_{k+1}c_{l-1} \ , \qquad \text{for } k \geq 0 \ . \end{aligned}$$

This model arises from a continuous-time Markov jump process on a complete graph, where a particle jumps from a cluster with k particles to a cluster with l particles with a rate given by the interaction kernel $K(k,l)$. As the size of the graph L and the number of particles N diverge with $N/L \to \rho \in [0,\infty)$, the k-cluster fractions converge to the above system [29]. Although the equations for c_k are deterministic, they have a more diffusive nature because of their stochastic origin. In particular, even starting with clusters of the same size the evolution will not be stationary for any reasonable choice of K, which already is a major difference to the models from the previous chapters. We then consider the case of the product kernel $K(k,l) = (kl)^\lambda$ for $\lambda \in [0,2)$ introduced in [7]. For this special case we can identify a unique self-similar profile g_λ in a suitable continuum limit and show that every solution $c_k(t)$ to the EDG equations with first moment ρ converges in a weak sense[1] to the self-similar solution with mass ρ.

In conclusion, this thesis demonstrates how fundamental features regarding the nature of interaction in coarsening models affect the general behavior of these models. Systems with local interactions are still far from understood, and for what it is worth, our contribution in this area rather emphasizes the mathematical difficulties with these models instead of solving them. Considering mean-field interaction instead of local interaction makes it at least possible to directly study the statistics of this system and find self-similar solutions with relative ease, although the general well-posedness can still be challenging and one cannot expect universal long-time behavior. At the moment, mean-field models of a more diffusive nature like exchange-driven growth (which directly comes from a stochastic microscopic model) seem like good candidates for models where it is possible to rigorously establish universal coarsening behavior. We give a first result in this direction and feel that there is a lot of potential for further research, especially including other classes of interaction kernels like zero-range interaction.

[1]To be exact, the convergence holds for a suitably rescaled empirical measure associated to a solution.

1. Special solutions to a nonlinear coarsening model with local interactions

The contents of this chapter were published in [16].

1.1. Introduction and results

1.1.1. Local coarsening models

Discrete mass transfer models with local interactions have been studied by several authors in different contexts. They have applications in physics such as the growth and coarsening of sand ripples in [32] or the clustering in granular gases [57], while also serving as approximations or toy models for more complex coarsening scenarios such as the evolution of droplets in dewetting films [27, 26] and grain growth [34]. The model that we study consists of an infinite number of mass points on a one-dimensional lattice that exchange mass with their nearest neighbours. In the symmetric case that we consider, the evolution is governed by the following system of ODEs,

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t,k) = F(x(t,k-1)) - 2F(x(t,k)) + F(x(t,k+1)), \tag{1.1}$$

where the right hand side represents the net mass flux at a site k which receives and transfers mass from its neighbours at rates controlled by the flux function F. This system can also be interpreted as the spatially discrete nonlinear PDE $\partial_t x = \Delta F(x)$. The monotonicity properties of F are crucial for the qualitative behaviour of solutions and depend on the application, as an increasing flux function will lead to mass diffusion and a decreasing flux function will lead to aggregation and coarsening. A combination of both is also possible, for example in models that were investigated in [19, 18].

In this chapter we consider the coarsening model proposed in [33], with flux function

$$F(x) = F_\beta(x) = -\frac{\beta}{|\beta|}x^\beta,$$

where $-\infty < \beta < 0$ or $0 < \beta \leq 1$ (In contrast to [33] we use the exponent β instead of $-\beta$). This largely resembles the sand ripple scenario [32], although we will refer to the lattice points as *particles* from now on. Distinctive features of the model are the infinite number of particles and the *vanishing rule: Particles that reach mass zero are*

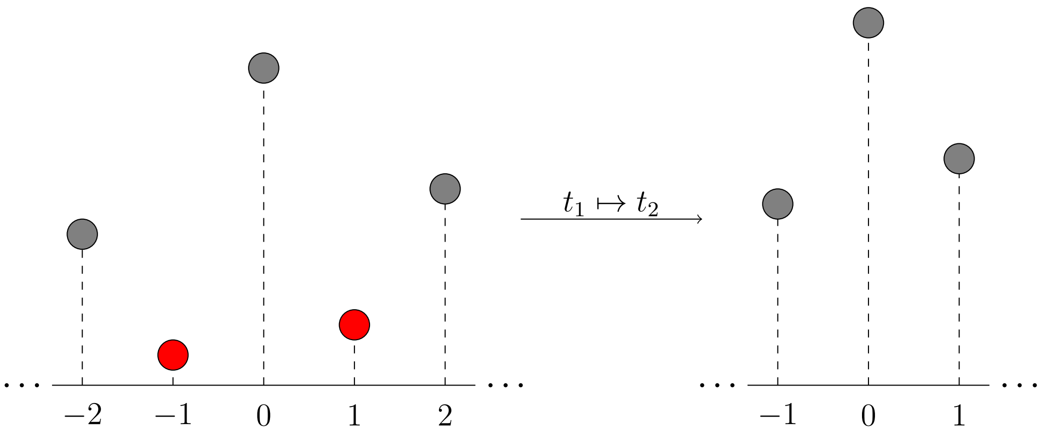

Figure 1.1.: Small particles vanish at t_2 and the average mass increases.

deleted from the system and the remaining particles are relabeled accordingly. This way small particles vanish from the system while transferring their mass to the rest of the system, which leads to a growth of the average particle size and an overall coarsening of the system.

With this particular choice of the flux function (except when $\beta = 1$), the equation has an invariant scaling: If $x = x(t, k)$ is a solution, then

$$x_\lambda(t, k) = \lambda^{\frac{1}{\beta - 1}} x(\lambda t, k)$$

is another solution. Thus, if $\langle x \rangle$ denotes a suitable lenght-scale, we expect that

$$\langle x \rangle \sim t^{\frac{1}{1-\beta}}. \tag{1.2}$$

In the case $\beta = 1$ the mean-field analysis in [32] indicates that $\langle x \rangle \sim \exp(\lambda t)$, where λ is not universal but depends on the initial distribution.

The problem in the mathematical analysis of such models is to rigorously establish such coarsening rates. The method of Kohn and Otto [40] has proved very useful in several situations to obtain (weak) upper bounds. In this context, $\langle x \rangle$ is usually some negative Sobolev norm and the dynamics have a gradient flow structure. In [19, 18], the method is successfully applied to discrete forward-backward diffusion equations similar to our setting, where the system is viewed as H^{-1} gradient flow with respect to the energy

$$E(x) = \sum_k \Phi(x_k),$$

where $\Phi' = F$. Unfortunately we can not adapt the method for two reasons. Firstly, the energy is not finite due to the presence of infinitely many particles. Secondly, the vanishing and relabeling of particles is not compatible with the gradient flow structure. It is also interesting to compare the coarsening exponents of our model and [19, 18]. In [18],

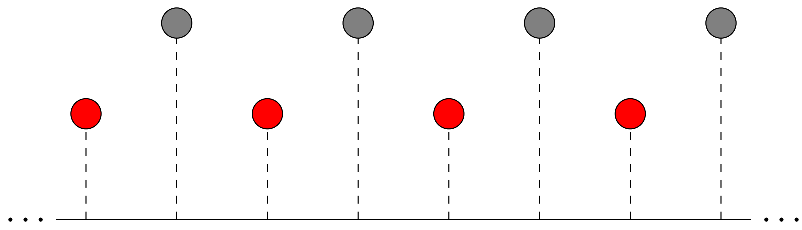

Figure 1.2.: After the smaller particles have vanished, the configuration is constant.

the generic coarsening rate is estimated as $\langle x\rangle \lesssim t^{\frac{1}{3}}$, while [19] yields the refined estimate $\langle x\rangle \lesssim t^{\frac{1}{3-\alpha}}$ under the assumption that $\Phi(x) \sim x^{\alpha}$, for $x >> 1$, $\alpha \in [0,1)$. The difference in exponents to our model looks contradictory, but there is also a key difference between the models: In [19, 18], the equation is assumed to be parabolic at 0 and particles do not vanish. Instead, an increasing number of small particles forms a bulk. The main coarsening mechanism then is diffusion between large particles across the bulk. In our model on the other hand there is no generic bulk because small particles vanish from the system and the particular coarsening exponent is a consequence of scaling.

Regarding upper coarsening bounds, the simple structure of our model enables us to apply more elementary arguments: For positive β, the right hand side of equation (1.1) can be estimated to obtain

$$\dot{x} \leq 2x^{\beta},$$

which can be integrated to yield the desired bound in the ℓ^{∞}-norm. For negative β the equation gives $\dot{x} \geq -2x^{\beta}$, which can be used to derive a weak upper bound, see Proposition 2.4 in [33]. Furthermore, the numerical simulations and heuristics in [33] demonstrate that single particles can grow linearly (thus faster than the scaling law) in time, showing that an ℓ^{∞}-bound cannot be expected in this case.

On the other hand, not much is known about the validity of lower bounds. As will be demonstrated below, there are many non-constant initial configurations which become stationary after a finite time due to the vanishing of particles. An easy example for this is a 2-periodic configuration of large and small particles. During the evolution, the large particles grow and the small particles shrink until disappearing at the same time, at which all large particles will be left with the same size and the evolution stops (see Figure 1.2).

The problem of classifying all initial data for which some form of a lower coarsening bound holds is completely open. The main result of this chapter is the existence of initial data and corresponding solutions with scale-characteristic coarsening rates, where $\langle x\rangle$ is a suitable average of the configuration, see Theorem 1.2. Our general ansatz is to reverse time, which transforms the equation into a nonlinear discrete parabolic equation which behaves much better and can be analysed by means of Harnack-type positivity estimates (see [9]) and parabolic regularity theory (see [48] for the continuum theory and [25] for the discrete analogue). It should be mentioned that the solutions that we construct coarsen in a very organised manner, whereas numerical simulations and heuristics that were done in [33] indicate that the generic coarsening behaviour is more disorganised. Nevertheless we believe that this result is valuable because our solutions are, as far as we know, the first rigorous examples of indefinite coarsening in a model with local interactions. The fact that the corresponding initial data can be arbitrarily close to constant data (see Corollary 1.4) shows that they are at least relevant on a topological level. On the other hand, the existence of such non-generic solutions yields obstructions when trying to quantify generic behaviour. As already mentioned in [33], a clarification of the notion of disorder seems necessary.

1.1.2. Setup and notation

We consider a discrete infinite number of particles with non-negative mass on a one-dimensional lattice. That means each configuration is an element of the space

$$\ell^\infty_+(\mathbb{Z}) = \{x = x(k) \in \ell^\infty(\mathbb{Z}) : \ x(k) \geq 0\}\,. \tag{1.3}$$

As described above, particles with zero mass will be deleted from the system during the evolution. However, relabeling the particle indices whenever a particle vanishes can be problematic. On the one hand, relabeling can be ambiguous, for example the vanishing times might not be in order or could have accumulation points. On the other hand the solution will be discontinuous in time. Thus it is more convenient to leave the configuration unchanged and update the interaction term on the right-hand side of equation (1.1) instead. For this purpose we define the *nearest living neighbour indices*

$$\sigma_+(x,k) = \inf\{l > k : x(l) > 0\},$$
$$\sigma_-(x,k) = \sup\{l < k : x(l) > 0\},$$

where we just write $\sigma_\pm(k)$ if there is no danger of confusion. Also we define the ordinary discrete Laplacian Δ and the *living particles Laplacian* Δ_σ as

$$\Delta x = x(k-1) - 2x(k) + x(k+1),$$
$$\Delta_\sigma x(k) = (x(\sigma_-(k)) - 2x(k) + x(\sigma_+(k))) \cdot \mathbf{1}_{\{x(k)\neq 0\}}.$$

Then the evolution of the system is governed by the following equation:

$$\begin{cases} \partial_t x = \Delta_\sigma F_\beta(x) & \text{in } (0,\infty)\times\mathbb{Z}, \\ x(0,\cdot) = x_0, \end{cases} \tag{1.4}$$

with $x_0 \in \ell^\infty_+$ and

$$F_\beta(x) = -\frac{\beta}{|\beta|}x^\beta, \ \ \beta \in (-\infty, 1] \setminus \{0\},$$

with $F_\beta(0) := 0$ for $\beta < 0$. The only drawback is that the right-hand side of (1.4) is no longer continuous, hence we have to use a concept of mild solutions, as in [33].

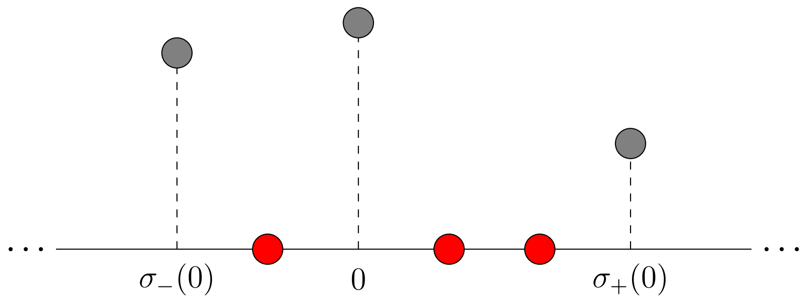

Figure 1.3.: Vanished particles remain in the physical domain, only neighbour relations σ_+, σ_- change.

Definition 1.1. *Let $0 < T \leq \infty$. We say that $x : [0, T) \to \ell^\infty_+(\mathbb{Z})$ is a solution to problem (1.4) if the following conditions are satisfied:*

1. *$t \mapsto x(t, k)$ is continuous on $[0, T)$ and $x(0, k) = x_0(k)$ for every $k \in \mathbb{Z}$.*
2. *$t \mapsto F_\beta(x(t, k))$ is locally integrable on $[0, T)$ for every $k \in \mathbb{Z}$.*
3. *For every $0 \leq t_1 < t_2 < T$ and every $k \in \mathbb{Z}$ we have*
$$x(t_1, k) - x(t_2, k) = \int_{t_1}^{t_2} \Delta_\sigma F_\beta(x)(s, k) \, \mathrm{d}s.$$

The second condition is automatically satisfied if β is positive. For the existence of solutions we refer to [33], where the case $\beta < 0$ is discussed. We expect a similar result to hold for positive β but since we are only concerned with special solutions anyway we will give no proof here. More important for our result is the well-posedness of the time-reversed evolution

$$\partial_t u = \Delta G_\beta(u),$$

with $G_\beta(u) = -F_\beta(u)$, which is the discrete analogue of a fast diffusion equation. This is addressed in Section 1.4, see Theorem 1.18.

It is easy to check that the evolution (1.4) conserves the average mass

$$\langle x \rangle = \lim_{N \to \infty} \frac{1}{2N + 1} \sum_{k=-N}^{N} x(k).$$

This is not really meaningful, since vanished and living particles are treated the same. To adequately measure the coarsening process, one has to average only over the living particles. Consequently we define

$$L_{\sigma_+, N} = \bigcup_{k=1}^{N} \{\sigma_+^{(k)}(0)\},$$
$$L_{\sigma_-, N} = \bigcup_{k=1}^{N} \{\sigma_-^{(k)}(0)\},$$

as sets of the first N positive, respectively negative living particle particle indices ($\sigma_+^{(k)}$ denoting the k-fold composition) and set

$$L_{\sigma,N} = \begin{cases} L_{\sigma_+,N} \cup L_{\sigma_-,N}, \text{ if } x(0) = 0, \\ L_{\sigma_+,N} \cup L_{\sigma_-,N} \cup \{0\}, \text{ if } x(0) > 0. \end{cases}$$

Then we can define the *living particle means*

$$\begin{aligned} \langle x \rangle_{\sigma,N} &= \frac{1}{|L_{\sigma,N}|} \sum_{k \in L_{\sigma,N}} x(k), \\ \langle x \rangle_\sigma^+ &= \limsup_{N\to\infty} \langle x \rangle_{\sigma,N}, \\ \langle x \rangle_\sigma^- &= \liminf_{N\to\infty} \langle x \rangle_{\sigma,N}. \end{aligned}$$

Since mass is transferred from small to large particles and the small particles eventually vanish, we expect the living particle means to grow in time.

1.1.3. Main result

In the main result of the chapter we show that there exist solutions where the average particle size grows with the characteristic rate that is indicated by scaling:

Theorem 1.2. *Let $\beta \in (-\infty, 0) \cup (0, 1]$ and F_β be defined as above. Then the following statements hold:*

1. *For every $\beta \in (-\infty, 0) \cup (0, 1)$ there exists $x_0 \in \ell_+^\infty(\mathbb{Z})$ and a solution to equation (1.4) with initial data x_0 that satisfies*

$$\begin{aligned} \langle x \rangle_\sigma^- &\gtrsim t^{\frac{1}{1-\beta}}, \\ ||x||_\infty &\lesssim t^{\frac{1}{1-\beta}}. \end{aligned}$$

2. *For $\beta = 1$ there exists $0 < \lambda \leq 2$, $x_0 \in \ell_+^\infty(\mathbb{Z})$ and a solution to equation (1.4) with initial data x_0 that satisfies*

$$\begin{aligned} \langle x \rangle_\sigma^- &\gtrsim \exp(\lambda t), \\ ||x||_\infty &\lesssim \exp(\lambda t). \end{aligned}$$

Here, $\gtrsim$ and $\lesssim$ mean that the corresponding inequalities hold up to a multiplicative constant that depends only on β.

We briefly describe the strategy of the proof. The key observation is that the time-reversed system corresponding to equation (1.4) is a nonlinear parabolic equation where particles are inserted instead of vanishing, which is much easier to handle. Thus the idea is to make a more or less explicit construction in the time-reversed setting and then reverse time again to obtain a sequence of approximate solutions $x^{(n)}$ which solve (1.4) and eventually converge to a solution with the desired properties. Each solution $x^{(n)}$ will be constructed in n steps, starting in the future time T_n (with $T_n \to \infty$), where the particle sizes are of order θ^n for some $\theta > 1$. We then insert particles to lower the average particle size to order θ^{n-1} and run the time-reversed evolution, equilibrating the system until all particle sizes are of order θ^{n-1}. The procedure is then iterated, going from sizes of order θ^{n+1-j} to θ^{n-j}, until after n steps all particles sizes are of order one (see Figure 1.4). A suitable compactness argument for $n \to \infty$ then yields a solution x on $[0, \infty)$ to equation (1.4).

In order to achieve the desired coarsening rate the time-span to equilibrate in the j-th step has to be of order $\theta^{(1-\beta)(n+1-j)}$, which is a-priori not clear. Due to scaling however, every step is equivalent to the problem of inserting particles into a configuration u_0 of order one (denoted by $u_0 \mapsto \Psi_* u_0$) such that after evolving the system under the backward equation for a uniform timespan T the particles are of order θ^{-1}. More precisely, we will prove the following result, which is the heart of the argument:

Lemma 1.3 (Key Lemma)**.** *Let $\beta \in (-\infty, 0) \cup (0, 1]$ and $G_\beta = -F_\beta$. Then for every $\varepsilon > 0$ there exists $T = T(\beta, \varepsilon) > 0$, such that the following holds: For every $u_0 \in \ell^\infty_+(\mathbb{Z})$ with $\frac{1}{2} \leq u_0 \leq 1$ there exists a creation operator Ψ_* and a solution u of the equation*

$$\begin{cases} \partial_t u = \Delta G_\beta(u) & \text{in } (0, \infty) \times \mathbb{Z}, \\ u(0, \cdot) = \Psi_* u_0, \end{cases} \tag{1.5}$$

that satisfies

$$\left| u(T, \cdot) - \frac{1}{2} \right| \leq \varepsilon.$$

The precise meaning of $\Psi_* u_0$ will be explained in the next section. In particular, if we set $\theta^{-1} = 1/2 + \varepsilon$ with $\varepsilon \leq 1/6$ then u will satisfy the desired estimate

$$\frac{1}{2}\theta^{-1} \leq u(T, \cdot) \leq \theta^{-1}.$$

The main idea to prove the lemma is to insert particles such that $1/2 - \varepsilon \leq \Psi_* u_0 \leq 1/2 + \varepsilon$ holds in an averaged sense. Since the backward equation is a diffusion, it is expected that the system equilibrates and average-wise estimates eventually induce point-wise estimates after a certain timespan, see Lemma 1.13. Note that due to the freedom of choice in the parameter ε, the back-in-time construction can generate initial data that are arbitrarily flat, demonstrating the instability of constant data:

Corollary 1.4. *Let $c > 0$. Then for every $\varepsilon > 0$ there exist initial data x_0 as in Theorem 1.2 such that*

$$||x_0 - c||_\infty \leq \varepsilon.$$

Before proving these results we introduce the formalism Ψ_* for the insertion of particles and explain the general construction of solutions to the coarsening equation.

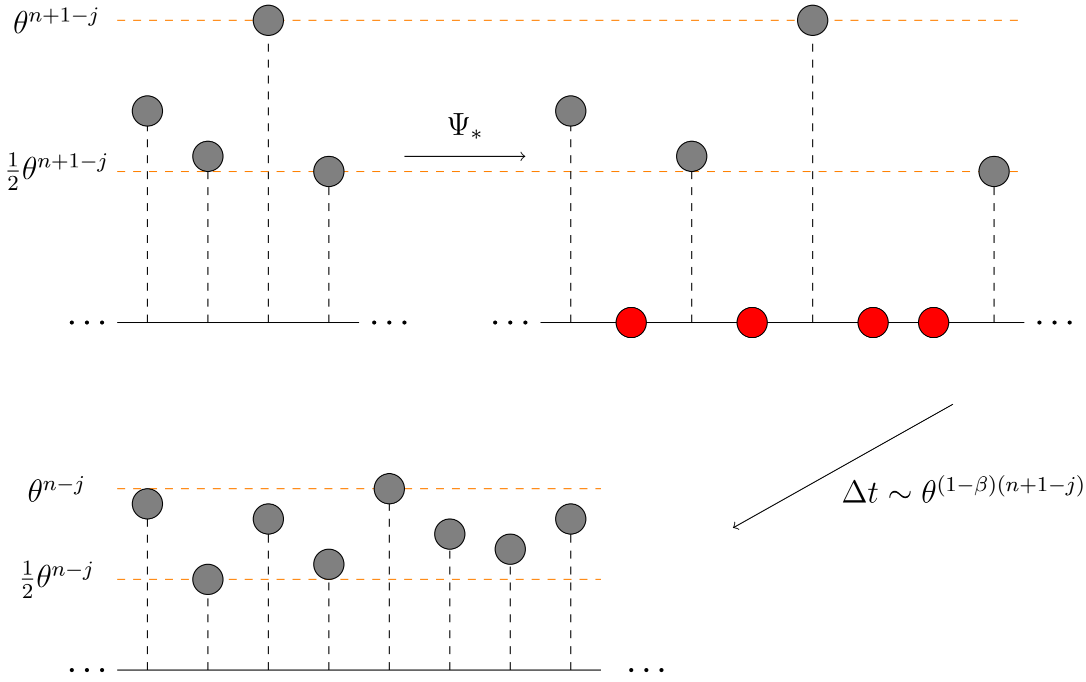

Figure 1.4.: The j-th step in the back-in-time construction.

Outline

The rest of the chapter is organised as follows. In Section 1.2 we describe how to construct (local-in-time) solutions to equation (1.4). The general idea is to choose some terminal data $x(T, \cdot)$ and go backward in time from there. The crucial observation is that the vanishing of particles corresponds to the *creation* of particles if time is reversed. Additionally, since the living particles do not carry any information of the vanished particles in the forward-in-time equation, new particles can be created at arbitrary times τ_j and positions $\{\Psi_*^{(j)}\}$ (for notation see next section) in the backward equation. This gives the necessary freedom to construct solutions with desirable properties. Hence, each data triple $(x(T, \cdot), \{\tau_j\}, \{\Psi_*^{(j)}\})$ gives rise to a solution of equation (1.4) on the interval $[0, T]$.

In Section 1.3 we use the local solutions from Section 1.2 to carry out the proof of Theorem 1.2 using the strategy that we outlined above. First we show that we can modify the local average of any suitable initial configuration by inserting particles into the configuration (see Lemma 1.10), which is an elementary argument. Then we analyze equation (1.5) by rewriting it as a linear discrete evolution equation in divergence form. By a suitable positivity estimate (see Lemma 1.11) we show that the equation is uniformly parabolic away from $t = 0$, which enables the use of discrete Nash-Aronson regularity estimates (see Theorem 1.30) to prove the desired equilibrium property (see Lemma 1.13) for the backward equation. The rest of the proof is a relatively straight forward compactness argument.

In Section 1.4 we carry out the proofs of some necessary technical results such as existence of solutions for the time-reversed setting, Harnack inequalities and application of the discrete Nash-Aronson estimates to our setting.

1.2. Construction of solutions

1.2.1. Insertion of particles

First we fix the notation for the insertion of particles. Basically, we need a precise way to insert zeroes into a given sequence of numbers. The most practical way to do this is via push-forward of a suitable increasing map $\Psi : \mathbb{Z} \to \mathbb{Z}$. This map can be defined by the corresponding sequence of "jumps". We make the following definition:

Definition 1.5. *Let $d : \mathbb{Z} \to \mathbb{N}_0$ be a given sequence of jumps. Then define the corresponding deformation as*

$$\Psi : \mathbb{Z} \to \mathbb{Z}$$

$$\Psi(k) = k + \sum_{m=0}^{k} d(m).$$

Now for every $x \in \ell_+^\infty(\mathbb{Z})$ we define the push-forward sequence $\Psi_* x$ as

$$\Psi_* x(\Psi(k)) = x(k),$$

for all $k \in \mathbb{Z}$ and $\Psi_* x(l) = 0$ if $l \notin \operatorname{Im}(\Psi)$.

With this definition, if x represents a particle configuration, then $\Psi_* x$ represents the same configuration with new mass-zero particles created. To be more precise, the condition $d(k) = l$ exactly means that we are inserting l new particles between the k-th and the $(k \pm 1)$-th particle, (depending on the sign of k). We will refer to the mapping Ψ_* as *particle creation operator* and, to keep notation as compact as possible, not explicitly refer to the deformation Ψ or the specific jump sequence d any more, but rather just state where particles are inserted. This is potentially ambiguous, for instance, "creating a particle between each two living particles" can be achieved by different d, potentially translating the original living particles. However, in the following sections these ambiguities do not affect the arguments, hence we will ignore them.

1.2.2. Back-in-time construction

Next we describe how to obtain a solution from a given terminal configuration x_{ter}, an increasing sequence of vanishing/creation times $\{\tau_j\}_{j=1,..,n}$ and corresponding creation operators $\{\Psi_*^{(j)}\}_{j=1,..,n}$. We define the solution piecewise by iteratively using the backward evolution (1.5) on $[\tau_{j-1}, \tau_j]$ after inserting particles at $t = \tau_{j-1}$ and continuing the procedure. To be precise, we define $u^{(j)}$ on the interval $[\tau_{j-1}, \tau_j]$ to be a solution of the following problem:

$$\begin{cases} \partial_t u^{(j)} = \Delta G_\beta(u^{(j)}) \text{ in } (\tau_{j-1}, \tau_j] \times \mathbb{Z}, \\ u^{(j)}(\tau_{j-1}) = \Psi_*^{(j)} \left[u^{(j-1)}(\tau_{j-1}) \right], \end{cases} \tag{1.6}$$

for $j = 1, .., n$, with $\tau_0 := 0$, $u^{(0)}(\tau_0) := x_{\text{ter}}$ and $G_\beta = -F_\beta$. We should note that by a solution we mean a classical solution, i.e $u^{(j)} \in C^0([\tau_{j-1}, \infty), \ell_+^\infty(\mathbb{Z}))$, for every $k \in \mathbb{Z}$ we have $u^{(j)}(\cdot, k) \in C^1((\tau_{j-1}, \infty))$ and the equation holds pointwise. Well-posedness of this problem is a-priori not clear, especially for the case $\beta < 0$. For the moment we just assume that the equation is solvable and focus on carrying out the construction of solutions to the coarsening equation. In Theorem 1.18 we give a sufficient condition on the initial data for existence of solutions that is easily verified for the data considered in the next section.

Reversing the time direction we obtain piecewise solutions of our original equation. However, one has to compose $u^{(j)}$ with the creation operators once more, since vanished particles remain in the "physical" domain in the original evolution (1.4). To be more precise, we set

$$x^{(j)}(t) = \left(\prod_{l=1}^{j-1} \Psi_*^{(n+1-l)} \right) \left[u^{(n+1-j)}(\tau_n - t) \right],$$

which lets us glue the solutions together in a continuous way:

$$x(t) = x^{(j)}(t), \text{ if } t \in [\tau_n - \tau_{n+1-j}, \tau_n - \tau_{n-j}),$$

for $j = 1, .., n$. Using $u^{(j)}(\tau_{j-1}) = \Psi_*^{(j)} \left[u^{(j-1)}(\tau_{j-1}) \right]$ it is easy to check that x defined this way is continuous in time. The next lemma shows that x is indeed a solution to our original equation:

Lemma 1.6. *Let Ψ_* be a creation operator as above. Then we have*

1. $\sigma_\pm(\Psi_* x, \Psi(k)) = \Psi(\sigma_\pm(x, k))$ *for every* $x \in \ell_+^\infty$ *and* $k \in \mathbb{Z}$.

2. $[\Delta_\sigma, \Psi_*]x = (\Delta_\sigma \Psi_* - \Psi_* \Delta_\sigma)\, x = 0$ *for every* $x \in \ell_+^\infty$.

3. $\langle \Psi_* x \rangle_{\sigma,N} = \langle x \rangle_{\sigma,N}$ *for every* $N > 0$ *and* $x \in \ell_+^\infty$.

Proof. 1. It suffices to prove the claim for σ_+, the other case is completely analogous. Because Ψ is strictly increasing, we have $\Psi(\sigma_+(x, k)) > \Psi(k)$. We also have

$$\Psi_* x(\Psi(\sigma_+(x, k))) = x(\sigma_+(x, k)) > 0,$$

which shows $\sigma_+(\Psi_* x, \Psi(k)) \le \Psi(\sigma_+(x,k))$. For the other inequality, we note that $\Psi_* x(l) > 0$ implies that $l = \Psi(m)$ for some $m \in \mathbb{Z}$. In this case we have

$$0 < \Psi_* x(l) = x(m)$$

which implies $m \ge \sigma_+(x,k)$, and because Ψ is increasing we conclude

$$l = \Psi(m) \ge \Psi(\sigma_+(x,k)),$$

which proves the first assertion.

2. Let $l = \Psi(k)$. We apply the identity in *1.* to get

$$\begin{aligned}\Delta_\sigma \Psi_* x(l) &= (\Psi_* x(\sigma_-(\Psi_* x, l)) - 2\Psi_* x(l) + \Psi_* x(\sigma_+(\Psi_* x, l))) \cdot \mathbf{1}_{\{\Psi_* x(l) \neq 0\}} \\ &= (x(\sigma_-(k)) - 2x(k) + x(\sigma_+(k))) \cdot \mathbf{1}_{\{x(k) \neq 0\}} \\ &= \Delta_\sigma x(k) = \Psi_* \Delta_\sigma x(l).\end{aligned}$$

On the other hand, if $l \notin \mathrm{Im}(\Psi)$, the identity is trivial.

3. Obvious from the definition. □

With the second statement of the above lemma, it is not difficult to verify that the sequence x we have constructed above solves equation (1.4):

Corollary 1.7. *Let x_{ter}, $\{\tau_j\}$ and $\{\Psi_*^{(j)}\}$ be given and x be constructed as above. If $t \mapsto F_\beta(x(t,k))$ is locally integrable for every $k \in \mathbb{Z}$, then x is a (mild) solution to equation* (1.4) *on $[0, \tau_n)$.*

Proof. Since x is continuous and piecewise smooth by construction, it suffices to show that $\partial_t x = \Delta_\sigma F_\beta(x)$ holds pointwise on all intervals $[\tau_n - \tau_{n+1-j}, \tau_n - \tau_{n-j})$. Indeed, we calculate

$$\begin{aligned}\partial_t x^{(j)}(t) &= \left(\prod_{l=1}^{j-1} \Psi_*^{(n+1-l)}\right) \left[\partial_t u^{(n+1-j)}(\tau_n - t)\right] \\ &= \left(\prod_{l=1}^{j-1} \Psi_*^{(n+1-l)}\right) \left[\Delta_\sigma F_\beta(u^{(n+1-j)})(\tau_n - t)\right] \\ &= \Delta_\sigma F_\beta \left(\left(\prod_{l=1}^{j-1} \Psi_*^{(n+1-l)}\right) \left[u^{(n+1-j)}(\tau_n - t)\right] \right) = \Delta_\sigma F_\beta \left(x^{(j)}\right).\end{aligned}$$

Here we used that Δ_σ commutes with creation operators by the previous lemma, as well as composition with the function F_β. □

Remark 1.8. *The above construction scheme implies the existence of many initial data and corresponding solutions to the coarsening equation which become stationary after a finite time. Indeed, x as above has this property if we pick x_{ter} to be a constant sequence.*

Because there is much freedom in the choice of particle creations and vanishing times this means that finding conditions on initial data such that lower coarsening bounds hold is a difficult task and remains an open problem. In the construction for the proof of Theorem 1.2 we will in fact choose $x_{\text{ter}}(k) = \theta^n$ *so that each approximate solution becomes stationary. Because* $\theta^n \to \infty$ *and* $\tau_n \to \infty$ *the limit solution however will grow indefinitely. The details will be explained in the next section.*

1.3. Proof of Theorem 1.2

We divide the full proof of Theorem 1.2 into four main steps. In the first step we show how to insert particles to modify the local average in a uniform way. The second step is to prove a long-time diffusive property of the backward equation which, together with the first step, will enable us to prove Lemma 1.3. In the third step the construction of the approximate sequence $x^{(n)}$ is thoroughly carried out. Finally we use a compactness argument to pass to the limit and obtain a solution with the desired properties, finishing the proof.

1.3.1. Step 1: Average modification by particle insertion

Definition 1.9 (Local Averages)**.** *Let* $x \in \ell^\infty(\mathbb{Z})$*. Then define the associated sequence of local averages as*

$$\Lambda(x, k, N) = \frac{1}{2N+1} \sum_{l=-N}^{N} x(k-l).$$

In the next lemma we show how to modify the local averages of a given sequence by inserting particles in a suitable way:

Lemma 1.10 (Particle insertion)**.** *Let* $u_0 \in \ell^\infty_+(\mathbb{Z})$ *with*

$$\frac{1}{2} \leq u_0 \leq 1.$$

Then for every $\varepsilon > 0$ *there exists a creation operator* Ψ_* *and* $N_0 \in \mathbb{N}$ *such that*

$$\left| \Lambda(\Psi_* u_0, \cdot, N) - \frac{1}{2} \right| \leq \varepsilon,$$

for $N \geq N_0$*. Furthermore, if* d *is the jump sequence associated to* Ψ_**, then* $||d||_\infty$ *is finite and depends only on* ε*.*

Proof. Let (λ_i) be an equidistant partition of the interval $[1/2, 1]$ with $|\lambda_i - \lambda_{i+1}| \leq \varepsilon$. We give an explicit scheme for the particle insertion as follows: We divide $\mathbb{Z}$ into disjoint blocks of particles with length K, where K is determined later:

$$\mathbb{Z} = \bigcup_{j \in \mathbb{Z}} B_j,$$

with $B_j = \{jK, ..., (j+1)K-1\}$. Let Λ_j denote the average mass in B_j with respect to u_0. We define the deformation Ψ by inserting L_i (determined later) particles to the right of $(j+1)K-1$ whenever he have

$$\lambda_i \leq \Lambda_j \leq \lambda_{i+1}.$$

This gives rise to a new partition of $\mathbb{Z}$ into blocks $\tilde{B}_j$ with varying lenghts $K+L_i$, where $\tilde{B}_j$ contains all elements of $\Psi(B_j)$ and the next L_i numbers that are not elements of $\mathrm{Im}(\Psi)$. We call a block with L_i inserted particles a *block of the i-th kind*. Then the average mass $\tilde{\Lambda}_j$ of such a block with respect to $\Psi_* u_0$ is by construction

$$\tilde{\Lambda}_j = \frac{1}{K+L_i}\left(\sum_{k\in B_j} u_0(k)\right) = \frac{K}{K+L_i}\Lambda_j,$$

which gives

$$\lambda_i \frac{K}{K+L_i} \leq \tilde{\Lambda}_j \leq \lambda_{i+1}\frac{K}{K+L_i} := \lambda_{i+1}\theta_i.$$

Because $1/2 \leq \lambda_i \leq 1$ we can, if K is large enough, choose $L_i \leq K$ such that

$$\left|\lambda_i\theta_i - \frac{1}{2}\right| \leq \mathcal{O}(\varepsilon),$$

and because λ_i and λ_{i+1} are close we also have

$$\left|\lambda_{i+1}\theta_i - \frac{1}{2}\right| \leq \mathcal{O}(\varepsilon).$$

This implies that the average mass of every block $\tilde{B}_j$ can be estimated as

$$\left|\tilde{\Lambda}_j - \frac{1}{2}\right| \leq \mathcal{O}(\varepsilon).$$

Next we calculate $\Lambda(\Psi_* u_0, k, N)$ for $N \gg K$ and arbitrary $k \in \mathbb{Z}$. Denote by n_i the number of blocks of the i-th kind in the domain of summation, that is $\{k-N, ..., k+N\}$. This implies that

$$|\{k-N, ..., k+N\}| = 2N+1 = \sum_i (K+L_i)n_i + \mathcal{O}(K).$$

Then we divide the summation in $\Lambda(\Psi_* u_0, k, N)$ into summation over the respective blocks and the rest of the particles in $\{k-N, ..., k+N\}$, whose number, and thus total mass R, is of order K. Thus we have

$$\begin{aligned}\Lambda(\Psi_* u_0, k, N) &= \frac{1}{2N+1}\left(\sum_{\text{sum over blocks}} + R\right)\\ &= (1+\mathcal{O}(K/N))\frac{\sum_{\text{sum over blocks}}}{\sum_i (K+L_i)n_i} + \mathcal{O}(K/N).\end{aligned}$$

By the estimates on the average masses of the blocks we have

$$\left(\frac{1}{2}-\varepsilon\right)\sum_i (K+L_i)n_i \le \sum_{\text{sum over blocks}} \le \left(\frac{1}{2}+\varepsilon\right)\sum_i (K+L_i)n_i,$$

which implies the desired estimate if $K/N \le \mathcal{O}(\varepsilon)$. Because $L_i \le K$ by construction the jump sequence satisfies $d \le K$ and K depends only on ε. □

1.3.2. Step 2: Estimate for the backward equation

The basic idea to analyse equation (1.5) is to view it as a discrete parabolic equation in divergence form with time-dependent coefficients. More precisely, with the finite difference operators

$$\partial^+ u(k) = u(k+1) - u(k),$$
$$\partial^- u(k) = u(k) - u(k-1),$$

we calculate

$$\partial_t u = \Delta G_\beta(u) = \partial^-\partial^+(G_\beta(u)) = \partial^-(a\partial^+ u),$$

where

$$a(t,k) = a_\beta(t,k) = \frac{G_\beta(u(t,k+1)) - G_\beta(u(t,k))}{u(t,k+1)-u(t,k)}.$$

The coefficient a is strictly positive and bounded from below if u is bounded from above but becomes singular at $u=0$, except for $\beta = 1$, where $a = 1$. Because of this we want to work with solutions that are bounded from above and below:

Lemma 1.11. *(Positivity estimate) Let $\beta \in (-\infty, 0) \cup (0,1)$ and $u_0 \in \ell^\infty_+(\mathbb{Z})$ with $\frac{1}{2} \le u_0 \le 1$. Let Ψ_* be a creation operator with associated jump sequence d that satisfies $d(k) \le L$. Then there exists a (classical) solution to equation (1.5) with initial data $\Psi_* u_0$. Furthermore, we have $u \le 1$ and*

$$u(t,\cdot) \ge c\left(1 \wedge t^{\frac{1}{1-\beta}}\right), \tag{1.7}$$

where c depends only on β and L.

Proof. Because the jump sequence satisfies $d \le L$, the distance between particles that have mass at least $1/2$ is at most $L+1$. Then the result follows directly from Theorem 1.18, since the above considerations imply $\Psi_* u_0 \in \mathcal{P}_{L+1,\frac{1}{2}}$. □

The lemma implies that there exists a solution u such that equation (1.5) is immediately strictly parabolic. Before we turn to the analysis of linear parabolic equations we establish uniform Hölder continuity. This is important for the stability of local averages for small times and later for the compactness of the approximating sequence.

Lemma 1.12. *(Uniform Hölder continuity) Let u_0, β, Ψ_* and u be as above. Then the following statements hold:*

1. *For $\beta < 0$ and $T > 0$ there exists $C = C(\beta, L, T)$ such that*
$$|u(t_2,k) - u(t_1,k)| \leq C|t_2 - t_1|^{\frac{1}{1-\beta}},$$
for all $0 < t \leq T$ and $k \in \mathbb{Z}$.

2. *For $\beta \in (0,1]$ we have*
$$|u(t_2,k) - u(t_1,k)| \leq 4|t_2 - t_1|,$$
for all $t > 0$ and $k \in \mathbb{Z}$.

Proof. First we note that due to equation (1.5) we have for $t_2 > t_1$:
$$|u(t_2,k) - u(t_1,k)| \leq \int_{t_1}^{t_2} |\Delta G_\beta(u)(s,k)| \,\mathrm{d}s.$$
The estimate $u \leq 1$ implies $|\Delta G_\beta(u)| \leq 4$ if $\beta \in (0,1]$. For negative β we use the lower bound (1.7) to get $|\Delta G_\beta(u)(s,k)| \leq C(\beta, L, T)s^{\frac{\beta}{1-\beta}}$ on each compact interval $[0,T]$. Then the desired inequality follows by using the estimates on $\Delta G_\beta(u)$ in the above identity and the elementary inequality $a^{\frac{1}{1-\beta}} - b^{\frac{1}{1-\beta}} \leq (a-b)^{\frac{1}{1-\beta}}$ for $a \geq b$ and $\beta < 0$. □

The next step is the long-time diffusivity result for linear equations, making use of discrete parabolic Hölder regularity (see Section 1.4.4 for details).

Lemma 1.13. *(Long-time estimate) Let $a : [0,\infty) \to \ell^\infty_+(\mathbb{Z})$ with $0 < \lambda_1 \leq a \leq \lambda_2$ and $a(\cdot,k) \in C^0([0,\infty))$ for every $k \in \mathbb{Z}$. Let $u_0 \in \ell^\infty_+(\mathbb{Z})$ and assume that there exist positive constants c_1, c_2 such that $c_1 \leq \Lambda(u_0, \cdot, N) \leq c_2$ for $N \geq N_0$. Let u be the solution of*
$$\begin{cases} \partial_t u = \partial^-(a\partial^+ u) = \mathcal{L}(t)u \\ u(0,\cdot) = u_0. \end{cases} \tag{1.8}$$
Then for every $\varepsilon > 0$ there exists $T = T(\varepsilon) > 0$, depending only on N_0 and the bounds on a, such that
$$c_1 - \varepsilon \leq u(t,k) \leq c_2 + \varepsilon$$
for all $t \geq T$, $k \in \mathbb{Z}$.

Proof. Since spatial translation does not change the type of equation and the bounds on a, it suffices to estimate $u(t,0)$. Let $\phi = \phi(t,k,s,l)$ denote the full fundamental solution to equation (1.8). Then we have
$$u(t,0) = \sum_l \phi(t,0,0,l)u_0(l) = \int_{\mathbb{R}} U(t,\xi)u_0(\lfloor t^{\frac{1}{2}}\xi \rfloor)\,\mathrm{d}\xi,$$

with

$$U(t,\xi) = t^{\frac{1}{2}}\phi(t,0,0,\lfloor t^{\frac{1}{2}}\xi\rfloor).$$

Let $\varepsilon > 0$. By Corollary 1.34 there exist $T, \delta > 0$ depending only on ε and the bounds on a such that $U(t,\cdot)$ can be approximated in L^1 by step-functions (which can be chosen to be positive since ϕ is positive) of step-width δ up to an error of ε for $t \geq T$. Let $\chi = \sum_k a_k \mathbf{1}_{I_k}$ be such a step-function, then we calculate

$$\begin{aligned}\int_{\mathbb{R}} \chi(\xi) u_0(\lfloor t^{\frac{1}{2}}\xi\rfloor)\,\mathrm{d}\xi &= \sum_k a_k \int_{I_k} u_0(\lfloor t^{\frac{1}{2}}\xi\rfloor)\,\mathrm{d}\xi \\ &= \sum_k |I_k| a_k \frac{1}{t^{\frac{1}{2}}|I_k|} \int_{t^{\frac{1}{2}} I_k} u_0(\lfloor \xi\rfloor)\,\mathrm{d}\xi.\end{aligned}$$

Since $c_1 \leq \Lambda(u_0,\cdot,N) \leq c_2$ for $N \geq N_0$, we have

$$c_1 - \varepsilon \leq \frac{1}{2R}\int_{[-R,R]} u_0(\lfloor \xi - \eta\rfloor)\,\mathrm{d}\eta \leq c_2 + \varepsilon,$$

for large enough R. Hence we have

$$(c_1-\varepsilon)||\chi||_{L^1} \leq \int_{\mathbb{R}} \chi(\xi)u_0(\lfloor t^{\frac{1}{2}}\xi\rfloor)\,\mathrm{d}\xi \leq (c_2+\varepsilon)||\chi||_{L^1},$$

for large enough $t \geq T$, since $|I_k| \geq \delta$. For such t and χ approximating U we calculate

$$\begin{aligned}\int_{\mathbb{R}} U(t,\xi)u_0(\lfloor t^{\frac{1}{2}}\xi\rfloor)\,\mathrm{d}\xi &\geq (c_1-\varepsilon)||\chi||_{L^1} - ||u_0||_\infty ||U(t,\cdot)-\chi||_{L^1} \\ &= c_1 + \mathcal{O}(\varepsilon),\end{aligned}$$

where we used $||\chi||_{L^1} = ||U(t,\cdot)||_{L^1} + \mathcal{O}(\varepsilon) = 1 + \mathcal{O}(\varepsilon)$ in the last step. The other bound is analogous. $\square$

Combining these two results we can prove the Key Lemma 1.3:

Proof of Lemma 1.3. Let $\varepsilon > 0$. By Lemma 1.10 there exists Ψ_* (with $||d||_\infty$ depending on ε) and $N_0 = N_0(\varepsilon)$ such that

$$\frac{1}{2} - \varepsilon \leq \Lambda(\Psi_* u_0, \cdot, N) \leq \frac{1}{2} + \varepsilon,$$

for $N \geq N_0$. Because of uniform Hölder continuity (see Lemma 1.12) there exists $t_1 = t_1(\beta,\varepsilon) > 0$ such that

$$\frac{1}{2} - 2\varepsilon \leq \Lambda(u(t_1,\cdot),\cdot,N) \leq \frac{1}{2} + 2\varepsilon.$$

On the other hand, Lemma 1.11 implies that

$$u(t_1, \cdot) \geq \delta = \delta(\beta, \varepsilon) > 0,$$

for all $t \geq t_1$. In particular, equation (1.5) is strictly parabolic on (t_1, ∞) (for $\beta = 1$ this step is obsolete). Then according to Lemma 1.13 there exists $t_2 = t_2(\beta, \varepsilon)$ such that

$$\frac{1}{2} - 3\varepsilon \leq u(T, \cdot) \leq \frac{1}{2} + 3\varepsilon,$$

with $T = t_2 + t_1$. □

1.3.3. Step 3: Approximating sequence

We will now construct the approximating sequence $x^{(n)}$, using the technique established in Section 1.2. Thus for every n we have to specify the terminal data, creation times $\tau_{n,j}$ and creation operators $\Psi_*^{(n,j)}$. For the rest of the section we fix $0 < \varepsilon \leq 1/6$ and define

$$\theta^{-1} = \frac{1}{2} + \varepsilon.$$

Let T be as in Lemma 1.3 and set

$$x_{\text{ter}}^{(n)} = \theta^n,$$

a constant sequence. Now $\tau_{n,j}$ and $\Psi_*^{(n,j)}$ are constructed iteratively. We choose $\Psi_*^{(n,1)}$ according to Lemma 1.3 applied to the constant sequence 1, then by the statement of the lemma, the fact that $1/2 - \varepsilon \geq \theta^{-1}/2$ and the scaling properties of the equation we have

$$\frac{1}{2}\theta^{n-1} \leq u^{(n,1)}(T\theta^{(1-\beta)n}, \cdot) \leq \theta^{n-1},$$

where $u^{(n,1)}$ is a solution to the backward equation (1.5) with initial data $\Psi_*^{(n,1)} x_{\text{ter}}^{(n)}$ according to Lemma 1.3. Consequently we set

$$\tau_{n,1} = T\theta^{(1-\beta)n}.$$

Iterating the procedure, for given $\tau_{n,j-1}$ and $u^{(n,j-1)}$ with

$$\frac{1}{2}\theta^{n-j+1} \leq u^{(n,j-1)}(\tau_{n,j-1}, \cdot) \leq \theta^{n-j+1},$$

we apply Lemma 1.3 to the rescaled sequence

$$\theta^{-n+j-1} u^{(n,j-1)}(\tau_{n,j-1}, \cdot),$$

which yields a creation operator $\Psi_* =: \Psi_*^{(n,j)}$ and a solution $u^{(n,j)}$ to

$$\begin{cases} \partial_t u^{(n,j)} = \Delta G_\beta(u^{(n,j)}) \text{ in } (\tau_{j-1}, \tau_j] \times \mathbb{Z}, \\ u^{(n,j)}(\tau_{n,j-1}) = \Psi_*^{(n,j)} \left[u^{(n,j-1)}(\tau_{n,j-1}) \right] \end{cases},$$

that, by scaling, satisfies

$$\frac{1}{2}\theta^{n-j} \leq u^{(n,j)}(\tau_{n,j}, \cdot) \leq \theta^{n-j}, \tag{1.9}$$

with

$$\tau_{n,j} = \tau_{n,j-1} + T\theta^{(1-\beta)(n+1-j)}.$$

As described in the previous section, this procedure yields a solution $x^{(n)}$ on the interval $[0, \tau_{n,n}]$ to the coarsening equation. The local integrability condition for negative β is satisfied due to (1.7). Since $x^{(n)}(\tau_{n,n})$ is constant up to vanished particles, the solution can be extended to $[0, \infty)$. Let

$$\begin{aligned} T_n &= \tau_{n,n}, \\ t_j &= T_n - \tau_{n,n-j} = T \sum_{k=n+1-j}^{n} \theta^{(1-\beta)(n+1-k)} = T \sum_{m=1}^{j} \theta^{(1-\beta)m}. \end{aligned} \tag{1.10}$$

The numbers t_j are exactly the times where particles can vanish and are the same for all n. In particular, the vanishing times of the limit will be contained in the set $\{t_j\}$. We summarize the properties of of $x^{(n)}$ that follow directly by construction:

$$\partial_t x^{(n)} = \Delta_\sigma F_\beta(x^{(n)}) \text{ in } (0, \infty) \times \mathbb{Z}, \tag{1.11}$$

$$x^{(n)}(t,k) = \theta^n \text{ for all } t \geq T_n \text{ and } x^{(n)}(t,k) > 0, \tag{1.12}$$

$$x^{(n)}(t,k) \leq \theta^j \text{ for all } t_{j-1} \leq t \leq t_j \text{ and } k \in \mathbb{Z}, \tag{1.13}$$

$$x^{(n)}(t_j,k) \geq \frac{1}{2}\theta^j \text{ for all } 1 \leq j \leq n \text{ and } k \in \mathbb{Z} \text{ with } x^{(n)}(t_j,k) > 0. \tag{1.14}$$

1.3.4. Step 4: Passage to the limit

Before using an appropriate compactness argument on the approximating sequence it is also necessary to control the decay of particles near their vanishing times uniformly since the particle interaction is discontinuous at $x(k) = 0$.

Lemma 1.14. *Let $x^{(n)}$ be defined as above and $\beta \neq 1$. For every $j > 0$ there exists $C = C(j, \beta)$ and $\varepsilon = \varepsilon(j, \beta) > 0$, such that for all particles $k \in \mathbb{Z}$ that vanish at $t = t_j$ we have*

$$x^{(n)}(t,k) \geq C(t_j - t)^{\frac{1}{1-\beta}},$$

for $t \in [t_j - \varepsilon, t_j]$. For $\beta = 1$ the statement holds in the same way except we have the lower bound

$$x^{(n)}(t,k) \geq C \exp(-2(t_j - t)) I_L(2(t_j - t)),$$

where I_L is the L-th modified Bessel function of the first kind and L depends only on β.

Proof. The construction of $x^{(n)}$ and Lemma 1.11 directly imply the statement for $\beta \neq 1$. The inequality for $\beta = 1$ follows from Theorem 1.18 in the same way as Lemma 1.11. $\square$

With the previous preparation we can prove the main result:

Proof of Theorem 1.2. Using the explicit construction of the sequence $x^{(n)}$ and Lemma 1.12 it is easy to see that $x^{(n)}(\cdot, k)$ is uniformly Hölder continuous on $[0, T]$ for every $k \in \mathbb{Z}$ and $T > 0$ defined above. By the Arzela-Ascoli Theorem we can, after an exhaustion $T_N \nearrow +\infty$ and a diagonal argument, extract a subsequence (not renamed) and a limit $x = x(t, k)$ such that

$$x^{(n)}(t, k) \to x(t, k) \text{ as } n \to \infty \text{ for all } t \in [0, \infty) \text{ and } k \in \mathbb{Z},$$

and $x(\cdot, k) \in C([0, \infty))$ for all $k \in \mathbb{Z}$. Let

$$\eta_k^{(n)} = \inf\{t > 0 : \; x^{(n)}(t, k) = 0\}$$

denote the vanishing time of the k-th particle. By another diagonal argument we can further restrict ourselves to a subsequence such that the particle vanishing times $\{\eta_k^{(n)}\}$ converge:

$$\eta_k^{(n)} \to \eta_k \in \{t_j\}_{j=0,..,\infty} \cup +\infty \text{ as } n \to \infty \text{ for all } k \in \mathbb{Z}.$$

In fact if $\eta_k^{(n)}$ is bounded we even have $\eta_k^{(n)} = \eta_k$ for n large enough because the set $\{t_j\}$ is discrete. Otherwise we can assume $\eta_k^{(n)} \to +\infty$ increasingly. If $\eta_k = t_j$ we check that this is indeed the vanishing time of $x(\cdot, k)$:

$$0 = \lim_{n\to\infty} x^{(n)}(t, k) = x(t, k),$$

for every $t > \eta_k$. Furthermore, by Lemma 1.14 we have for $\beta \neq 1$

$$x^{(n)}(t, k) \geq C(t_j - t)^{\frac{1}{1-\beta}} > 0 \text{ for all } t \in [t_j - \varepsilon, t_j),$$

for large enough n, hence also $x(t, k) > 0$ for $t < t_j$, and the analogous argument works for $\beta = 1$ with the corresponding estimate. Next we show that x solves equation (1.4). Fix $k \in \mathbb{Z}$, $0 \leq j_1 < j_2$ and integrate equation (1.4) with $x^{(n)}$ from s_1 to s_2, where $t_{j_1} \leq s_1 < s_2 \leq t_{j_2}$, to obtain

$$x^{(n)}(s_2, k) - x^{(n)}(s_1, k) = \int_{s_1}^{s_2} \Delta_\sigma F_\beta \left(x^{(n)}\right)(t, k)\, \mathrm{d}t.$$

By construction, the function $t \mapsto \sigma_\pm(x^{(n)}(t, \cdot), k)$ is constant on (s_1, s_2). Furthermore, by the construction of the sequence $x^{(n)}$, the number of values of $\sigma_\pm(x^{(n)}(t, \cdot), k)$ regarded as a sequence in n is finite, hence we can assume this to be independent of n after taking another subsequence, in other words

$$\sigma_\pm(x^{(n)}(t, \cdot), k) = \sigma_\pm(x(t, \cdot), k) \text{ for all } s_1 < t < s_2.$$

Let $s_1 < t < s_2$. If $x(t,k) > 0$, then also $x^{(n)}(t,k) > 0$ for large n and by the point-wise convergence of $x^{(n)}$ and the above identity we conclude

$$\Delta_\sigma F_\beta(x^{(n)})(t,k) \to \Delta_\sigma F_\beta(x(t,k)) \text{ as } n \to \infty.$$

On the other hand, $x(t,k) = 0$ implies $\eta_k \leq t_{j_1}$ and $x^{(n)}(t,k) = 0$, since $\eta_k^{(n)} = \eta_k$ for n large, hence we also get $\Delta_\sigma F_\beta(x^{(n)})(t,k) \to \Delta_\sigma F_\beta(x(t,k))$. Then we can apply dominated convergence, where we use that $x^{(n)}$ is locally bounded in the case $\beta > 0$ and the lower bound from Lemma 1.14 in the case $\beta < 0$, and pass to the limit in the above integral identity to conclude

$$x(s_2,k) - x(s_1,k) = \int_{s_1}^{s_2} \Delta_\sigma F_\beta(x)(t,k)\,\mathrm{d}t,$$

which shows that x is a solution to the coarsening equation.

It remains to show that x satisfies the desired bounds. We first consider the case $\beta \neq 1$. By (1.14) we have

$$x^{(n)}(t_j,k) \geq \frac{1}{2}\theta^j,$$

for all $j \in \mathbb{N}$ and living particles, and by convergence the same inequality holds in the limit $n \to \infty$. In particular we have

$$\langle x(t_j)\rangle_\sigma^- \geq \frac{1}{2}\theta^j.$$

On the other hand, it is easy to check that $\langle x\rangle_\sigma^-$ is conserved between particle vanishings, hence we have

$$\langle x(t)\rangle_\sigma^- \geq \frac{1}{2}\theta^j,$$

whenever $t_j \leq t \leq t_{j+1}$. Because of (1.10) we have

$$t_j \sim \theta^{(1-\beta)j}.$$

This means that $t_j \leq t \leq t_{j+1}$ implies

$$\frac{1}{1-\beta}\log_\theta(t) - C \leq j \leq \frac{1}{1-\beta}\log_\theta(t) + C,$$

and consequently

$$\langle x(t)\rangle_\sigma^- \gtrsim t^{\frac{1}{1-\beta}}.$$

The upper bound on $||x||_\infty$ follows in the same way by (1.13). For the case $\beta = 1$ the same argument applies, but in this case we have

$$t_j = jT,$$

which leads to

$$\langle x(t)\rangle_\sigma^- \gtrsim \theta^{\frac{t}{T}} = \exp(\lambda t).$$

The restriction $\lambda \leq 2$ follows from the fact that equation (1.4) for $\beta = 1$ implies $\dot{x} \leq 2x$. □

The proof of Corollary 1.4 follows easily with a very similar argument:

Proof of Corollary 1.4. Let $\varepsilon > 0$. It suffices to consider the case $c = 1/2$. We apply the same construction as in the proof of Theorem 1.2 (with potentially different ε in the definition of θ) to get an approximate solution $x^{(n)}$, but in the iteration scheme we apply an additional step to $u^{(n,n)}$, satisfying $1/2 \leq u^{(n,n)} \leq 1$ according to (1.9). Using Lemma 1.3 with ε from above yields $\tilde{T}$, only depending on β and ε and a solution $u^{(n,n+1)}$ to the backward equation that satisfies

$$\left| u^{(n,n+1)}(\tau_{n,n} + \tilde{T}, \cdot) - \frac{1}{2} \right| \leq \varepsilon.$$

Then the sequence $x^{(n)}$ has the same properties as in the proof of Theorem 1.2 with the addition that the initial data are in an ε-ball around $1/2$ for all n by the above inequality, which gives the desired result after sending n to infinity. □

Remark 1.15.

- *For $\beta \in (0,1)$ the achieved growth rate is optimal, because equation (1.4) implies $\dot{x} \leq 2x^\beta$, which can be integrated to obtain $||x||_\infty \lesssim t^{\frac{1}{1-\beta}}$.*
- *The fact that it was possible to choose $\tau_{n,j} - \tau_{n,j-1} = T\theta^{(1-\beta)(n+1-j)}$ in the iteration step was crucial to obtain the desired growth rates. By comparison principle however, the estimate (1.9) also remains valid if we choose a much larger time-span between particle insertions. This means that the above method can be adapted to produce solutions that are unbounded but grow arbitrarily slowly.*
- *For convenience we chose $x_{\text{ter}}^{(n)}$ to be constant in the approximation scheme. For the construction however we only used that $\frac{1}{2}\theta^n \leq x_{\text{ter}}^{(n)} \leq \theta^n$ so that one can use arbitrary sequences satisfying these bounds in our construction scheme to produce solutions with the same coarsening behaviour.*

1.4. Analysis of the discrete FDE

Here we we address all technical results that were used in the previous sections and either prove them or give a reference. In the first three parts we discuss aspects of the discrete fast diffusion equation, while the rest of the section contains results about parabolic Hölder regularity in the discrete setting.

1.4.1. The equation $\partial_t u = \Delta G_\beta(u)$

We consider the Cauchy problem for the discrete fast diffusion equation

$$\begin{cases} \partial_t u = \frac{\beta}{|\beta|} \Delta u^\beta = \Delta G_\beta(u) & \text{in } (0,\infty) \times \mathbb{Z}, \\ u(0,\cdot) = u_0, \end{cases} \tag{1.15}$$

with $\beta \in (-\infty, 0) \cup (0, 1]$, $u_0 \in \ell^\infty_+(\mathbb{Z})$ and

$$\Delta u = u(k-1) - 2u(k) + u(k+1).$$

We are concerned with the long-time existence of classical solutions:

Definition 1.16. *A function $u : [0,\infty) \to \ell^\infty_+(\mathbb{Z})$ is a solution to problem (1.15) if the following conditions are satisfied:*

1. *$t \mapsto u(t,\cdot)$ is in $C^0([0,\infty); \ell^\infty_+(\mathbb{Z}))$ and $u(0,\cdot) = u_0$.*
2. *For every $k \in \mathbb{Z}$ we have $u(\cdot,k) \in C^1((0,\infty); \mathbb{R}_{>0})$ and*

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t,k) = \Delta G_\beta(u)(t,k),$$

for all $k \in \mathbb{Z}$ and $t > 0$.

For positive β it is not hard to prove the existence of a solution for any kind of initial data, since G_β is bounded at zero and one has simple a-priori estimates due to the comparison principle (see below). For negative β the existence of a sufficiently regular solution for arbitrary data cannot be expected due to $G_\beta(x)$ becoming singular at $x = 0$. Similar to the result in [33], we have to restrict to initial data which satisfy a certain positivity condition:

Definition 1.17. *For $u \in \ell^\infty_+(\mathbb{Z})$ and $d > 0$ let*

$$\sigma_+(u,k,d) = \inf\{l > k : u(l) \geq d\}.$$

Then for any $L \in \mathbb{N}$ and $d > 0$ we define

$$\mathcal{P}_{L,d} = \left\{ u \in \ell^\infty_+ : \sup_{k \in \mathbb{Z}} \sigma_+(u,k,d) \leq L \right\}.$$

In other words, $u \in \mathcal{P}_{L,d}$ means that particles with large mass cannot be very far apart. This is also relevant for the case $\beta > 0$ since it allows us to prove certain Harnack-type positivity estimates. We have the following result:

Theorem 1.18. *Let $\beta \in (-\infty, 0) \cup (0, 1]$, and consider initial data $u_0 \in \mathcal{P}_{L,d}$. Then the following statements hold:*

1. *If $\beta \in (-\infty, 0) \cup (0, 1)$, there exists a positive constant $c = c(L, d, \beta, ||u_0||_\infty)$ and a solution u to equation (1.15) on $[0, \infty)$ with initial data u_0 satisfying*

$$u(t,k) \geq c\left(1 \wedge t^{\frac{1}{1-\beta}}\right), \text{ for all } k \in \mathbb{Z}. \tag{1.16}$$

2. *If $\beta = 1$, the same statement holds with estimate (1.16) replaced by*

$$u(t,k) \geq c \exp(-2t) I_L(2t),$$

   where $I_L(t)$ denotes the L-th modified Bessel function of the first kind.

3. *Comparison principle: If $c_1 \leq u_0 \leq c_2$, then u satisfies these bounds for all times.*

In this and the next two sections we give a full proof of the above result. The general strategy to prove existence of solutions for equation (1.5) is to use regularization and standard ODE theory. Instead of infinitely many particles with non-negative mass we first consider a periodic N-particle ensemble where particles have strictly positive mass. The first important a-priori estimate is the comparison principle:

Lemma 1.19 (Finite positive ensemble)**.** *Let $\mathbb{T}_N$ denote the one-dimensional periodic lattice with N lattice points. Let $u_0 \in \ell^\infty_+(\mathbb{T}_N)$ with $0 < \delta \leq u_0 \leq C$. Then there exists a unique solution $u : [0, \infty) \to \ell^\infty_+(\mathbb{T}_N)$ of (1.15) with $\delta \leq u(t, \cdot) \leq C$.*

Proof. The proof is very similar to Lemma 2 in [18] and a standard maximum principle argument. Because $u_0 \geq \delta$, standard ODE theory gives the existence and uniqueness of a smooth solution u on the time interval $[0, t^*]$ to equation (1.15) with $\delta/2 \leq u(t, \cdot) \leq 2C$ for some positive t^*. For small $\varepsilon > 0$ we then consider the solution u_ε of the modified problem

$$\begin{aligned} \partial_t u_\varepsilon &= \Delta G_\beta(u_\varepsilon) + \varepsilon, \\ u_\varepsilon(0, \cdot) &= u_0, \end{aligned}$$

that exists on the same time interval as u and satisfies the same bounds after possibly making t^* smaller. Because G_β is smooth on $[\delta/2, 2C]$ we have that $u_\varepsilon \to u$ uniformly on $[0, t^*]$. We claim that u_ε attains its minimum over $[0, t^*] \times \mathbb{T}_N$ at $t = 0$. If not, there exists $t_0 \in (0, t^*]$ and k_1 such that $u_\varepsilon(t_1, k_1)$ is the absolute minimum. Consequently we get

$$0 \geq \partial_t u_\varepsilon(t_1, k_1) = \Delta G_\beta(u_\varepsilon)(t_1, k_1) + \varepsilon \geq \varepsilon,$$

a contradiction. Here we used that G_β is increasing. Hence $u_\varepsilon(t, \cdot) \geq \delta$ for all $t \in [0, t^*]$ and, sending $\varepsilon \to 0$, the same bound holds for u. The same argument for the maximum where $+\varepsilon$ is replaced with $-\varepsilon$ yields that $u \leq C$. Iterating from $t = t^*$, we see that the solution can be extended to $[0, \infty)$ and always satisfies the desired bounds. $\square$

From this result we can easily pass to the limit as $N \to \infty$ to obtain solutions for infinite numbers of particles:

Corollary 1.20 (Infinite positive ensemble)**.** *Let $u_0 \in \ell^\infty_+(\mathbb{Z})$ with $0 < \delta \le u_0 \le C$. Then there exists a solution $u : [0,\infty) \to \ell^\infty_+(\mathbb{Z})$ of (1.15) with $\delta \le u(t,\cdot) \le C$.*

Proof. This is a standard compactness argument. We choose $u_0^{(N)}$ to be N-periodic such that $u_0^{(N)}(k) \to u_0(k)$ for each $k \in \mathbb{Z}$. Let $u^{(N)}$ be the corresponding solutions from the above lemma. Then due to the a-priori bounds $\delta \le u^{(N)} \le C$ and equation (1.15) we have

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}u^{(N)}\right| = \left|\Delta G_\beta\left(u^{(N)}\right)\right| \le K(\delta, C),$$

hence the solutions are uniformly Lipschitz continuous. Applying the Arzela-Ascoli Theorem and a diagonal argument we can extract a convergent subsequence (not relabeled) such that $u^{(N)}(\cdot,k) \to u(\cdot,k)$ uniformly on compact time intervals, where $u(\cdot,k) \in C^0([0,\infty))$. In particular u satisfies the same bounds as $u^{(N)}$. Integrating (1.15) in time and passing to the limit (which is possible due to the a-priori bounds) then yields

$$u(t,k) = u_0(k) + \int_0^t \Delta_\sigma G_\beta(u)(s,k)\,\mathrm{d}s.$$

This in turn shows that $u(\cdot,k)$ is continuously differentiable and solves (1.15) pointwise. Again, the bounds on u yield Lipschitz continuity in $\ell^\infty_+(\mathbb{Z})$. □

For our purpose, we need the existence of solutions in particular for initial data with mass zero particles. The general strategy is to approximate the initial data by regularized data via

$$u_{0,\delta} = u_0 \vee \delta.$$

The above existence result then yields long-time solutions u_δ with initial data $u_{0,\delta}$. In the case $\beta > 0$ one can pass to the limit $\delta \to 0$ in the same manner as above, since G_β is bounded at zero, yielding a general existence result:

Corollary 1.21 (Existence for positive β)**.** *Let $\beta \in (0,1]$ and $u_0 \in \ell^\infty_+(\mathbb{Z})$. Then there exists a solution $u : [0,\infty) \to \ell^\infty_+(\mathbb{Z})$ of (1.15).*

Alternatively it is likely possible to prove this result directly via an infinite dimensional fixed-point method. Since we need Corollary 1.20 for the case $\beta < 0$ anyway, the above method is the fastest for our purpose. In the next section we deal with the negative β case, including existence and the positivity estimate (1.16). Then we prove the positivity estimate for positive β, completing the proof of Theorem 1.18.

1.4.2. Existence of solutions for $\beta < 0$

In the following we always assume $\beta < 0$. The key idea to prove existence of solutions is to exploit the fact that regions which are enclosed by large particles (called *traps*) are screened from the rest of the particles, very similar to [33]. One important difference however is the fact that the backward equation does not yield a-priori estimates on the persistence of traps. We make the following definition:

Definition 1.22. *We say that a solution u to equation (1.15) with initial data $u_0 \in \mathcal{P}_{L,d}$ has the persistence property on $[0,T]$ if $u(0,k) \geq d$ implies $u(t,k) \geq \frac{d}{2}$ for all $t \in [0,T]$.*

By making use of the theory for the coarsening equation developed in [33] we have the following result concerning Hölder regularity:

Lemma 1.23. *There exist constants $T' = T'(\beta, d)$ and $C = C(\beta, L) > 0$ such that the following holds: If a solution u to equation (1.15) with initial data $u_0 \in \mathcal{P}_{L,d}$ has the persistence property on $[0,T]$ and $T \leq T'$, then*

$$|u(t_2,k) - u(t_1,k)| \leq C|t_2 - t_1|^{\frac{1}{1-\beta}},$$

for all $t_2, t_1 \in [0,T]$ and $k \in \mathbb{Z}$.

Proof. We consider the time reversed function

$$x(s,k) = u(T-s,k),$$

then x is a solution to the coarsening equation for $0 \leq s \leq T$ that satisfies $x(0,\cdot) \in \mathcal{P}_{L,\frac{d}{2}}$. Applying Lemma 3.3 from [33] (if $T \leq T^*(\beta, \frac{d}{2}) =: T'$) yields the desired Hölder continuity for x, and thus also for u. $\square$

From this result we derive the first a-priori estimate:

Lemma 1.24. *Let $u_{0,\delta}$ be as above and let u_δ be the corresponding solution of equation (1.15), which exists by Lemma 1.20. Then there exists $T = T(L,d) > 0$ such that u_δ has the persistence property on $[0,T]$.*

Proof. First we note that because the solution satisfies $u_\delta \geq \delta$ for all times we have the Lipschitz estimate

$$|\partial_t u_\delta| \leq 4\delta^\beta.$$

This means that $u_\delta(0,k) \geq d$ implies $u_\delta(t,k) \geq \frac{d}{2}$ for $0 \leq t \leq t_0$, where

$$t_0 = \frac{d}{8\delta^\beta}.$$

Let T be the largest time such that $u_\delta(0,k) \geq d$ implies $u_\delta(t,k) \geq \frac{d}{2}$ on $[0,T]$. By the above considerations we already know that $T > 0$. If $T \leq T'(\beta, d)$ we can apply Lemma 1.23 and get

$$u_\delta(t,k) \geq u_\delta(0,k) - Ct^{\frac{1}{1-\beta}}.$$

If $u_\delta(0,k) \geq d$, this implies $u_\delta(T,k) \geq d - CT^{\frac{1}{1-\beta}}$. On the other hand, by the definition of T there exists such a k with $u_\delta(T,k) \leq \frac{3d}{4}$, hence

$$\frac{3d}{4} \geq d - C(L)T^{\frac{1}{1-\beta}},$$

which gives a lower bound for T in terms of L and d. $\square$

The next a-priori estimate is crucial to get uniform Hölder bounds on u_δ, as well as integral bounds which are needed to pass to the limit.

Lemma 1.25. *Let u_δ be as above. Then there exists $c = c(\beta, L, d) > 0$ and $t^* = t^*(\beta, L, d) > 0$ such that*

$$u_\delta(t, \cdot) \geq c t^{\frac{1}{1-\beta}},$$

for $0 \leq t \leq t^*$.

Proof. We apply a very similar argument as in the proof of Lemma 3.5 in [33]. If the statement is false, there exist sequences $u^{(n)}$, $t_n \to 0$, $\delta_n \to 0$ and $k_n \in \mathbb{Z}$ such that

$$u_{\delta_n}^{(n)}(t_n, k_n) \leq \frac{1}{n} t_n^{\frac{1}{1-\beta}}.$$

By translation invariance we can assume that $k_n = k_0$ is constant. We rescale and define

$$v_n(s, k) = t_n^{\frac{1}{\beta-1}} u_{\delta_n}^{(n)}(t_n s, k).$$

Then v_n is a solution to equation (1.15) with $v_n(1, k_0) \to 0$. Additionally we have $v_n(0, \cdot) \in \mathcal{P}_{L,d}$ and v_n satisfies the persistence property on $[0, 1]$ for large n by Lemma 1.24. Since $t_n^{\frac{1}{\beta-1}} d \to \infty$ and $T' \to \infty$ as $d \to \infty$ we also have that v_n is uniformly Hölder continuous by Lemma 1.23. Let B be the largest set of consecutive indices containing k_0 such that

$$\liminf_{n\to\infty} v_n(1, k) = 0,$$

for $k \in B$. Observe that we have $|B| \leq L$ due to $t_n^{\frac{1}{\beta-1}} d \to \infty$ and the persistence property. Let l_-, l_+ be the nearest particle index to the left, respectively to the right of B. We restrict to a subsequence such that $v_n(1, k) \to 0$ for $k \in B$ and $v_n(1, l_\pm) \geq \lambda > 0$. If we define the local mass $M_n(s)$ as

$$M_n(s) = \sum_{k \in B} v_n(s, k),$$

then an elementary calculation gives

$$\frac{\mathrm{d}}{\mathrm{d}s} M_n(s) = v_n^\beta(s, l_- + 1) - v_n^\beta(s, l_-) + v_n^\beta(s, l_+ - 1) - v_n^\beta(s, l_+). \tag{1.17}$$

Due to uniform Hölder continuity and particles in B going to zero there exists $\varepsilon > 0$ such that

$$v_n^\beta(s, l_\pm) \leq \frac{1}{2} v_n^\beta(s, l_\pm \mp 1),$$

for $s \in [1-\varepsilon, 1]$ and large enough n. Using equation (1.17) on this time interval we obtain

$$2\frac{\mathrm{d}}{\mathrm{d}s}M_n(s) \geq v_n^\beta(s, l_- + 1) + v_n^\beta(s, l_+ - 1) \geq M_n^\beta(s),$$

which, after integrating from $1-\varepsilon$ to 1 yields

$$M_n(1) \geq \tilde{\varepsilon} > 0,$$

which gives a contradiction after sending n to infinity. $\square$

This result gives us the a-priori estimates we need to pass to the limit:

Corollary 1.26 (Hölder continuity)**.** *Let u_δ and t^* be as above. Then there exists $C = C(\beta, L, d)$ such that*

$$|u_\delta(t_2, k) - u_\delta(t_1, k)| \leq C|t_2 - t_1|^{\frac{1}{1-\beta}},$$

for all $t_2, t_1 \in [0, t^*]$ and $k \in \mathbb{Z}$.

Proof. Let $t_2, t_1 \in [0, t^*]$, $t_2 > t_1$. We integrate (1.15) in time and estimate

$$\begin{aligned}|u(t_2, k) - u(t_1, k)| &\leq \int_0^t |\Delta G(u_\delta)(s, k)| \,\mathrm{d}s \lesssim \int_{t_1}^{t_2} s^{\frac{\beta}{1-\beta}} \,\mathrm{d}s \\ &\sim t_2^{\frac{1}{1-\beta}} - t_1^{\frac{1}{1-\beta}} \leq (t_2 - t_1)^{\frac{1}{1-\beta}},\end{aligned}$$

where we used Lemma 1.25 to estimate $|\Delta G_\beta(u_\delta)|$. $\square$

With these preparations we can prove the first statement of Theorem 1.18 for negative β:

Proof of existence and positivity bound for $\beta < 0$. Let $u_{0,\delta}$ and u_δ as above. By Corollary 1.26 and the Arzela-Ascoli Theorem there exists a subsequence $\delta \to 0$ such that $u_\delta \to u$ uniformly on $[0, t^*]$. Moreover, by Lemma 1.25 we have

$$u_\delta^\beta(t, k) \leq c^\beta t^{\frac{\beta}{1-\beta}},$$

which implies $u_\delta^\beta \to u^\beta$ in $L^1([0, t^*])$. Using this we can pass to the limit in the integral equation

$$u_\delta(t, k) = u_{0,\delta}(k) - \int_0^t \Delta G(u_\delta)(s, k) \,\mathrm{d}s,$$

showing that u is a solution to the backward equation with initial data u_0 on $[0, t^*]$. Since the lower bound from Lemma 1.25 also holds in the limit, we can extend the solution from t^* to arbitrary times via comparison principle, which also changes the lower bound to $\sim 1 \wedge t^{\frac{1}{1-\beta}}$ for large times. $\square$

1.4.3. Harnack-type inequality for $0 < \beta \leq 1$

In the previous part Lemma 1.25 was crucial to prove existence of a solution. The result of the lemma, together with the positivity condition $\mathcal{P}_{L,d}$ can be interpreted as a Harnack-type inequality, see [9]. For $0 < \beta < 1$ a similar result holds, the equation however behaves differently and the indirect proof does not work here. We will pursue another approach and show the inequality directly with explicit constants, handling the case $\beta = 1$ separately. The key observation is that a large particle next to a small particle will always induce growth on the small particle, despite the size of the other neighbour of the small particle. This decouples the equation in a sense and we only need to study the local problem:

Lemma 1.27 (Local Problem). *Let $0 < \beta < 1$ and $T > 0$. Consider two functions $F \in C^0([0,T];[0,\infty))$ and $u \in C^1([0,T];[0,\infty))$ which satisfy*

$$\begin{aligned} F(t) &\geq c t^{\frac{\beta}{1-\beta}}, \\ \dot{u}(t) &\geq F(t) - 2u^\beta(t), \end{aligned}$$

on $[0,T]$. Then u satisfies

$$u(t) \geq \eta_1(c) t^{\frac{1}{1-\beta}},$$

on the interval $[0,T]$, where η_1 is a positive strictly increasing function which depends only on β.

Proof. We define the rescaled function

$$v(t) = t^{\frac{1}{\beta-1}} u(t),$$

on the half-open interval $(0,T]$. Then it suffices to show that v is bounded from below by $\eta_1(c)$. We calculate

$$\begin{aligned} t\dot{v}(t) &= t\left(t^{\frac{1}{\beta-1}}\dot{u}(t) + \frac{1}{\beta-1} t^{\frac{1}{\beta-1}-1} u(t)\right) \\ &\geq t\left(t^{\frac{1}{\beta-1}} F(t) - 2t^{\frac{1}{\beta-1}} u^\beta(t) + \frac{1}{\beta-1} t^{\frac{1}{\beta-1}-1} u(t)\right) \\ &= t^{\frac{\beta}{\beta-1}} F(t) - 2v^\beta(t) - \frac{1}{1-\beta} v(t) \\ &=: t^{\frac{\beta}{\beta-1}} F(t) - \theta(v(t)), \end{aligned}$$

in particular the assumption on F implies that

$$t\dot{v}(t) \geq c - \theta(v(t)). \tag{1.18}$$

Since the function θ is strictly increasing on $[0,\infty)$ we can define the inverse function $\eta_1 = \theta^{-1}$, which is also strictly increasing. We claim that $v \geq \eta_1(c)$ on $(0,T]$. If this is not true, there is $\varepsilon > 0$ and $t^* \in (0,T]$ such that $v(t^*) \leq \eta_1(c) - \varepsilon$. In particular we have

$$c - \theta(v(t^*)) \geq \tilde{\varepsilon} > 0,$$

for some $\tilde{\varepsilon} > 0$. But then the differential inequality (1.18) implies that $v(t) \leq \eta_1(c) - \varepsilon$, and hence $c - \theta(v(t)) \geq \tilde{\varepsilon}$ for all $t \in (0,t^*]$. Dividing by t and integrating (1.18) in time gives

$$v(t^*) - v(t) \geq \varepsilon \log\left(\frac{t^*}{t}\right),$$

for all $0 < t \leq t^*$. Sending t to zero then gives a contradiction. □

The above lemma enables us to prove a Harnack type inequality:

Lemma 1.28 (Harnack type inequality)**.** *Let u be a solution to equation (1.15) with $0 < \beta < 1$ and initial data $0 \leq u_0 \leq 1$. Then we have*

$$u(t,k) \geq \eta(|k-l|)(t-s)^{\frac{1}{1-\beta}},$$

for all $k,l \in \mathbb{Z}$ and $0 \leq t-s \leq t^*(u(s,l))$. The function η is strictly positive and the function t^* is non-negative, strictly increasing with $t^*(u) = 0$ iff $u = 0$. Furthermore, both functions depend only on β.

Proof. Due to translation invariance in space and time it suffices to consider the case $s = 0$ and $l = 0$. We will make an iterative argument, using Lemma 1.27 in each step. First we note that due to $u_0 \leq 1$ and the comparison principle we have the Lipschitz estimate

$$|\dot{u}| \leq 4,$$

in particular

$$u(t,0) \geq u_0(0) - 4t.$$

The case $u_0(0) = 0$ is trivial. If $u_0(0) > 0$, the Lipschitz estimate implies

$$u(t,0) \geq t^{\frac{1}{1-\beta}},$$

whenever $u_0(0) - 4t - t^{\frac{1}{1-\beta}} > 0$. If we set $f(t) = 4t + t^{\frac{1}{1-\beta}}$ then t^* is defined as the inverse of f. Thus the above lower bound holds for $0 \leq t \leq t^*(u_0(0))$. For $u(t,1)$ we have

$$\begin{aligned}\dot{u}(t,1) &= u^\beta(t,0) - 2u^\beta(t,1) + u^\beta(t,2)\\ &\geq u^\beta(t,0) - 2u^\beta(t,1).\end{aligned}$$

This means that $u^\beta(t,0)$ and $u(t,1)$ satisfy the assumptions of Lemma 1.27 with $T = t^*(u_0(0))$ and $c = 1$. Thus we have

$$u(t,1) \geq \eta_1(1) t^{\frac{1}{1-\beta}},$$

on $[0, t^*(u_0(0))]$. Now we can successively apply the same argument to the pairs of functions $(u^\beta(t,1)$,$u(t,2))$,...,$(u^\beta(t,k-1)$,$u(t,k))$, for $k \in \mathbb{N}$. The argument for $-k$ is the same. Then the desired inequality follows with the function $\eta = \eta(r)$ $(r \in \mathbb{N})$ defined as

$$\begin{aligned} \eta(r) &= \eta_1^{(r)}(1), \\ \eta(0) &= 1, \end{aligned}$$

where $\eta_1^{(r)}$ means that η is r times composed with itself. □

Lemma 1.29. *Let u be a solution to the constant coefficient linear equation (1.15) with $\beta = 1$. Then for every $k \in \mathbb{Z}$ and $N \in \mathbb{N}$ we have*

$$u(t,k) \geq M(u_0,k,N) \exp(-2t) I_N(2t),$$

where $I_N(t)$ denotes the N-th modified Bessel function of the first kind and

$$M(u_0,k,N) = \sum_{l=-N}^{N} u_0(k-l)$$

denotes the local initial mass.

Proof. In the linear constant-coefficient case we can give an explicit formula by Fourier-analysis (cf. [1, p. 376]). We write

$$\hat{u}(t,\theta) = \sum_{k=-\infty}^{k=+\infty} u(t,k) \exp(-ik\theta),$$

taking the time derivative on both sides and using the equation then yields

$$\begin{aligned} \partial_t \hat{u}(t,\theta) = \exp(-i\theta) &\sum_{k=-\infty}^{k=+\infty} u(t,k) \exp(-ik\theta) \\ &- 2 \sum_{k=-\infty}^{k=+\infty} u(t,k) \exp(-ik\theta) \\ &+ \exp(i\theta) \sum_{k=-\infty}^{k=+\infty} u(t,k) \exp(-ik\theta) \\ &= 2(\cos(\theta) - 1) \hat{u}(t,\theta). \end{aligned}$$

We solve this ODE in t with initial data $\hat{f}$ to obtain

$$\hat{u}(t,\theta) = \hat{f}(\theta) \exp(2t(\cos(\theta) - 1)),$$

which gives the discrete heat kernel

$$\begin{aligned}\phi(t,k) &= \exp(-2t)\frac{1}{2\pi}\int_{-\pi}^{\pi}\exp(2t\cos(\theta)-ik\theta)\,\mathrm{d}\theta\\ &= \exp(-2t)\frac{1}{\pi}\int_{0}^{\pi}\exp(2t\cos(\theta))\cos(k\theta)\,\mathrm{d}\theta\\ &= \exp(-2t)I_k(2t),\end{aligned}$$

where I_k is the kth modified Bessel function of the first kind. Then the desired inequality follows directly by the standard representation

$$u(t,k) = \sum_{l\in Z} u_0(k-l)\phi(t,l),$$

the fact that ϕ is decreasing in the second argument and the obvious estimate. □

We summarize the findings of this section and prove the remaining statements of Theorem 1.18:

Proof of positivity estimate for $0 < \beta \leq 1$. First we consider the case $\beta \neq 1$. It suffices to consider the case $u_0 \leq 1$ by scaling (this means that for general data the constants get an additional dependence on $||u_0||_\infty$). Because $u_0 \in \mathcal{P}_{L,d}$, for every $k \in \mathbb{Z}$ there exists k' with $u_0(k') \geq d$ and $|k-k'| \leq L$. Then Lemma 1.28 with $s=0$ and $l=k'$ yields

$$u(t,k) \geq \min_{j=1,..,L} \eta(j)t^{\frac{1}{1-\beta}},$$

for $0 \leq t \leq t^*(d)$ because t^* is monotone. For $\beta = 1$ we note that $u_0 \in \mathcal{P}_{L,d}$ implies that $M(u,k,L) \geq 2d$, since there are at least two terms in the sum that are greater than or equal to d by definition of $\mathcal{P}_{L,d}$. Then the statement follows directly from Lemma 1.29. □

1.4.4. Nash-Aronson estimates and Hölder continuity

For $u = u(k) \in \ell^\infty(\mathbb{Z})$ we define the forward and backward difference operators

$$\begin{aligned}\partial^+ u(k) &= u(k+1) - u(k)\\ \partial^- u(k) &= u(k) - u(k-1).\end{aligned}$$

For $a = a(t,k)$ with $0 < c_1 \leq a \leq c_2$ and $a(\cdot,k) \in C^0([0,\infty))$ we consider the discrete analogue to a parabolic evolution equation in divergence form:

$$\begin{cases}\partial_t u = \partial^-(a\partial^+ u) =: \mathcal{L}(t)u\\ u(0,\cdot) = u_0.\end{cases} \tag{1.19}$$

We denote by $\phi(t,k,s,l)$ the fundamental solution to (1.19), in other words, $\phi(\cdot,\cdot,s,l)$ is the solution to the above equation starting at time s with initial data $\phi_0(k) = \delta_{kl}$. Since $\mathcal{L}(t)$ is a bounded operator from $\ell^2(\mathbb{Z})$ to $\ell^2(\mathbb{Z})$, ϕ can be written as

$$\phi(t,k,s,l) = \left\langle \exp\left(\int_s^t \mathcal{L}(r)\,\mathrm{d}r\right)\delta_l, \delta_k \right\rangle,$$

where δ_k are the canonical basis vectors in $\ell^2(\mathbb{Z})$. The general solution starting at time $t = s$ to (1.8) is then given by

$$u(t,k) = \sum_{l\in\mathbb{Z}^d} \phi(t,k,s,l)u_0(l).$$

We also define the reduced fundamental solution

$$\psi(t,k) = \phi(t,k,0,0) = \phi(t,0,0,k),$$

and the corresponding "macroscopic" rescaled function

$$\begin{aligned} &U : \mathbb{R} \to [0,+\infty) \\ &U(t,\xi) = t^{\frac{1}{2}}\psi(t,\lfloor t^{\frac{1}{2}}\xi\rfloor). \end{aligned}$$

We have the following Nash-Aronson estimates on the fundamental solution:

Theorem 1.30. *There exist constants $t_0 > 0$, $C > 0$ and $\alpha > 0$, depending only on the bounds on a, such that the following statements hold:*

- *Aronson estimate:*

$$\psi(t,k) \le \frac{C}{1\vee t^{\frac{1}{2}}}\exp\left(-\frac{|k|}{1\vee t^{\frac{1}{2}}}\right), \tag{1.20}$$

  for every $k\in\mathbb{Z}$ and $t \ge 0$.

- *Nash continuity estimate:*

$$|\psi(t,k) - \psi(t,l)| \le \frac{C}{t^{\frac{1}{2}}}\left(\frac{|k-l|}{t^{\frac{1}{2}}}\right)^{\alpha}, \tag{1.21}$$

  for every $k,l\in\mathbb{Z}$ and $t \ge 0$.

Proof. Here we cite the results from Appendix B of [25]. Inequality (1.20) is precisely the statement of Proposition B.3. For the second inequality (1.21) we first note that (1.20) implies $|\psi| \lesssim t^{-\frac{1}{2}}$. Then the desired estimate at a time t^* follows from Proposition B.6 applied at $t = s = t^*/2$ with $f = \psi(t^*/2,\cdot)$ and the semigroup property. □

These estimates have important consequences for the function U. Inequality (1.20) implies that

$$U(t, \xi) \le \Phi(\xi),$$

for some integrable function Φ. In particular the function family $U(t, \cdot)$ are tight probability measures. On the other hand, the estimate (1.21) implies that the function $U(t, \cdot)$, which is a step-function by definition, becomes Hölder continuous in the following sense:

Definition 1.31 (Approximate Hölder Continuity)**.** *Let $\{f_n\} \subset L^\infty(\mathbb{R})$ be a sequence of functions. Then $\{f_n\}$ is said to be **approximately Hölder continuous** with exponent $\alpha \in (0, 1]$ if for every $\varepsilon > 0$ there exists $n = n(\varepsilon)$ such that $|x - y| \ge \varepsilon$ implies*

$$|f_n(x) - f_n(y)| \le C|x - y|^\alpha, \tag{1.22}$$

for $n \ge n(\varepsilon)$ and a universal positive constant C.

The important observation is that Hölder continuity on the discrete microscopic level implies approximate Hölder continuity on the macroscopic scale:

Lemma 1.32. *The function $U(t, \cdot)$ is approximately Hölder continuous as $t \to \infty$. Furthermore, the constants C, α and $t = t(\varepsilon)$ only depend on the bounds of the coefficient a in (1.19).*

Proof. By the estimate (1.21) from Theorem 1.30 we calculate

$$|U(t, \xi) - U(t, \eta)| \lesssim \left(\frac{|\lfloor t^{\frac{1}{2}} \xi \rfloor - \lfloor t^{\frac{1}{2}} \eta \rfloor|}{t^{\frac{1}{2}}} \right)^\alpha.$$

Since

$$\frac{|\lfloor t^{\frac{1}{2}} \xi \rfloor - \lfloor t^{\frac{1}{2}} \eta \rfloor|}{t^{\frac{1}{2}}} = |\xi - \eta| + \mathcal{O}(t^{-\frac{1}{2}}),$$

we get the desired estimate for $|\xi - \eta| \gtrsim t^{-\frac{1}{2}}$. □

The next result is of crucial importance for the main result of the chapter. Denote by $\mathcal{T}_\delta(\mathbb{R})$ the set of step-functions with step-width at least δ. Then we have:

Lemma 1.33. *Let $(f_n) \subset L^1(\mathbb{R})$ be tight and approximately Hölder continuous. Then for every $\varepsilon > 0$ there exists n_0 and $\delta > 0$, such that for every f_n with $n \ge n_0$ there exists $\chi \in \mathcal{T}_\delta(\mathbb{R})$ with*

$$||f_n - \chi||_{L^1(\mathbb{R})} \le \varepsilon.$$

Proof. Let $\varepsilon > 0$. Because (f_n) is tight in L^1 there exists $R > 0$ such that

$$\int_{|x|\geq R} |f_n(x)| \,\mathrm{d}x \leq \varepsilon,$$

hence it suffices to approximate (f_n) in L^∞. By approximate Hölder continuity there exists n_0 such that

$$|f_n(x) - f_n(y)| \leq C|x-y|^\alpha,$$

for $n \geq n_0$ and $|x-y| \geq R^{-\frac{1}{\alpha}}\varepsilon$. This means that the piecewise-constant interpolation χ of f_n with step-width $R^{-\frac{1}{\alpha}}\varepsilon$ approximates f_n uniformly up to an error of $R^{-1}\varepsilon^\alpha$, hence

$$||f_n - \chi||_{L^1} \lesssim \varepsilon + \varepsilon^\alpha.$$

□

Combining the last two lemmas we obtain the following corollary, which is used in the proof of the main result:

Corollary 1.34. *For every $\varepsilon > 0$ there exists $T > 0$ and $\delta > 0$, such that for every $U(t,\cdot)$ with $t \geq T$ there exists $\chi \in \mathcal{T}_\delta(\mathbb{R})$ with*

$$||U(t,\cdot) - \chi||_{L^1(\mathbb{R})} \leq \varepsilon.$$

Furthermore, T and δ only depend on ε and the bounds on a.

Proof. Approximate Hölder continuity was already established, while tightness in L^1 follows from the estimate (1.20) of Theorem 1.30. The dependence of the constants is easily checked revisiting the proofs of the previous two lemmas. □

2. Well-posedness and self-similar solutions of singular LSW equations

2.1. Introduction and results

2.1.1. Singular LSW equations

In this chapter we investigate a well-known model of coarsening processes after Lifshitz, Slyozov and Wagner [47, 59], LSW for short, where particles exchange mass at a rate which depends on a certain system average, the so called mean-field. For finitely many particles with corresponding particle masses or sizes $x_1, ..., x_N$ the time evolution is given by the ODE

$$\dot{x}_j = \begin{cases} a(x_j)\theta(t) - b(x_j), & \text{if } x_j > 0, \\ 0, & \text{if } x_j = 0, \end{cases} \tag{2.1}$$

where a, b are given functions depending on the physical model and θ is the mean-field which is determined by the condition that the total mass of the system $\sum x_j$ should be conserved. Consequently, the value of the mean-field is given by

$$\theta = \frac{\sum_{x_j>0} b(x_j)}{\sum_{x_j>0} a(x_j)}. \tag{2.2}$$

Thus in the system (2.1) all living particles (i.e particles of non-zero size) interact with each other through the mean-field and particles with size 0 vanish from the system, which leads to growth of the average particle size, hence coarsening. If the particles are not discrete but instead described in terms of a size distribution function $f(t, x)$, the time evolution is given by the first order continuity equation

$$\partial_t f = \partial_x((b - a\theta)f), \tag{2.3}$$

with mean-field

$$\theta = \frac{\int_0^\infty b\, f \, \mathrm{d}x}{\int_0^\infty a\, f \, \mathrm{d}x}. \tag{2.4}$$

In the classical case of Ostwald ripening we have $a(x) = x^{\frac{1}{3}}, b(x) = 1$, but also more complicated examples are possible. Typically a, b are required to behave like certain power laws at $x = 0$ and $x = \infty$, where the exponents depend on the space dimension

(3D or 2D particles). The well-posedness of equation (2.3) was established in [51] for the case of Ostwald ripening and subsequently generalized by [43, 44, 52] to a fairly large class of functions a, b and also more general mass constraints. One physical scenario not covered yet is the case of two-dimensional diffusion-controlled growth, where $a(x) = 1$, $b(x) = x^{-\frac{1}{2}}$. The existing theory fails here due to the singular nature of $b(x)$ at $x = 0$.

In this chapter we are interested in the more general singular case where the the coefficients are given by $a(x) = 1$, $b(x) = x^{-\beta}$ for some $\beta > 0$. Then the LSW system (2.1) becomes

$$\dot{x}_j = \left(\theta - x_j^{-\beta}\right) \cdot \mathbf{1}_{\{x_j>0\}}, \quad \theta = \frac{\sum_{x_j>0} x_j^{-\beta}}{|\{j : x_j > 0\}|}, \tag{2.5}$$

respectively

$$\partial_t f = \partial_x \left((x^{-\beta} - \theta) f\right), \quad \theta = \frac{\int_0^\infty x^{-\beta} f \, \mathrm{d}x}{\int_0^\infty f \, \mathrm{d}x}. \tag{2.6}$$

Our main result Theorem 2.3 is that a general measure solution can be approximated in the L^1-Wasserstein distance by a sequence of empirical measure solutions stemming from the discrete equation (2.5). In particular this result implies the existence of general measure-valued solutions for this class of models, which was not covered before. Apart from existence of general solutions, we also investigate self-similar solutions. We find that there exists a one-parameter family of self-similar solutions which all have compact support but only one of them is smooth, which is a rather typical phenomenon for LSW-type equations.

Besides extending the existing theory for LSW equations, we are also interested in the connections between equation (2.6) and other models. First we note that equation (2.5), respectively (2.6) is the mean-field version of the coarsening model from Chapter 1 with negative exponent (here we use $-\beta$ instead of β for convenience). Recall that in this model, the configuration space is $\ell^\infty_+(\mathbb{Z})$ and particles interact via

$$\dot{x}_j = \begin{cases} x_{\sigma^-(j)}^{-\beta} - 2x_j^{-\beta} + x_{\sigma^+(j)}^{-\beta}, & \text{if } x_j > 0, \\ 0, & \text{if } x_j = 0. \end{cases} \tag{2.7}$$

The expressions $\sigma^-(j), \sigma^+(j)$ denote the nearest particle index to the left, respectively to the right of j that has non-zero mass. As in the mean-field case, the particle configuration $\{x_j\}_{j\in\mathbb{Z}}$ coarsens as smaller particles vanish and the average particle size grows. However, due to the local nature of the model, there is no closed equation for the size-distribution. A natural question is whether the statistics of this system behave like a solution to equation (2.6) if the spatial distribution of particles is homogeneous enough.

Another aspect of the LSW equation is the connection to the model of exchange-driven growth (EDG), which we briefly describe here (for a more exhaustive treatment we refer to Chapter 3 which is fully dedicated to the study of EDG with product kernel). Consider atomic particles distributed among sites in a completely connected graph and the stochastic jump process where a a particle at a site with k particles jumps to a site with l at a rate $K(k,l)$. In the limit of diverging particle number N and graph size L with $N/L \to \rho$ the dynamics are described by the statistical master equation

$$\begin{aligned} \dot{c}_k &= \sum_{l\geq 1} K(l,k-1)c_l c_{k-1} - \sum_{l\geq 1} K(k,l-1)c_k c_{l-1} && (2.8)\\ &\quad -\sum_{l\geq 1} K(l,k)c_l c_k + \sum_{l\geq 1} K(k+1,l-1)c_{k+1}c_{l-1}\ , \qquad \text{for } k\geq 0\ , && (2.9) \end{aligned}$$

where c_k is the fraction of sites with k particles. The derivation of the above equation from the stochastic particle system was made rigorous in [29]. Recently, EDG gained some interest with results regarding well-posedness [20] and long-time behavior [21, 54], see also Chapter 3. Motivated by the singular LSW equation (2.6) and the model from Chapter 1, one natural candidate for the the rate kernel is

$$K(k,l) = \begin{cases} k^{-\beta}, & \text{if } k,l \neq 0,\\ 0, & \text{else.} \end{cases}$$

This kernel is homogeneous in k and due to $K(k,0) = 0$, empty sites are virtually removed from the system, which leads to the same coarsening mechanism as in the LSW model. Moreover, due to the negative power β, particles tend to aggregate which promotes coarsening even more. For this kernel the master equation (2.8) reads

$$\begin{aligned} \dot{c}_0 &= (1-c_0)c_1,\\ \dot{c}_1 &= -\left((1-c_0) + \sum_{l=1}^{\infty} l^{-\beta}\right) c_1 + (1-c_0)2^{-\beta}c_2,\\ \dot{c}_k &= \left(\sum_{l=1}^{\infty} l^{-\beta}\right) c_{k-1} - \left((1-c_0)k^{-\beta} + \sum_{l=1}^{\infty} l^{-\beta}\right) c_k + (1-c_0)(k+1)^{-\beta}c_{k+1},\ k\geq 2. \end{aligned}$$

Here, the factor $1-c_0 = \sum_{l=1}^{\infty} c_l$ represents the number of non-empty sites. It appears on the right-hand side because in the stochastic limit the particle interaction is normalized with respect to all sites in the graph and not only the non-empty sites. Since we want empty sites to not affect the system anymore, one possibility is to change the aforementioned normalization. Another possibility is to simply absorb the factor $(1-c_0)$ in a time change, which leads to the same modified master equation:

$$\begin{aligned} \dot{c}_0 &= c_1,\\ \dot{c}_1 &= -\left(1+\theta[c]\right) c_1 + 2^{-\beta}c_2,\\ \dot{c}_k &= \theta[c]c_{k-1} - \left(k^{-\beta} + \theta[c]\right) c_k + (k+1)^{-\beta}c_{k+1},\ k\geq 2, \end{aligned}$$

where

$$\theta[c] = \frac{1}{1-c_0}\sum_{l\geq 1} l^{-\beta}c_l = \frac{\sum_{l\geq 1} l^{-\beta}c_l}{\sum_{l\geq 1} c_l}$$

can be interpreted as the average jump rate of the non-empty sites. We further set $u(k) = c_k$ for $k \geq 1$ and $u(0) = 0$ to obtain

$$\begin{cases} \partial_t u = \partial^+(k^{-\beta}u) - \partial^-\left(\theta[u]u\right), \\ u(t,0) = 0, \end{cases} \tag{2.10}$$

where $\partial^+ u(k) = u(k+1) - u(k)$, $\partial^- u(k) = u(k) - u(k-1)$ are the forward and backward discrete difference operators. We see that equation (2.10) has a very similar formal structure to equation (2.6), so the latter can be interpreted as a continuum version of the EDG model for this specific choice of the rate Kernel. For the Becker-Döring model [3] which is similar to EDG, a connection to LSW is well known [49, 55]. However, in this case the LSW equation only describes the macroscopic evolution of large clusters in a system above critical density. Microscopically, the bulk of the system relaxes to an equilibrium state with critical density, where the macroscopic vanishing of a particle corresponds to the microscopic cluster becoming part of the bulk equilibrium. In our case there is no non-trivial microscopic equilibrium and vanishing happens already on the atomic level. At the moment we do not know if solutions to (2.10) behave like solutions to (2.6) for large times, something that is true in the case of the symmetric product kernel with vanishing particles (see Chapter 3). Since equation (2.10) has a more diffusive nature than the LSW equation we conjecture that the long-time behavior of (2.10) is not described by (2.6) but some regularized version with a second order term. This is is also possible future work.

2.1.2. Results

Before stating our main result, we define the notion of solutions for the LSW equation (2.6) in terms of the weak formulation for general measures. Apart from the theoretical advantage of working in an abstract framework, it is also physically reasonable for a system to contain macroscopic fractions of particles with the same size, hence the size-distribution can no longer be described by a density function but a more general measure that may contain Diracs.

For spatial variables we use the notation $\mathbb{R}_+ = (0, \infty)$, $\overline{\mathbb{R}}_+ = [0, \infty)$. Then $\mathcal{P}_1(\overline{\mathbb{R}}_+)$ denotes the space of probability measures on $\overline{\mathbb{R}}_+$ with finite first moment. In accordance with [52] we view $\mathcal{P}_1(\overline{\mathbb{R}}_+)$ as a subset of $C_c(\overline{\mathbb{R}}_+)^*$ endowed with the weak-$*$ topology. Another natural choice of topology is the one induced by the L^1-Wasserstein metric (see next section). Testing equation (2.6) with a smooth function and integrating in space and time, we obtain a suitable weak formulation of the equation:

Definition 2.1 (Weak solutions)**.** *Let $T > 0$. A weak-$*$ continuous map $\mu\colon [0,T) \to \mathcal{P}_1(\overline{\mathbb{R}}_+)$ is a solution to the LSW equation (2.6) with mean-field $\theta \in L^1_{loc}([0,T))$ if the following conditions are satisfied:*

1. *For every $t \in [0,T)$ it holds*

$$\int_{\mathbb{R}_+} x \, \mathrm{d}\mu_t = \int_{\mathbb{R}_+} x \, \mathrm{d}\mu_0.$$

2. For every $\psi \in C_c^\infty((0,T)\times\mathbb{R}_+)$ it holds

$$\int_0^T \int_{\overline{\mathbb{R}}_+} \partial_t\psi + (\theta - x^{-\beta})\partial_x\psi \,\mathrm{d}\mu_t\,\mathrm{d}t = 0.$$

If not explicitly specified, we always assume $T = \infty$.

Remark 2.2. *There are some subtleties in the weak formulation of the LSW equation that we briefly want to address:*

- *Instead of explicitly defining the mean-field as in (2.6), the function θ is determined by the first condition, i.e the conservation of the first moment. The advantage of this formulation is that it is more stable when considering sequences of solutions. Formally we recover the formula (2.6) by using the test functions $\psi(t,x) = \eta(t)x$, $\eta \in C_c^\infty((0,T))$, which can be made rigorous by approximation.*
- *The fact that we only consider test functions vanishing at $x = 0$ reflects the fact that equation (2.6) does not conserve the total measure on $\mathbb{R}_+$. Thus the condition that the solution is a probability measure at all times is merely a normalization (which will be useful nonetheless) that can always be satisfied by placing a Dirac measure at the origin that compensates the loss of measure.*
- *It is clear from the definition that for all $x_0 > 0$, the constant measure $\mu_t = \delta_{x_0}$ is a solution with mean field $\theta = x_0^{-\beta}$. Furthermore, if μ_t conserves the first moment and satisfies the weak formulation but $\mu_t(\overline{\mathbb{R}}_+) = \lambda(t) < 1$, then $\tilde{\mu}_t = \mu_t + (1-\lambda(t))\delta_0$ is a solution in the sense of the definition (see comment above).*

Next we note that if $x_1(t), ..., x_N(t)$ are functions satisfying the discrete system of equations (2.5), then the corresponding empirical measure

$$\mu_t^N = \frac{1}{N}\sum_{j=1}^N \delta_{x_j(t)}$$

is a measure solution to the LSW equation in the sense of our definition. Hence for initial data in the class of empirical measures, the well-posedness of the equation reduces to the study of ODE. For general initial data, the natural approach is to approximate by empirical measures. Let d_1 denote the L^1-Wasserstein metric on probability measures. Then we have the following result:

Theorem 2.3. *Let μ^N be a sequence of empirical measure solutions to equation (2.6) with mean-fields θ^N and $\mu_0 \in \mathcal{P}_1(\overline{\mathbb{R}}_+)$ such that $d_1(\mu_0^N, \mu_0) \to 0$, $N \to \infty$. Then there exists a solution μ_t to equation (2.6) with initial data μ_0, mean-field $\theta \in L^p_{loc}([0,\infty))$, $p \in (1, 1+\beta^{-1})$ and a subsequence $N \to \infty$ such that for every $T > 0$ we have:*

$$\sup_{t\in[0,T]} d_1(\mu_t^N, \mu_t) \to 0, \quad \theta^N \rightharpoonup \theta \text{ in } L^p((0,T)), \text{ as } N \to \infty.$$

Because every measure can be approximated by empirical measures we conclude the following existence result:

Corollary 2.4. *For every $\mu_0 \in \mathcal{P}_1(\overline{\mathbb{R}}_+)$ there exists a solution μ_t to equation (2.6) with initial data μ_0 and mean-field $\theta \in L^p_{loc}([0,\infty))$, $p \in (1, 1+\beta^{-1})$.*

Unfortunately, it is not clear at the moment whether solutions are unique. Due to the singularity of $x^{-\beta}$, the standard approaches to prove uniqueness like in [52] fail. Interestingly, it was shown in [33] that solutions to the local model (2.7) are not unique. However, the proof makes heavy use of the local structure of the model, so it is not clear at the moment whether uniqueness might also fail for the mean-field equation (2.6).

Another natural question is for which initial data there exists a solution where $\theta \in L^\infty_{loc}([0,\infty))$. It is easy to see that in the discrete system (2.5), the mean-field blows up whenever a particle vanishes, hence if the initial data have atoms then the mean-field will in general not be bounded. The natural conjecture is that a solution with bounded mean-field exists for measures with a density function and finite initial mean-field. We also suspect that at least solutions with bounded mean-field are unique and plan to continue research in this direction in future work.

Next we are interested in the existence of special solutions that have a self-similar form. A dimensional analysis of equation (2.6) suggests that $x \sim t^{\frac{1}{\beta+1}}$. This scaling law together with the conservation of the first moment leads to the following ansatz for a solution:

$$f(t,x) = t^{-\frac{2}{1+\beta}} \Phi(t^{-\frac{1}{1+\beta}} x),$$

where $\Phi = \Phi(z)$ needs to be determined. Using this ansatz in equation (2.6) leads to a differential equation for Φ:

$$\frac{2}{1+\beta}\Phi(z) + \frac{1}{1+\beta} z\Phi'(z) = \beta z^{-\beta-1}\Phi(z) - (z^{-\beta} - \theta[\Phi])\Phi'(z), \tag{2.11}$$

with $\theta[\Phi]$ as in (2.6). It turns out that the solution is uniquely characterized by the value $\theta[\Phi]$ and a continuum of values is admissible:

Proposition 2.5. *There exist constants $\theta_{min}, \theta_{max}$ depending on β, such that the following holds: For every $\vartheta \in [\theta_{min}, \theta_{max})$ there exists a (up to normalization) unique self-similar profile Φ with $\theta[\Phi] = \vartheta$ and compact support $[0, z_*]$, where z_* depends on ϑ, β. Furthermore,*

$$\Phi(z) \sim c_1 z^\beta, \text{ as } z \to 0^+,$$
$$\Phi(z) \sim \begin{cases} c_2(z_* - z)^r, \text{ as } z \to z_*^- \text{ if } \vartheta < \theta_{max}, \\ c_2 e^{-\frac{R}{z_*-z}}, \text{ as } z \to z_*^- \text{ if } \vartheta = \theta_{min}, \end{cases}$$

where $c_1, c_2, R, r > 0$ depend on β, r with $r \to \infty$ as $\vartheta \to \theta_{min}^+$ and $r \to 0$ as $\vartheta \to \theta_{max}$. These are the only possible self-similar profiles.

For other values of $\theta[\Phi]$, the solution does either not exist globally in space or has infinite first moment, which is not permitted in our definition of solutions. We obtain a one-parameter family of self-similar solutions to equation (2.6) and only at the critical parameter $\vartheta = \theta_{min}$ the solution is smooth. This observation makes the question for the long-time behavior of equation (2.6) quite delicate, since for generic initial data it is not clear which profile is selected for large times, if at all. We expect it is possible to perform an analysis similar to [53], where the authors demonstrate that for the classical LSW equation, there is a dense set of initial data that do not converge to any self-similar solution.

Outline

The rest of this chapter is organized as follows. In Section 2 we give a precise description of our framework. The key step is to formulate the LSW equation (2.6) in terms of the size-ranking function $x(t, \cdot)$ associated with the probability measure μ_t like in [52]. The empirical measure approximation then corresponds to the approximation of x by step functions whose values evolve according to the discrete LSW equation (2.5), for which we introduce some basic properties and notation.

In Section 3 we give the full proof of Theorem 2.3. The main difficulty lies in the fact that the mean-field θ for the discrete system does not stay bounded but blows up whenever a particle vanishes, in contrast to the situation in [52]. To obtain L^p estimates for θ we therefore have to characterize the rate at which particles vanish, which is done by a careful inductive argument which only works for a small volume fraction at the tail end of the particle distribution, see Proposition 2.12. To extend the estimates to all particles we also prove a vanishing property for local averages in Proposition 2.16. Then we can derive local L^p estimates for $x_j^{-\beta}$ (see Lemma 2.20) that can be extended to arbitrary time intervals and yield the desired compactness in Proposition 2.19. We then pass to the limit in Proposition 2.23. Here, an important step is to prove that the vanishing times of approximating solutions converge to the vanishing time of the limiting function to deal with the discontinuous function $\mathbf{1}_{\{x_j>0\}}$ on the right-hand side, after which the limit is carried out similarly to [52, Proposition 6.1].

Finally, in Section 4 we study the self-similar solutions to equation (2.6) and prove Proposition 2.5. The ODE (2.11) can be analyzed by relatively elementary techniques. Here the key observation is that the nonlocal term $\theta[\Phi]$ can be freely chosen, after which the equation becomes a simple first order ODE.

2.2. Setup and basic properties

2.2.1. Size rankings

In the class of empirical measures, equation (2.6) for the measure $\mu_t^N = \frac{1}{N}\sum_{j=1}^N \delta_{x_j(t)}$ directly translates to the system of ODE (2.5) for the sizes x_j. Size ranking functions are the appropriate objects to extend this correspondence naturally to arbitrary measures.

We give a brief overview about size-rankings, and most statements of this subsection can be found in [52] and the references therein, which we refer to for a more exhaustive treatment of this topic.

Definition 2.6. *We define the Banach space L^1_d as the following closed subspace of $L^1((0,1))$:*

$$L^1_d = \{x\colon (0,1] \to \mathbb{R}\colon x \in L^1((0,1)), x(1) = 0,\ x \text{ is decreasing and right continuous}\}.$$

Then we define the map $x\colon \mathcal{P}_1(\overline{\mathbb{R}}_+) \to L^1_d$ by

$$x[\mu](\varphi) = \sup\{y\colon \mu([y,\infty)) > \varphi\}. \tag{2.12}$$

The function $x[\mu]$ is called the size ranking associated to μ and for convenience we write $x[\mu](\varphi) = x(\varphi)$ if there is no danger of confusion.

In other words, the size-ranking x associated to μ is the right-continuous inverse of the tail distribution $F_\mu(y) = \mu([y,\infty))$. Furthermore, $x(\varphi)$ is characterized by the identity

$$\int_{\mathbb{R}_+} g \,\mathrm{d}\mu = \int_0^1 g(x(\varphi)) \,\mathrm{d}\varphi, \tag{2.13}$$

which holds for all continuous functions $g\colon \mathbb{R}_+ \to \mathbb{R}$ with compact support, and by approximation also for a wider class of functions. In particular we have

$$\int_{\mathbb{R}_+} x \,\mathrm{d}\mu = \int_0^1 x(\varphi) \,\mathrm{d}\varphi, \quad \int_{\mathbb{R}_+} x^{-\beta} \,\mathrm{d}\mu = \int_0^1 x(\varphi)^{-\beta} \cdot \mathbf{1}_{\{x(\varphi)>0\}} \,\mathrm{d}\varphi,$$

where both sides of the equations can be infinity in the second identity. The usefulness of the size ranking map lies in its connection to the Wasserstein metric d_1, which is defined by

$$d_1(\mu,\nu) = \inf_{\pi \in C(\mu,\nu)} \int_{\mathbb{R}_+^2} |x-y| \,\mathrm{d}\pi(x,y),$$

where $C(\mu,\nu)$ is the set of couplings of μ and ν. This metric is well-defined on $\mathcal{P}_1(\overline{\mathbb{R}}_+)$ and the following holds:

Proposition 2.7. *The metric space $(\mathcal{P}_1(\overline{\mathbb{R}}_+), d_1)$ is complete and a sequence μ_n converges to μ with respect to d_1 if and only if μ_n converges weakly-$*$ to μ and*

$$\int_{\overline{\mathbb{R}}_+} x \,\mathrm{d}\mu_n \to \int_{\overline{\mathbb{R}}_+} x \,\mathrm{d}\mu.$$

Furthermore, the following key result holds (cf. [52, Lemma 4.1 & 4.2]):

Proposition 2.8. *The map $x\colon \mathcal{P}_1(\overline{\mathbb{R}}_+) \to L^1_d$ is a bijection. Furthermore, for $\mu,\nu \in \mathcal{P}_1(\overline{\mathbb{R}}_+)$ it holds*

$$d_1(\mu,\nu) = \|x[\mu] - x[\nu]\|_{L^1}.$$

In the case of empirical measures $\mu = \frac{1}{N}\sum_{j=1}^{N}\delta_{x_j}$ with $x_1 \geq x_2 \geq ... \geq x_N$ the size ranking is simply given by

$$x(\varphi) = \sum_{j=1}^{N} x_j \cdot \mathbf{1}_{[\frac{j-1}{N},\frac{j}{N})}(\varphi), \tag{2.14}$$

which can easily be seen by computing

$$\int_{\mathbb{R}_+} g \,\mathrm{d}\mu = \frac{1}{N}\sum_{j=1}^{N} g(x_j) = \int_0^1 g\left(\sum_{j=1}^{N} x_j \cdot \mathbf{1}_{[\frac{j-1}{N},\frac{j}{N})}(\varphi)\right) \mathrm{d}\varphi.$$

Note that the size ranking only depends on the measure itself. If the sizes x_j were unordered, taking the size ranking is basically the operation of ordering the tuple $x_1, ..., x_N$, hence the name. Now if we express the discrete LSW equation (2.5) in terms of the size ranking for the corresponding empirical measure, we get

$$\dot{x}(t,\varphi) = \left(\theta(t) - x(t,\varphi)^{-\beta}\right) \cdot \mathbf{1}_{\{x(t,\varphi)>0\}}, \quad \theta(t) = \frac{\int_0^1 x(t,\varphi)^{-\beta} \cdot \mathbf{1}_{\{x(t,\varphi)>0\}} \,\mathrm{d}\varphi}{\int_0^1 \mathbf{1}_{\{x(t,\varphi)>0\}} \,\mathrm{d}\varphi}. \tag{2.15}$$

This formulation of the LSW equation can be directly generalized to arbitrary functions with values in L^1_d. As usual, it is more practical to define the mean-field implicitly through the mass constraint:

Definition 2.9. *Let $T > 0$. Then a pair of functions $x = x(t,\varphi)$, $\theta = \theta(t)$ is a solution to the LSW equation on $[0,T)$ if the following relations are satisfied:*

1. *We have $x \in C^0([0,T]; L^1_d)$, $\theta \in L^1((0,T))$ and $x(\cdot,\varphi)^{-\beta} \cdot \mathbf{1}_{\{x(\cdot,\varphi)>0\}} \in L^1((0,T))$ for all $\varphi \in (0,1)$.*
2. *For all $t \in [0,T]$ we have*

$$\int_0^1 x(t,\varphi) \;\mathrm{d}\varphi = \int_0^1 x(0,\varphi) \;\mathrm{d}\varphi. \tag{2.16}$$

3. *For all $t \in [0,T]$ and $\varphi \in (0,1)$ it holds*

$$x(t,\varphi) = x(0,\varphi) + \int_0^t \left(\theta(s) - x(s,\varphi)^{-\beta}\right) \cdot \mathbf{1}_{\{x(s,\varphi)>0\}} \;\mathrm{d}s. \tag{2.17}$$

For brevity we refer to solutions in the above sense simply as solutions to equation (2.17). The crucial result for this chapter is the fact that a curve of measures μ_t is a solution to the LSW equation (2.6) if the associated size rankings $x(t,\cdot)$ solve (2.17):

Proposition 2.10. *Let $\mu = \mu_t\colon [0,T) \to \mathcal{P}_1(\overline{\mathbb{R}}_+)$ be a curve of measures. If the associated size rankings $x(t,\cdot)$ solve equation (2.17) with mean-field θ, then μ_t is a solution to the LSW equation with mean-field θ on $[0,T)$.*

Since the proof is rather short we present it here. For a more detailed description of the equivalence between measures and size-rankings we refer to [52, Proposition 4.3].

Proof. Let $\psi \in C_c^\infty((0,T)\times\mathbb{R}_+)$. Since θ and $x(\cdot,\varphi)^{-\beta}\cdot\mathbf{1}_{\{x(\cdot,\varphi)>0\}}$ are in $L^1((0,T))$, equation (2.17) implies that $x(\cdot,\varphi)\in W^{1,1}((0,T))$ for every $\varphi\in(0,1)$. This implies that $t\mapsto\psi(t,x(t,\varphi))$ is weakly differentiable with

$$\begin{aligned}\frac{d}{dt}\psi(t,x(t,\varphi)) &= \partial_t\psi(t,x(t,\varphi)) + \dot{x}(t,\varphi)\partial_x\psi(t,x(t,\varphi))\\ &= \partial_t\psi(t,x(t,\varphi)) + \Big(\theta(t)-x(t,\varphi)^{-\beta}\Big)\cdot\mathbf{1}_{\{x(t,\varphi)>0\}}\partial_x\psi(t,x(t,\varphi))\\ &= \partial_t\psi(t,x(t,\varphi)) + \Big(\theta(t)-x(t,\varphi)^{-\beta}\Big)\,\partial_x\psi(t,x(t,\varphi)),\end{aligned}$$

since ψ has compact support. We integrate the above identity in φ and t, observing that the left-hand-side vanishes after integration in time:

$$\begin{aligned}0 &= \int_0^T\int_0^1 \partial_t\psi(t,x(t,\varphi)) + \Big(\theta(t)-x(t,\varphi)^{-\beta}\Big)\,\partial_x\psi(t,x(t,\varphi))\,\mathrm{d}\varphi\,\mathrm{d}t\\ &= \int_0^T\int_{\mathbb{R}_+}\partial_t\psi + (\theta - x^{-\beta})\partial_x\psi\,\mathrm{d}\mu_t\,\mathrm{d}t,\end{aligned}$$

where the last equality follows from (2.13). This also implies

$$\int_{\mathbb{R}_+} x\,\mathrm{d}\mu_t = \int_0^1 x(t,\varphi)\,\mathrm{d}\varphi = \int_0^1 x(0,\varphi)\,\mathrm{d}\varphi = \int_{\mathbb{R}_+} x\,\mathrm{d}\mu_0.$$

At last, the weak-$*$ continuity of μ_t follows from $x\in C^0([0,T);L^1_d)$, Proposition 2.7 and Proposition 2.8. □

2.2.2. Mean-field ODE for finitely many particles

In this subsection we consider the discrete LSW system (2.5), which will become important for our approximation of general solutions. First we discuss the existence of solutions for equation (2.5). By definition, particles with size zero don't contribute to the evolution, hence we assume without loss of generality that the initial particle sizes are all strictly positive. Then the equation becomes

$$\dot{x}_j = \theta(t) - x_j^{-\beta},\ j=1,...,N,\quad \theta(t) = \frac{1}{N}\sum_{j=1}^N x_j(t)^{-\beta}. \tag{2.18}$$

Since this system is a finite-dimensional ODE, a unique local solution exists up to some time τ, where the solution leaves the domain of definition. In our case this happens if a particle reaches size 0. Before the time τ the ordering of particle masses is conserved during the evolution, i.e $x_i(0)\geq x_j(0)$ implies $x_i(t)\geq x_j(t)$ for all $t\geq 0$. This is clear because as long as $x_i, x_j > 0$ we have

$$\frac{d}{dt}(x_i - x_j) = x_j^{-\beta} - x_i^{-\beta}.$$

Thus from now on we assume that $x_1 \geq x_2 \geq ... \geq x_N > 0$, which also avoids ambiguity when we idenitfy $x_1, ..., x_N$ with a size ranking function (see previous subsection). This also means that τ is the time when the smallest particle reaches size 0. When this happens, we can restart the evolution by considering equation (2.18) with $N-1$ instead of N (or $N-l$ instead of N if l particles vanish at the same time). This can be continued until potentially only one particle is left, at which point the solution becomes stationary. In any case, we get a solution for all times. This procedure can be summarized in the following system:

$$\begin{aligned} \dot{x}_j &= \theta(t) - x_j^{-\beta}, \ j = 1, ..., \lambda(t), \quad \theta(t) = \frac{1}{\lambda(t)} \sum_{j=1}^{\lambda(t)} x_j(t)^{-\beta}, \\ \lambda(t) &= \max\{j \in \{1, ..., N\} \colon x_j(t) > 0\}, \\ x_j(t) &= 0, \ j = \lambda(t) + 1, ..., N, \end{aligned} \tag{2.19}$$

which we will consider as the discrete LSW system from now on. It is easy to check that discrete solutions that are constructed in this manner are also solutions in the sense of equation (2.17) after identifying $x_1, ..., x_N$ with a size ranking. However, before the approximation procedure, the viewpoint (2.19) is more intuitive.

Next we establish some basic properties of the system (2.19). Inserting the trivial estimate $\theta \geq 0$ in the equation for $\dot{x}_j$ we get the differential inequality

$$\dot{x}_j \geq -x_j^{-\beta},$$

which can be integrated to obtain the lower bound

$$x_j(t_2) \geq \left(x_j(t_1)^{\beta+1} - (\beta+1)(t_2 - t_1)\right)^{\frac{1}{\beta+1}}, \tag{2.20}$$

whenever $0 \leq t_1 < t_2$ and $x_j(t_2) > 0$. This estimate has two important implications. The first one is a half-life estimate:

$$x_j(t) \geq \frac{1}{2} x_j(t_1), \qquad \text{if } t \leq t_1 + \frac{x_j(t_1)^{\beta+1}}{2(\beta+1)}. \tag{2.21}$$

Secondly, if the particle x_j vanishes at time τ, then setting $t_2 = \tau$ in (2.20) we get

$$x_j(t) \leq (\beta+1)^{\frac{1}{\beta+1}} (\tau - t)^{\frac{1}{\beta+1}}. \tag{2.22}$$

In the following we denote by τ_j the vanishing time of the particle j, i.e

$$\tau_j = \inf\{t \geq 0 \colon x_j(t) = 0\}. \tag{2.23}$$

Note that τ_j can be ∞ and by construction we have $x_j(t) > 0$ for $t < \tau_j$ and $x_j(t) = 0$ for $t \geq \tau_j$.

Next we consider the evolution of the volume fraction of "large" particles. To be precise, let $d > 0$ and denote

$$l_d(t) = \max\{l \in \{1, ..., N\} : x_l(t) \geq d\}. \tag{2.24}$$

For the analysis it turns out to be crucial that there is a macroscopic fraction of large particles at all times. In mathematical terms this is expressed as $l_d(t) \geq \kappa N$ for some $\kappa > 0$. One way to estimate $l_d(t)$ from below is to use the half-life estimate (2.21). This directly implies

$$l_{\frac{d}{2}}(t) \geq l_d(0), \quad \text{if } t \leq T^* := \frac{d^{\beta+1}}{2(\beta+1)}. \tag{2.25}$$

Hence if initially there are many large particles, the same will be true at least until T^* if we compromise by making the threshold d smaller. It is interesting to note that the analysis so far is the same as for the local equation (2.7) since it only relies on the lower bound $\dot{x}_j \geq x_j^{-\beta}$ that holds in both the mean-field and the local model. So far the threshold d is generic. We have another natural way of estimating l_d if we choose d in a way that is related to the initial average particle size

$$\rho = \frac{1}{N}\sum_{i=1}^{N} x_i(0). \tag{2.26}$$

In fact we have the following result:

Lemma 2.11. *Let $x_1, ..., x_N$ be a solution to equation (2.5) on the interval $[0, T]$ with $\rho > 0$. Then for all $R > 0$ we have*

$$l_{\frac{\rho}{2}}(t) \geq \frac{N}{R}\left(\frac{\rho}{2} - \frac{1}{N}\sum_{i=1}^{N}(x_i(t) - R)_+\right). \tag{2.27}$$

Proof. By definition of ρ and conservation of mass we have

$$\begin{aligned} N\rho = \sum_{i=1}^{N} x_i(0) = \sum_{i=1}^{N} x_i(t) &= \sum_{i \leq l_{\frac{\rho}{2}}(t)} x_i(t) + \sum_{i > l_{\frac{\rho}{2}}(t)} x_i(t) \\ &\leq l_{\frac{\rho}{2}}(t)R + \sum_{i \leq l_{\frac{\rho}{2}}(t)} (x_i(t) - R)_+ + (N - l_{\frac{\rho}{2}}(t))\frac{\rho}{2} \\ &\leq l_{\frac{\rho}{2}}(t)R + \sum_{i=1}^{N}(x_i(t) - R)_+ + N\frac{\rho}{2}. \end{aligned}$$

Rearranging the terms yields the desired inequality. □

Note that in terms of the size ranking $x(\varphi)$ associated with $x_1, ..., x_N$ we have

$$\rho = \int_0^1 x(0, \varphi)\, \mathrm{d}\varphi, \quad \frac{1}{N}\sum_{i=1}^{N}(x_i(t) - R)_+ = \int_0^1 (x(t, \varphi) - R)_+\, \mathrm{d}\varphi.$$

Then the above Lemma illustrates the fact that only scenario where a macroscopic lower bound on $\ell_{\frac{\rho}{2}}$ might be violated is if $x(t, \varphi)$ concentrates at the origin in finite time, a phenomenon known as gelation. In the upcoming sections we will prove that gelation cannot occur, which is one of the main a priori estimates for the discrete system.

2.3. Proof of Theorem 2.3

2.3.1. Vanishing behavior of small particles

The goal of this subsection is to show that small enough particles indeed vanish if there is a macroscopic fraction of large particles. Furthermore, to control $x_j^{-\beta}$ in a suitable L^p-space, not only the upper bound (2.22), but also a corresponding lower bound needs to hold. Throughout this section, $x_1, ..., x_N$ will always denote a solution to equation (2.19). Also recall the definition of the large particle index l_d as in (2.24). We will prove the following result:

Proposition 2.12. *Let $T > 0$, $d > 0$, $\kappa > 0$ such that $l_d \geq \kappa N$ on $[0, T]$. Then for every $t_0 \in (0, T)$ there exist $\delta > 0$, $\eta > 0$ and $c > 0$ such that the following holds: If $j \geq (1 - \delta)\lambda(s)$ and $0 < x_j(s) \leq \eta$ for some $s \in [0, t_0]$, then we have $\tau_j \leq T$ and*

$$x_j(t) \geq c(\tau_j - t)^{\frac{1}{\beta+1}}, \text{ for } s \leq t \leq \tau_j. \tag{2.28}$$

The constants δ, η, c depend on β, d, κ, T, t_0.

The above statement can easily be verified for the smallest particle:

Lemma 2.13. *Let $T > 0$, $d > 0$, $\kappa > 0$ such that $l_d \geq \kappa N$ on $[0, T]$. Then for every $t_0 \in (0, T)$ there exists $\eta > 0$ depending on β, d, κ, T, t_0 such that if $\lambda(s) = n$ and $x_n(s) \leq \eta$ for some $s \in [0, t_0]$, then $\tau_n \leq T$ and*

$$x_n(t) \geq c(\beta)\,(\tau_n - t)^{\frac{1}{\beta+1}}\,, \text{ for } s \leq t \leq \tau_n. \tag{2.29}$$

Proof. By definition $x_n(s)$ is the smallest non-zero particle size at time s. We use (2.5) and calculate

$$\dot{x}_n = n^{-1}\sum_{l=1}^{n} x_l^{-\beta} - x_n^{-\beta} = n^{-1}\left(\sum_{l \leq l_d} x_l^{-\beta} + \sum_{l > l_d} x_l^{-\beta}\right) - x_n^{-\beta}.$$

For particles in the first sum we use $x_k \geq d$, while for the rest of the particles we insert the estimate $x_l \geq x_n$, which yields

$$\begin{aligned} \dot{x}_n &\leq d^{-\beta} + \left(n^{-1}(n - l_d(0)) - 1\right) x_n^{-\beta} \\ &= d^{-\beta} - n^{-1} l_d x_n^{-\beta} \leq d^{-\beta} - \kappa x_n^{-\beta}. \end{aligned}$$

Thus for small enough η, $x_n(s) \leq \eta$ implies $d^{-\beta} \leq \frac{1}{2}\kappa x_n(s)^{-\beta}$, and thus

$$\dot{x}_n \leq -\frac{1}{2}\kappa x_n^{-\beta},$$

at time s and consequently also for $t \geq s$ since x decreases. Multiplying with x_n^{β} and integrating this differential inequality from t_1 to t_2 for $s \leq t_1 < t_2$ we obtain

$$x_n(t_2) \leq \left(x_n(t_1)^{\beta+1} - \frac{1}{2}(\beta + 1)(t_2 - t_1)\right)^{\frac{1}{\beta+1}}.$$

Setting $t_1 = s \le t_0$, we see that x_n vanishes before time T if $x_n(s)^{\beta+1} \le \frac{1}{2}(\beta+1)(T-t_0)$, which is satisfied after possibly making η smaller. Then setting $t_2 = \tau_j$ in the above inequality yields

$$x_n(t)^{\beta+1} \ge \frac{1}{2}(\beta+1)(\tau - t), \;\; s \le t \le \tau_j,$$

which gives the desired statement. □

The above Lemma is the prototypical result regarding vanishing of small particles. The key calculation here was to use the lower bound on l_d to replace the mean-field on the right-hand-side of (2.19) by a constant. The difficulty to extend this argument to arbitrary particles instead of only the smallest one lies in the fact that the mean-field blows up whenever a particle vanishes and so there is no hope to estimate the mean-field term by some constant. Instead we have to work with the residual mean-field which is defined as

$$\theta_j(t) = \frac{1}{\lambda(t)} \sum_{l=j+1}^{\lambda(t)} x_l(t)^{-\beta}. \tag{2.30}$$

Lemma 2.14. *Let $T > 0$, $d > 0$, $\kappa > 0$ such that $l_d \ge \kappa N$ on $[0,T]$. Then the following inequalities hold for all $j > \kappa N$ and $t \in [0, \tau_j)$:*

$$\dot{x}_j(t) \le \theta_j(t) + d^{-\beta} - \kappa x_j(t)^{-\beta}, \tag{2.31}$$
$$\dot{x}_j(t) \le \theta_j(t). \tag{2.32}$$

Proof. We rewrite the right-hand-side of (2.19) as

$$\dot{x}_j(t) = \frac{1}{\lambda(t)} \left(\sum_{l=1}^{l_d(t)} x_l(s)^{-\beta} + \sum_{l=l_d(t)+1}^{j} x_l(s)^{-\beta} + \sum_{l=j+1}^{\lambda(s)} x_l(s)^{-\beta} \right) - x_j(s)^{-\beta}.$$

On the first sum in the bracket we use $x_l \ge d$, while for the second sum we insert $x_l \ge x_j$. This yields

$$\frac{1}{\lambda(t)} \sum_{l=1}^{l_d(t)} x_l(t)^{-\beta} \le d^{-\beta},$$
$$\frac{1}{\lambda(t)} \sum_{l=l_d(0)+1}^{j} x_l(t)^{-\beta} \le \frac{j - l_d(0)}{\lambda(t)} x_j(t)^{-\beta} \le \left(1 - \frac{l_d(0)}{\lambda(t)}\right) x_j(t)^{-\beta} \le (1-\kappa) x_j(t)^{-\beta},$$

and hence inequality (2.31) follows. For the second inequality, we simply use the estimate $x_l \ge x_j$ also in the first sum. Then

$$\frac{1}{\lambda(t)} \sum_{l=1}^{j} x_l(t)^{-\beta} \le x_j(t)^{-\beta},$$

and (2.32) follows. □

Next we want to use the above Lemma to extend the estimate (2.29) to arbitrary particle indices with the crucial property that the constants in the estimates should not depend on the particle index or the total number of particles, since we want to consider the limit $N \to \infty$ later on. Note that particles will never continuously shrink unless they are smallest in size due to the blow-up of the mean-field. To account for this growth effect of vanishing particles, we can only consider a particles whose index j is not too far away from $\lambda(t)$. We then proceed inductively, with the following Lemma as the inductive step:

Lemma 2.15. *Let $T > 0$, $d > 0$, $\kappa > 0$ such that $l_d \geq \kappa N$ on $[0, T]$. Then for every $t_0 \in (0, T)$ and $c \leq \frac{\kappa}{4}(\beta+1)^{\frac{1}{\beta+1}}$ there exist $\delta > 0$, $\eta > 0$ such that the following holds: If $j \geq (1-\delta)\lambda(s)$, $0 < x_j(s) \leq \eta$ for some $s \in [0, t_0]$ and for all $l \in \{j+1, .., N\}$ with $\tau_l > s$ we have $\tau_l \leq T$ and*

$$x_l(t) \geq c(\tau_l - t)^{\frac{1}{\beta+1}}, \text{ for } s \leq t \leq \tau_l,$$

then we have $\tau_j \leq T$ and

$$x_j(t) \geq c(\tau_j - t)^{\frac{1}{\beta+1}}, \text{ for } s \leq t \leq \tau_j.$$

The constants δ, η depend on $\beta, d, \kappa, T, t_0, c$.

Proof. Recall the definition of the residual mean-field $\theta_j = \lambda^{-1} \sum_{l=j+1}^{\lambda} x_l^{-\beta}$. First we integrate θ_j from t_1 to t_2 for $s \leq t_1 < t_2 \leq T$, using $\lambda(t) \geq l_d(t) \geq \kappa N$, $j \geq (1-\delta)\lambda(t)$ and the estimates for x_l that hold by assumption:

$$\begin{aligned}\int_{t_1}^{t_2} \theta_j(t)\,\mathrm{d}t &\leq \frac{1}{\kappa N} \sum_{l \in [j+1, \lambda(t)] :\, x_l(t_1) > 0} \int_{t_1}^{t_2 \wedge \tau_l} c^{-\beta}(\tau_l - t)^{-\frac{\beta}{\beta+1}}\,\mathrm{d}t \\ &= \frac{1}{\kappa N} \sum_{l \in [j+1, \lambda(t)] :\, x_l(t_1) > 0} (\beta+1)c^{-\beta}\left((\tau_l - t_1)^{\frac{1}{\beta+1}} - (\tau_l - t_2 \wedge \tau_l)^{\frac{1}{\beta+1}}\right) \\ &\leq \frac{\lambda(t) - j}{\kappa N}(\beta+1)c^{-\beta}(t_2 - t_1)^{\frac{1}{\beta+1}} \leq \frac{\delta}{\kappa}(\beta+1)c^{-\beta}(t_2 - t_1)^{\frac{1}{\beta+1}}. \end{aligned} \tag{2.33}$$

This estimate together with the inequality (2.32) implies that

$$x_j(t) \leq x_l(s) + \frac{\delta}{\kappa}(\beta+1)c^{-\beta}T^{\frac{1}{\beta+1}}, \text{ for } s \leq t \leq T,$$

hence we can δ and η small enough so we have

$$\sup_{t \in [s,T]} x_j(t) \leq \tilde{\eta}, \tag{2.34}$$

for $\tilde{\eta}$ as small as we want. In particular we can achieve the inequality

$$d^{-\beta} \leq \frac{\kappa}{2} x_j(t)^{-\beta}, \text{ for } s \leq t \leq T.$$

Using this estimate in (2.31) we obtain

$$\dot{x}_j(t) \leq \theta_j(t) - \frac{\kappa}{2} x_j(t)^{-\beta}, \text{ for } s \leq t \leq T. \tag{2.35}$$

Multiplying this inequality by x_j^β and integrating, we get

$$\begin{aligned}\frac{1}{\beta+1}\left(x_j(t_2)^{\beta+1} - x_j(t_1)^{\beta+1}\right) &\leq \int_{t_1}^{t_2} x_j(t)^\beta \theta_j(t)\,\mathrm{d}t - \frac{\kappa}{2}(t_2 - t_1) \\ &\leq \tilde{\eta}^\beta \frac{\delta}{\kappa}(\beta+1)c^{-\beta} T^{\frac{1}{\beta+1}} - \frac{\kappa}{2}(t_2 - t_1),\end{aligned}$$

where we used the estimates (2.33) and (2.34). As in the proof of Lemma 2.13 we set $t_1 = s$ and observe that we can adjust η and δ such that $\tau_j \leq T$. To conclude the lower bound for x_j we integrate inequality (2.35) from $t \geq s$ to τ_j, using the estimate (2.33) on θ_j and (2.22) to estimate $x_j^{-\beta}$ from below:

$$\begin{aligned}-x_j(t) &\leq \frac{\delta}{\kappa}(\beta+1)c^{-\beta}(\tau_j - t)^{\frac{1}{\beta+1}} - \frac{\kappa}{2}(\beta+1)^{\frac{1}{\beta+1}}(\tau_j - t)^{\frac{1}{\beta+1}} \\ &= \left(\frac{\delta}{\kappa}(\beta+1)c^{-\beta} - \frac{\kappa}{2}(\beta+1)^{\frac{1}{\beta+1}}\right)(\tau_j - t)^{\frac{1}{\beta+1}}.\end{aligned}$$

Thus the desired lower bounds hold if

$$\frac{\kappa}{2}(\beta+1)^{\frac{1}{\beta+1}} - \frac{\delta}{\kappa}(\beta+1)c^{-\beta} \geq c.$$

Thus if $c \leq \frac{\kappa}{4}(\beta+1)^{\frac{1}{\beta+1}}$, we can once more adjust δ such that the desired inequality holds. □

With this preparation we can easily prove the statement from the beginning of this subsection.

Proof of Proposition 2.12. Let $t_0 \in (0, T)$. Then the desired statement holds for the smallest particle with some constants $\eta, c > 0$ by Lemma 2.13 and thus trivially also with some possibly smaller constants that match the ones from Lemma 2.15. Starting from the smallest particle, using the statement of Lemma 2.15 inductively for all particles j with $j \geq (1-\delta)\lambda(s)$ yields the desired result. □

2.3.2. Evolution of averages

In this section we show that local averages of the particle configuration $x_1, ..., x_N$ behave similarly to single particles, which, combined with the the results of the previous subsection, turns out to be quite useful. We introduce the notation

$$[n,m] = \{n, n+1, ..., m-1, m\}, \tag{2.36}$$

$$M_{[n,m]} = \sum_{l=n}^{m} x_l, \tag{2.37}$$

$$\rho_{[n,m]} = \frac{1}{m-n+1}\sum_{l=n}^{m} x_l, \tag{2.38}$$

$$\tau_{\varepsilon,j} = \inf\{t \geq 0 \colon |[j,\lambda(t)]| < \varepsilon|[j,N]|\}. \tag{2.39}$$

Since ε is typically very small, the number $\tau_{\varepsilon,j}$ can be interpreted as the time where almost all particles (compared to the initial number) in the tail end region $[j,N]$ have vanished. We then have a result that is very similar to Lemma 2.13.

Proposition 2.16. *Let $T > 0$, $d > 0$, $\kappa > 0$ such that $l_d \geq \kappa N$ on $[0,T]$. Then for each $t_0 \in (0,T)$ and $\varepsilon > 0$ there exist positive constants η, c such that the following holds: For each j with $0 < \rho_{[j,N]}(s) \leq \eta$ for some $s \in [0,t_0]$ we have $\tau_{\varepsilon,j} \leq T$ and*

$$\rho_{[j,N]}(t) \geq \left(\rho_{[j,N]}(\tau_{\varepsilon,j})^{\beta+1} + c(\tau_{\varepsilon,j} - t)\right)^{\frac{1}{\beta+1}}, \text{ for } s \leq t \leq \tau_{\varepsilon,j}. \tag{2.40}$$

The constants η, c depend on $\beta, d, \kappa, \varepsilon, t_0, T$.

Lemma 2.17. *Let $T > 0$, $d > 0$, $\kappa > 0$ such that $l_d \geq \kappa N$ on $[0,T]$. Then for all $j \in [1,N]$ and $t \in (0,T)$ with $j < \lambda(t)$ we have*

$$\frac{d}{dt}M_{[j,N]}(t) \leq \frac{\lambda(t)-j+1}{\lambda(t)} \cdot l_d(t) \cdot \left(d^{-\beta} - (\lambda(t)-j+1)^{\beta} M_{[j,N]}(t)^{-\beta}\right). \tag{2.41}$$

Proof. We take the time derivative of $M_{[j,N]}$ and use equation (2.5) to obtain

$$\begin{aligned} \frac{d}{dt}M_{[j,N]} &= \frac{d}{dt}\sum_{l=j}^{N} x_l = \sum_{l=j}^{\lambda}\left(\frac{1}{\lambda}\sum_{m=1}^{\lambda} x_m^{-\beta} - x_l^{-\beta}\right) \\ &= \frac{\lambda-j+1}{\lambda}\sum_{m=1}^{\lambda} x_m^{-\beta} - \sum_{l=j}^{\lambda} x_l^{-\beta} \\ &= \frac{\lambda-j+1}{\lambda}\sum_{l=1}^{j-1} x_l^{-\beta} - \frac{j-1}{\lambda}\sum_{l=j}^{\lambda} x_l^{-\beta}. \end{aligned}$$

We split the first summation into the region $[1, l_d]$ and $[l_d+1, j-1]$ to obtain

$$\begin{aligned} \frac{d}{dt}M_{[j,N]} &= \frac{\lambda-j+1}{\lambda}\sum_{l=1}^{l_d} x_l^{-\beta} + \frac{\lambda-j+1}{\lambda}\sum_{l=l_d+1}^{j-1} x_l^{-\beta} - \frac{j-1}{\lambda}\sum_{l=j}^{\lambda} x_l^{-\beta} \\ &=: \mathrm{I} + \mathrm{II} - \mathrm{III}. \end{aligned}$$

In the first term, we simply apply $x_l \geq d$ to obtain

$$\mathrm{I} \leq \frac{\lambda - j + 1}{\lambda} \cdot l_d \cdot d^{-\beta}.$$

To estimate the second term, recall that the particles are ordered, hence $x_l \geq x_j$ for all indices l in the second sum. This implies

$$\mathrm{II} \leq \frac{\lambda - j + 1}{\lambda}(j - l_d - 1)x_j^{-\beta} \leq \frac{\lambda - j + 1}{\lambda}(j - l_d - 1)M_{[j,N]}^{-\beta},$$

where we used the estimate $M_{[j,N]} \leq (N - j + 1)x_j$, which holds because the sequence x_j is decreasing. To estimate the term III from below, we note that the function $x \mapsto x^{-\beta}$ is convex. Hence we can apply the Jensen inequality to estimate

$$\begin{aligned}\sum_{l=j}^{\lambda} x_l^{-\beta} &= (\lambda - j + 1)\sum_{l=j}^{N} x_l^{-\beta}(\lambda - j + 1)^{-1} \leq (\lambda - j + 1)\left(\sum_{l=j}^{\lambda} x_l(\lambda - j + 1)^{-1}\right)^{-\beta} \\ &= (\lambda - j + 1)^{\beta+1} M_{[j,N]}^{-\beta},\end{aligned}$$

hence

$$\mathrm{III} \geq \frac{(\lambda - j + 1)^{\beta+1}}{\lambda}(j - 1)M_{[l,N]}^{-\beta}.$$

Putting together the estimates for the three terms yields the desired inequality. $\square$

Proof of Proposition 2.16. By definition, as long as $t < \tau_{\varepsilon,j}$ we have

$$\begin{aligned}d^{-\beta} - (\lambda(t) - j + 1)^{\beta} M_{[j,N]}(t)^{-\beta} &\leq d^{-\beta} - \varepsilon^{\beta}(N - j + 1)^{\beta} M_{[j,N]}(t)^{-\beta} \\ &= d^{-\beta} - \varepsilon^{\beta}\rho_{[j,N]}(t)^{-\beta}.\end{aligned}$$

Hence, if η is chosen small enough and $\rho_{[j,N]}(t) \leq \eta$, then the above expression is negative at $t = s$, and by (2.41), $\rho_{[j,N]}$ is decreasing for $s \leq t < \tau$. Then we further estimate the right-hand side of (2.41) to obtain

$$\frac{d}{dt}M_{[j,N]} \leq \kappa\varepsilon(N - j + 1)\left(d^{-\beta} - \varepsilon^{\beta}\rho_{[j,N]}^{-\beta}\right),$$

and dividing by the factor $N - j + 1$ yields the differential inequality

$$\frac{d}{dt}\rho_{[j,N]} \leq \kappa\varepsilon\left(d^{-\beta} - \varepsilon^{\beta}\rho_{[j,N]}^{-\beta}\right) \leq -\frac{1}{2}\kappa\varepsilon^{\beta+1}\rho_{[j,N]}^{-\beta},$$

if η is chosen small enough. Now the statement that $\tau_{\varepsilon,j} \leq T$ for small enough η follows along the same lines as in Lemma 2.13, noting that $\rho_{[j,N]}$ can only vanish after the time $\tau_{\varepsilon,j}$. Finally, multiplying the above inequality by $\rho_{[j,N]}^{\beta}$ and integrating from t to $\tau_{\varepsilon,j}$ for some $t \in (s, \tau_{\varepsilon,j})$ we get

$$\frac{1}{\beta + 1}\left(\rho_{[j,N]}(\tau_{\varepsilon,j})^{\beta+1} - \rho_{[j,N]}(t)^{\beta+1}\right) \leq -\frac{1}{2}\kappa\varepsilon^{\beta+1}(\tau - t),$$

which yields the desired inequality. $\square$

2.3.3. Compactness

Our goal in this subsection is to prove a compactness result for solutions to equation (2.5). For the approximation procedure it is convenient to formulate the compactness properties in terms of size rankings. Recall that the size ranking $x \in L^1_d$ associated with a decreasing sequence $x_1, .., x_N$ is the function

$$x(\varphi) = \sum_{j=1}^{N} x_j \cdot \mathbf{1}_{[\frac{j-1}{N}, \frac{j}{N})}(\varphi).$$

For size ranking solutions $x(t, \varphi)$ we denote by $\tau(\varphi) = \inf\{t \geq 0 \colon x(t, \varphi) = 0\}$ the vanishing time of $x(\cdot, \varphi)$. In the next subsection we will see that besides suitable L^p bounds for solutions, we also need a good control on the vanishing of particles, which motivates the following definition:

Definition 2.18. *Let $h \colon [0, \infty) \to [0, \infty)$. Then we say that h has the stable vanishing property with threshold $\eta \colon [0, \infty) \to (0, \infty)$ and modulus $\alpha \colon [0, \infty)^2 \to [0, \infty)$ if the following holds:*

1. *If $h(t) = 0$ for some $t \geq 0$, then $h(s) = 0$ for all $s \geq t$.*
2. *For all $t \geq 0$ we have $\alpha(t, \cdot) \in C^0([0, \infty))$ and $\alpha(t, 0) = 0$.*
3. *If $h(t) < \eta(t)$ for some $t \geq 0$, then for $\tau_h = \inf\{s \geq 0 \colon h(s) = 0\}$ it holds*

$$\tau_h \leq t + \alpha(t, h(t)). \tag{2.42}$$

We then prove the following result:

Proposition 2.19. *Let K_0 be a compact subset of $L^1([0, 1])$ with $0 \notin K_0$. Then if $x = x_1, .., x_N$ is a solution to equation (2.5) (identified with $x(t, \varphi)$ as above) with initial data $x(0, \cdot) \in K_0$, we have $x(t, \cdot) \in K_t$ for every $t > 0$, where K_t is a compact subset of $L^1([0, 1])$ that depends on K_0, t and β. Furthermore, the following statements hold:*

1. *For all $T > 0$, $\varphi \in [0, 1]$, $t_1, t_2 \in [0, T]$ and $p \in [1, 1 + \beta^{-1})$ we have*

$$\|x(\cdot, \varphi)^{-\beta} \cdot \mathbf{1}_{\{x(\cdot,\varphi)>0\}}\|_{L^p([0,T])} \leq C(K_0, T, \beta, p), \tag{2.43}$$
$$\|\theta\|_{L^p([0,T])} \leq C(K_0, T, \beta, p), \tag{2.44}$$
$$|x(t_2, \varphi) - x(t_1, \varphi)| \leq C(K_0, T, \beta, p)|t_2 - t_1|^{\frac{p-1}{p}}. \tag{2.45}$$

2. *For all $\varphi \in (0, 1)$, the function $x(\cdot, \varphi)$ has the stable vanishing property with threshold η and modulus α, where η, α depend on K_0 and β.*

The key step for the above result is the following local L^p estimate that builds on the results from the previous two subsections:

Lemma 2.20. *Let $x = x_1, .., x_N$ be a solution of equation (2.5) with $l_d(0) \geq \kappa N$ for some $d, \kappa > 0$. Then for every $p \in [1, 1+\beta^{-1})$ there exists a positive constant $C = C(\beta, d, \kappa, p)$ such that for every $j \in [1, N]$ it holds*

$$\|x_j^{-\beta} \cdot \mathbf{1}_{\{x_j>0\}}\|_{L^p([0,T^*/2])} \leq C, \tag{2.46}$$

where T^* is as in (2.25).

Proof. First we note that by the basic estimate (2.25) we have $l_{\frac{d}{2}} \geq \kappa N$ on $[0, T^*]$, where T^* only depends on β and d. Let $\delta, \eta, c > 0$ be according to Proposition 2.12 with $t_0 = T^*/2$. Then we apply Proposition 2.16 with $t_0 = T^*/2$, ε chosen depending on δ (see below) and make η smaller if necessary such that both statements apply with the same constant η. Now consider $j \in [1, N]$. In the following we denote by C any positive constant that may depend on β, d, κ, p. Let

$$t_1 = \begin{cases} \sup\{t \in [0, T^*/2] \colon x_j(t) > \eta\}, \text{ if } x_j(0) > \eta, \\ 0, \text{ if } x_j(0) \leq \eta. \end{cases}$$

Then by construction we have the estimate

$$\int_0^{t_1} x_j(t)^{-p\beta}\,\mathrm{d}t \leq \frac{T^*}{2}\eta^{-p\beta} \leq C.$$

If $t_1 = T^*/2$ we are done. If $t_1 < T^*/2$ we apply Proposition 2.16 at t_1, since $\rho_{[j,N]}(t_1) \leq x_j(t_1) \leq \eta$, which yields $\tau_{\varepsilon,j} \in [t_1, T^*)$. Then by (2.40) have

$$\int_{t_1}^{\tau_{\varepsilon,j}} x_j(t)^{-p\beta}\,\mathrm{d}t \leq C \int_{t_1}^{\tau} (\tau - t)^{-\frac{p\beta}{\beta+1}}\,\mathrm{d}t \leq C,$$

since $\frac{p\beta}{\beta+1} < 1$. If $\tau_{\varepsilon,j} \geq T^*/2$, we are done. If $\tau_{\varepsilon,j} < T^*/2$ we define

$$t_2 = \begin{cases} \sup\{t \in [\tau_{\varepsilon,j}, T^*/2] \colon x_j(t) > \eta\}, \text{ if } x_j(\tau_{\varepsilon,j}) > \eta, \\ \tau_{\varepsilon,j}, \text{ if } x_j(\tau_{\varepsilon,j}) \leq \eta. \end{cases}$$

Then on $[\tau_{\varepsilon,j}, t_2]$ we have the same estimate as on $[0, t_1]$ and if $t_2 = T^*/2$ we are done. Otherwise we proceed once again as follows. By definition of $\tau_{\varepsilon,j}$, we have $|[j, \lambda(t)]| \leq \varepsilon|[j, N]|$ for all $t \geq \tau_{\varepsilon,j}$, which means

$$\begin{aligned} &\lambda(t) - j + 1 \leq \varepsilon(N - j + 1) \leq \varepsilon(\kappa^{-1}\lambda(t) - j + 1), \\ \Rightarrow\quad &(1 - \varepsilon)(j - 1) \geq (1 - \kappa^{-1}\varepsilon)\lambda(t), \end{aligned}$$

hence we have $j \geq (1 - \delta)\lambda(t)$ if ε is chosen small enough depending on δ and κ. Then because $t_2 < T^*/2$ Proposition 2.12 is applicable and implies $\tau_j \leq T^*$ and the lower bound (2.28). Finally we set $t_3 = \tau_j \wedge T^*/2$ and estimate the integral on $[t_2, t_3)$ in the same way as on $[t_1, \tau_{\varepsilon,j})$, which finishes the proof. □

Corollary 2.21. *Let $x_1, .., x_N$ as above. Then for every $p \in [1, 1+\beta^{-1})$ there exists a positive constant $C = C(\beta, d, \kappa, p)$ such that*

$$\|\theta\|_{L^p([0,T^*/2])} \leq C, \tag{2.47}$$

$$|x_j(t_2) - x_j(t_1)| \leq C|t_2 - t_1|^{\frac{p-1}{p}}, \text{ for} j \in \{1, .., N\}, \; t_1, t_2 \in [0, T^*/2]. \tag{2.48}$$

Proof. We have

$$\|\theta\|_{L^p} = \left\| \lambda^{-1} \sum_{j=1}^N x_j^{-\beta} \cdot \mathbf{1}_{\{x_j>0\}} \right\|_{L^p} \leq \kappa^{-1} N^{-1} \sum_{j=1}^N \|x_j^{-\beta} \cdot \mathbf{1}_{\{x_j>0\}}\|_{L^p} \leq \kappa^{-1} C,$$

with C as in Proposition 2.20. The bound for $|x_j(t_2) - x_j(t_1)|$ then easily follows by integrating the differential equation for x_j and Hölder inequality. $\square$

From now on we choose $d = \frac{\rho}{2}$ in accordance with Lemma 2.11. Then we want to extend all of the previous analysis which was done on the time interval $[0, T^*]$, respectively $[0, T^*/2]$ to arbitrary time intervals. This is done by iterating the analysis from above. The key observation is that the timespan T^* in each step does only depend on ρ and not the volume fraction $l_{\frac{\rho}{2}}$, which we cannot uniformly control uniformly from below.

Proof of Proposition 2.19. Let $N > 0$ and $x = x_1, .., x_N$ be a solution to equation (2.5) with initial data $x(0, \cdot) \in K_0$ as identified above. Because K_0 is compact, we have $\operatorname{dist}(K_0, 0) \geq \rho > 0$. By construction we have

$$\int_0^1 x(0, \varphi) \, \mathrm{d}\varphi = \frac{1}{N} \sum_{j=1}^N x_j(0) \geq \rho.$$

Without loss of generality we assume that we have equality in the above line, as all relevant estimates only get better if ρ is larger. Again by compactness, we have

$$\lim_{R \to \infty} \sup_{y \in K_0} \int_0^1 (y(\varphi) - R)_+ \, \mathrm{d}\varphi = 0.$$

Hence there exists $R_0 > 0$ depending only on K_0 such that

$$\frac{\rho}{2} - \frac{1}{N} \sum_{i=1}^N (x_i(0) - R_0)_+ = \frac{\rho}{2} - \int_0^1 (x(0, \varphi) - R_0)_+ \, \mathrm{d}\varphi \geq \frac{\rho}{4},$$

and thus Lemma 2.11 implies that

$$l_{\frac{\rho}{2}}(0) \geq \frac{\rho N}{4 R_0} =: \kappa_0 N.$$

Then by Lemma 2.20 and Corollary 2.21 we have

$$\begin{aligned} \|x(\cdot, \varphi)^{-\beta} \cdot \mathbf{1}_{\{x(\cdot,\varphi)>0\}}\|_{L^p([0,T^*/2])} &\leq C_0, \text{ for } \varphi \in [0,1], \\ \|\theta\|_{L^p([0,T^*/2])} &\leq C_0, \\ x(t, \varphi) &\leq x(0, \varphi) + C_0, \text{ for } t \in [0, T^*/2], \end{aligned}$$

where C_0 only depends on K_0 and β. In particular we have

$$\int_0^1 (x(t,\varphi) - R)_+ \,\mathrm{d}\varphi \le \int_0^1 (x(0,\varphi) - R)_+ \,\mathrm{d}\varphi + (C_0 - R)_+,$$

Hence for $R = R_1 := \max\{C_0, R_0\}$ we again apply Lemma 2.11 and obtain

$$l_{\frac{\varrho}{2}}(t) \ge \frac{\rho N}{4R_1} =: \kappa_1 N \text{ for } t \in [0, T^*/2].$$

Thus we can again apply Corollary 2.21 (on the interval $[T^*/2, T^*]$) and get analogous estimates from above with different constants C_1, R_2. Iterating this argument, we see that for every $T > 0$ we have estimates

$$\begin{aligned} \|x(\cdot,\varphi)^{-\beta} \cdot \mathbf{1}_{\{x(\cdot,\varphi)>0\}}\|_{L^p([0,T])} &\le C(T,p), \text{ for } \varphi \in [0,1], \\ \|\theta\|_{L^p([0,T])} &\le C(T,p), \\ \int_0^1 (x(T,\varphi) - R)_+ \,\mathrm{d}\varphi &\le \int_0^1 (x(0,\varphi) - R)_+ \,\mathrm{d}\varphi + (C(T) - R)_+, \\ l_{\frac{\varrho}{2}}(T) &\ge \kappa(T)N. \end{aligned}$$

In particular the third line implies the desired compactness with K_t chosen as the set consisting of all solutions $x(t,\cdot)$ with initial data in K_0, and the Hölder bound follows from the L^p-bounds on $\theta, x(\cdot,\varphi)^{-\beta} \cdot \mathbf{1}_{\{x(\cdot,\varphi)>0\}}$ and equation (2.19).

To prove the stable vanishing property let $t \ge 0$ and $j \in \{1,..,N\}$ with $x_j(t) < \eta$, where η is chosen below. We want to use Proposition 2.12 and Proposition 2.16 with $T = t+1$ and $t_0 = t$. By the above considerations we have $l_{\frac{\varrho}{2}} \ge \kappa N$ on $[0,T]$ for some $\kappa > 0$. first we apply Proposition 2.16 with ε to be chosen below. Then $\tau_{\varepsilon,j} \le t+1$ and (2.40) implies that

$$\tau_{\varepsilon,j} \le t + C\rho_{[j,N]}(t)^{\beta+1} \le t + Cx_j(t)^{\beta+1}.$$

Here and in the following, we denote by C a constant that may depend on β, K_0, ε and t. Next we note that the Hölder continuity on $[0, t+1]$ implies

$$x_j(\tau_{\varepsilon,j}) \le x_j(t) + C(\tau_{\varepsilon,j} - t)^{\frac{p-1}{p}} \le x_j(t) + Cx_j(t)^{\frac{(\beta+1)(p-1)}{p}},$$

for some fixed $p \in (1, 1+\beta^{-1})$. As in the proof of Lemma 2.20 we note that $j \ge (1-\delta(\varepsilon))\lambda(s)$ for $s \ge t$ with $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$. Hence after choosing ε small enough we apply Proposition 2.12 with $T = t+2$, $t_0 = t+1$. If $x_j(t) \le \eta$ with η small enough, then $x_j(\tau_{\varepsilon,j})$ is small enough so the requirements for the Proposition are met. Hence $\tau_j \le t+2$ and by (2.28) and the preceding considerations we conclude

$$\begin{aligned} \tau_j &\le \tau_{\varepsilon,j} + Cx_j(\tau_{\varepsilon,j})^{\beta+1} \\ &\le t + Cx_j(t)^{\beta+1} + C\left(x_j(t) + Cx_j(t)^{\frac{(\beta+1)(p-1)}{p}}\right)^{\beta+1}, \end{aligned}$$

which implies the stable vanishing property for $x(\cdot,\varphi)$. □

2.3.4. Passage to the limit

Now want to prove well-posedness of the mean-field equation in L^1_d by approximation. Note that the subset of functions in L^1_d that are of the form (2.14) for some $N > 0$ are dense in L^1_d. We denote the set of such functions by $\mathcal{T}$. By the previous results of this section, equation (2.17) is well-posed in $\mathcal{T}$ and solutions are represented by finite sequences solving equation (2.5). We want to use the compactness for these solutions (Proposition 2.19) to pass to the limit. Because of the discontinuity of the function $x \to \mathbf{1}_{\{x>0\}}$, appearing on the right hand side of (2.17), we first need to show that vanishing times of approximating solutions converge to vanishing times of the limit function. This is a consequence of the stable vanishing property:

Lemma 2.22. *Let $h_n\colon [0,\infty) \to [0,\infty)$ be a sequence of functions that have the stable vanishing property. If $h\colon [0,\infty) \to [0,\infty)$ and $h_n \to h$, then $\tau_{h_n} \to \tau_h$ and h has the stable vanishing property.*

Proof. First we show that $\tau_{h_n} \to \tau_h$. It suffices to show that the limit of any converging subsequence (where convergence to ∞ is possible) is equal to τ_h. Thus we assume without relabeling that $\tau_{h_n} \to \tau \in [0,\infty]$. Then for $t > \tau$ we have

$$h(t) = \lim_{n\to\infty} h_n(t) = 0,$$

because $t > \tau_{h_n}$ for large enough n. Next we show that $h(t) > 0$ for all $t < \tau$. If this were not true, then there exists $t' < \tau$ such that $\lim_{n\to\infty} h_n(t') = 0$, in particular $h_n(t') < \eta(t')$ if n is large. Then the stable vanishing property of h_n implies

$$\tau_{h_n} \le t' + \alpha(t', h_n(t')) \to t', \ n \to \infty,$$

because of $\alpha(t', 0) = 0$ and continuity, which contradicts $\tau_{h_n} \to \tau > t'$. This shows that $h(t) > 0$ if $t < \tau$ and $h(t) = 0$ if $t > \tau$, hence $\tau = \tau_h$. It is left to show that h has the stable vanishing property. If $h(t) < \eta(t)$, then $h_n(t) < \eta(t)$ for n large. Then we have

$$\tau_{h_n} \le t + \alpha(t, h_n(t)),$$

so letting $n \to \infty$ and using $\tau_{h_n} \to \tau_h$ yields the desired property. □

We can now finally state and prove the result regarding approximation of general solutions, which is done in similar fashion as [52, Proposition 6.1], although in our case we have to be a bit more careful when passing to the limit in the equation due to the singular and discontinuous terms on the right-hand side of the equation. Then Theorem 2.3 is a direct consequence of the following result and Proposition 2.8:

Proposition 2.23. *Let $\{x^{(n)}(t,\cdot)\} \subset \mathcal{T}$ be a sequence of solutions to equation (2.17) with mean-fields $\theta^{(n)}$ and assume that there exists $x_0 \in L^1_d$, $x_0 \neq 0$, such that $x^{(n)}(0,\cdot) \to x_0$ in L^1. Then there exists $x \in C^0([0,\infty); L^1_d)$, $\theta \in L^p_{loc}([0,\infty))$ for all $p \in (1, 1+\beta^{-1})$ and a subsequence $n \to \infty$ such that for all $T > 0$ we have*

$$\sup_{t\in[0,T]} \|x^{(n)}(t,\cdot) - x(t,\cdot)\|_{L^1((0,1))} \to 0, \quad \theta^{(n)} \rightharpoonup \theta \text{ in } L^p((0,T)),$$

and x, θ solve the mean-field equation (2.17) on $[0,\infty)$.

Proof. Troughout the proof subsequences are not relabeled for convenience. Because $x^{(n)}(0,\cdot) \to x_0 \neq 0$ in L^1, the assumption for Proposition 2.19 is satisfied with $K_0 = \{x^{(n)}(0,\cdot)\} \cup \{x_0\}$, where we assume w.l.o.g that $x^{(n)}(0,\cdot) \neq 0$. Then Proposition 2.19 implies that for each $t \geq 0$, the sequence $x^{(n)}(t,\cdot)$ is precompact in L^1 and by (2.45) it holds

$$\int_0^1 |x^{(n)}(t_2,\varphi) - x^{(n)}(t_1,\varphi)| \,\mathrm{d}\varphi \leq C|t_2 - t_1|^{\frac{p-1}{p}},$$

which implies equicontinuity of the sequence in $C^0([0,\infty);L^1(0,1))$. Hence by applying Arzela-Ascoli Theorem on compact time intervals and a diagonal argument there exists $x \in C^0([0,\infty);L^1_d)$ such that $\sup_{t\in[0,T]}\|x^{(n)}(t,\cdot) - x(t,\cdot)\|_{L^1((0,1))} \to 0$ for every $T > 0$. In particular the L^1 convergence implies that the mass constraint is satisfied by the limit:

$$\int_0^1 x(t,\varphi)\,\mathrm{d}\varphi = \int_0^1 x_0(\varphi)\,\mathrm{d}\varphi.$$

By weak compactness and the bound (2.44) there exists $\theta \in L^p_{loc}((0,\infty))$ such that $\theta^{(n)} \rightharpoonup \theta$ in $L^p((0,T))$ for every $T > 0$. To show that x, θ solve equation (2.17), we have to exploit another compactness property of the sequence $x^{(n)}$ which follows from monotonicity. Indeed, since $x^{(n)}(t,\cdot)$ is decreasing for every $t \geq 0$, we have for $\phi \in (0,1)$ that

$$\phi x^{(n)}(t,\phi) \leq \int_0^1 x^{(n)}(t,\varphi)\,\mathrm{d}\varphi \leq C.$$

This implies that the sequence $x^{(n)}(t,\cdot)$ is bounded and decreasing on $[\phi,1)$ for every $\phi > 0$. Thus by Helly's selection theorem and a diagonal argument there exists $\tilde{x}(t,\varphi)$ such that $x^{(n)}(t,\varphi) \to \tilde{x}(t,\varphi)$ for every $\varphi \in (0,1)$. By a further diagonal argument such a $\tilde{x}(t,\varphi)$ exists for all rational t, and finally by Hölder continuity in time for all $t \geq 0$. In particular we have for all $t \geq 0, \varphi \in (0,1)$ that

$$x^{(n)}(t,\varphi) \to \tilde{x}(t,\varphi).$$

Note that due to $\sup_{t\in[0,T]}\|x^{(n)}(t,\cdot) - x(t,\cdot)\|_{L^1((0,1))} \to 0$ we have $x(t,\cdot) = \tilde{x}(t,\cdot)$ almost everywhere for every $t \geq 0$, but $\tilde{x}$ might not be right-continuous. First we show that $\tilde{x}, \theta$ satisfy identity (2.17). Since $x^{(n)}$ is a solution we have for every $t \geq 0$ and $\varphi \in (0,1)$

$$x^{(n)}(t,\varphi) = x^{(n)}(0,\varphi) + \int_0^t \left(\theta^{(n)}(s) - x^{(n)}(s,\varphi)^{-\beta}\right) \cdot \mathbf{1}_{\{x^{(n)}(s,\varphi)>0\}}\,\mathrm{d}s.$$

We want to pass to the limit in the above identity. By Lemma 2.22 we know that the vanishing times $\tau_n(\varphi)$ of $x^{(n)}(\cdot,\varphi)$ converge to the vanishing time $\tilde{\tau}(\varphi)$ of $\tilde{x}(\cdot,\varphi)$. In particular we have $\mathbf{1}_{\{x^{(n)}(s,\varphi)>0\}} \to \mathbf{1}_{\{\tilde{x}(s,\varphi)>0\}}$. Together with the weak convergence of $\theta^{(n)}$ this implies

$$\int_0^t \theta^{(n)}(s) \cdot \mathbf{1}_{\{x^{(n)}(s,\varphi)>0\}}\,\mathrm{d}s \to \int_0^t \theta(s) \cdot \mathbf{1}_{\{\tilde{x}(s,\varphi)>0\}}\,\mathrm{d}s.$$

For the second term in the integral note that for each fixed φ,

$$x^{(n)}(\cdot,\varphi)^{-\beta}\cdot\mathbf{1}_{\{x^{(n)}(\cdot,\varphi)>0\}}\to\tilde{x}(\cdot,\varphi)^{-\beta}\cdot\mathbf{1}_{\{\tilde{x}(\cdot,\varphi)>0\}},$$

almost everywhere on $(0,t)$. Together with the bound (2.43) this implies that the above convergence holds weakly in $L^p((0,t))$, and thus

$$\int_0^t x^{(n)}(s,\varphi)^{-\beta}\cdot\mathbf{1}_{\{x^{(n)}(s,\varphi)>0\}}\,\mathrm{d}s\to\int_0^t\tilde{x}(s,\varphi)^{-\beta}\cdot\mathbf{1}_{\{\tilde{x}(s,\varphi)>0\}}\,\mathrm{d}s,$$

showing that $\tilde{x},\theta$ satisfy eq. (2.17). It is left to show that the mean-field ODE not only holds for $\tilde{x},\theta$, but also for x,θ. To do this, note that because $x(t,\varphi)=\tilde{x}(t,\varphi)$ for almost all φ, both functions being decreasing and x being right-continuous we have

$$x(t,\varphi)=\lim_{\varepsilon\to0^+}\tilde{x}(t,\varphi+\varepsilon),\ \text{ for all }\varphi\in(0,1).$$

Thus the desired result follows if we can pass to the limit $\varepsilon\to0^+$ in the equation

$$\tilde{x}(t,\varphi+\varepsilon)=\tilde{x}(0,\varphi+\varepsilon)+\int_0^t\left(\theta(s)-\tilde{x}(s,\varphi+\varepsilon)^{-\beta}\right)\cdot\mathbf{1}_{\{\tilde{x}(s,\varphi+\varepsilon)>0\}}\,\mathrm{d}s. \tag{2.49}$$

The argument here is very similar to the limit $n\to\infty$ from above. For this we have to establish that $\tilde{\tau}(\varphi+\varepsilon)\to\tau=\tau(\varphi)=\inf\{t\geq0\colon x(t,\varphi)=0\}$ as $\varepsilon\to0^+$. This is again a consequence of Lemma 2.22, since $\tilde{x}$ inherits the stable vanishing property from $x^{(n)}$ by pointwise convergence, hence the vanishing times of the sequence $\tilde{x}(\cdot,\varphi+\varepsilon_n)$ for $\varepsilon_n\to0$ converge to the vanishing time of $x(\cdot,\varphi)$. Also all a priori bounds for $x^{(n)}$ carry over to $\tilde{x}$, and so the limit $\varepsilon\to0^+$ in equation (2.49) is carried out as before, which finishes the proof. □

2.4. Existence and uniqueness of self-similar profiles

To find self-similar solutions to equation (2.6), we make the ansatz $\varphi(t,x)=t^{-\frac{2}{1+\beta}}\Phi(t^{-\frac{1}{1+\beta}}x)$, so that

$$\int_0^\infty x\varphi(t,x)\,\mathrm{d}x=\int_0^\infty z\Phi(z)\,\mathrm{d}z,$$
$$\theta[\varphi(t,\cdot)]=t^{-\frac{\beta}{1+\beta}}\theta[\Phi].$$

Here we require that Φ has finite first moment, $\theta[\Phi]<\infty$ and $\Phi\in W^{1,1}_{loc}(\mathbb{R}_+)$ so φ is a weak solution in the sense of Definition 2.1. With $z=t^{-\frac{1}{1+\beta}}x$ we then calculate

$$\begin{aligned}\partial_t\varphi&=-t^{-\frac{2}{1+\beta}-1}\left(\frac{2}{1+\beta}\Phi(z)+\frac{1}{1+\beta}z\Phi'(z)\right),\\ \partial_x\left((x^{-\beta}-\theta[\varphi])\varphi\right)&=-\beta x^{-\beta-1}\varphi+(x^{-\beta}-\theta[\varphi])\partial_x\varphi\\ &=t^{-\frac{2}{1+\beta}-1}\left(-\beta z^{-\beta-1}\Phi(z)+(z^{-\beta}-\theta[\Phi])\Phi'(z)\right).\end{aligned}$$

Thus, φ is a solution if and only if Φ satisfies

$$\frac{2}{1+\beta}\Phi(z) + \frac{1}{1+\beta}z\Phi'(z) = \beta z^{-\beta-1}\Phi(z) - (z^{-\beta} - \theta[\Phi])\Phi'(z), \tag{2.50}$$

or

$$\Phi' = \frac{\beta z^{-\beta-1} - \frac{2}{1+\beta}}{z^{-\beta} - \theta[\Phi] + \frac{1}{1+\beta}z}\Phi. \tag{2.51}$$

First we investigate for which values of $\theta[\Phi]$ there can exist a reasonable solution to equation (2.51). To this end we introduce for $\vartheta \in [0,\infty)$ the function

$$\lambda_\vartheta(z) = \frac{\beta z^{-\beta-1} - \frac{2}{1+\beta}}{z^{-\beta} - \vartheta + \frac{1}{1+\beta}z}, \tag{2.52}$$

and consider the differential equation

$$\Phi'(z) = \lambda_\vartheta(z)\Phi(z). \tag{2.53}$$

Then we have the following results:

Lemma 2.24. *There exist constants $\theta_{min}, \theta_{max}$ depending only on β such that for every $\vartheta \in [\theta_{min}, \theta_{max})$ there exists $z_* \in (0,\infty)$ such that the following holds: Every non-zero and non-negative solution Φ to equation (2.53) stays strictly positive on $(0, z_*)$, vanishes at z_* and satisfies*

$$\Phi(z) \sim c_1 z^\beta, \text{ as } z \to 0^+,$$
$$\Phi(z) \sim \begin{cases} c_2(z_* - z)^r, \text{ as } z \to z_*^- \text{ if } \vartheta < \theta_{min}, \\ c_2 e^{-\frac{R}{z_*-z}}, \text{ as } z \to z_*^- \text{ if } \vartheta = \theta_{min}, \end{cases}$$

where $c_1, c_2, R, r > 0$ depend on β, r with $r \to \infty$ as $\vartheta \to \theta_{min}^+$ and $r \to 0$ as $\vartheta \to \theta_{max}$.

Proof. Let $\lambda_1(z) = \beta z^{-\beta-1} - \frac{2}{1+\beta}$, $\lambda_2(z) = z^{-\beta} - \vartheta + \frac{1}{1+\beta}z$, so that

$$\lambda_\vartheta(z) = \frac{\lambda_1(z)}{\lambda_2(z) - \vartheta}.$$

Then λ_2 decreases from infinity at $z = 0$, attains its minimum at $z = z_{min} = (\beta(1+\beta))^{\frac{1}{1+\beta}}$ and then increases, while λ_1 decreases and changes its sign from positive to negative at $z = \frac{1}{2}z_{min}$. We then set

$$\theta_{min} = \lambda_2(z_{min}),$$
$$\theta_{max} = \lambda_2\left(\frac{1}{2}z_{min}\right),$$

and for $\vartheta \in [\theta_{min}, \theta_{max})$ we choose z_* to be the smallest point such that $\lambda_2(z_*) - \vartheta = 0$. Due to the above discussion z_* is well defined and we have

$$\frac{1}{2} z_{min} < z_* \leq z_{min},$$

and z_* is a decreasing function of ϑ with $z_* = z_{min}$ if $\vartheta = \theta_{min}$. Then the function λ_ϑ satisfies

$$\begin{aligned} \lambda_\vartheta(z) > 0, \text{ for } z \in \left(0, \frac{1}{2} z_{min}\right), \\ \lambda_\vartheta(z) < 0, \text{ for } z \in \left(\frac{1}{2} z_{min}, z_*\right). \end{aligned}$$

Next we consider the asymptotic behavior of λ_ϑ as $z \to 0^+$ and $z \to z_*^-$. First we calculate

$$\lambda_\vartheta(z) = \frac{\beta z^{-1} - \frac{2}{1+\beta} z^\beta}{1 - \vartheta z^\beta + \frac{1}{1+\beta} z^{\beta+1}} = \frac{\beta z^{-1}}{1 - \vartheta z^\beta + \frac{1}{1+\beta} z^{\beta+1}} + \mathcal{O}(1) = \frac{\beta}{z} + \mathcal{O}(z^{\beta-1}), \text{ as } z \to 0^+.$$

For the asymptotics at z_* we use Taylor expansion for λ_1, λ_2:

$$\begin{aligned} \lambda_\vartheta(z) &= \frac{\lambda_1(z_*) + \lambda'(z_*)(z - z_*) + \mathcal{O}((z - z_*)^2)}{\lambda_2'(z_*)(z - z_*) + \frac{1}{2}\lambda_2''(z_*)(z - z_*)^2 + \mathcal{O}((z - z_*)^3)} \\ &= \frac{\lambda_1(z_*)}{\lambda_2'(z_*)(z - z_*)} + \mathcal{O}(1), \text{ as } z \to z_*^-, \end{aligned}$$

if $\lambda_2'(z_*) \neq 0$, which is the case for $\vartheta > \theta_{min}$. In the boundary case $\vartheta = \theta_{min}$ we have $z_* = z_{min}$, hence $\lambda_2'(z_*) = 0$. A short calculation shows that $\lambda_2''(z_{min}) > 0$, thus in this case we have

$$\lambda_\vartheta(z) = \frac{2\lambda_1(z_*)}{\lambda_2''(z_*)(z - z_*)^2} + \frac{2\lambda_1'(z_*)}{\lambda_2''(z_*)(z - z_*)} + \mathcal{O}(1), \text{ as } z \to z_*^-.$$

To conclude the asymptotic behavior of the solution at $z = 0, z = z_*$ we choose some point $z_0 \in (0, z_*)$ with $\Phi(z_0) > 0$. Then by (2.51) Φ is given by

$$\Phi(z) = \Phi(z_0) \exp\left(\int_{z_0}^{z} \lambda_\vartheta(\xi) \, \mathrm{d}\xi\right),$$

which implies $\Phi > 0$ on $(0, z_*)$ as well as the desired asymptotics by plugging in the asymptotic expressions for λ_ϑ from above. Regarding the behavior of the exponent $r = \frac{\lambda_1(z_*)}{\lambda_2'(z_*)}$, we note that $z_* \to \frac{1}{2} z_{min}$ as $\vartheta \to \theta_{max}$ and $z_* \to z_{min}$ as $\vartheta \to \theta_{min}$, hence

$$\frac{\lambda_1(z_*)}{\lambda_2'(z_*)} \to 0, \text{ as } z_* \to \frac{1}{2} z_{min}, \quad \frac{\lambda_1(z_*)}{\lambda_2'(z_*)} \to \infty, \text{ as } z_* \to z_{min}.$$

This finishes the proof. $\square$

Lemma 2.25. *For $\vartheta < \theta_{min}$, every solution to eq. (2.53) stays strictly positive on $(0, \infty)$ and satisfies $\Phi(z) \sim z^{-2}$, as $z \to \infty$. For $\vartheta \geq \theta_{max}$ every solution blows up at some finite point.*

Proof. If $\vartheta < \theta_{min}$, $\lambda_2 - \vartheta > 0$, and hence λ_ϑ is positive on a finite interval, then becomes negative and behaves asymptotically like $2z^{-1}$ as $z \to \infty$, which implies that $\Phi(z)$ decreases like z^{-2} at infinity. If $\vartheta = \theta_{max}$, then λ_1 and $\lambda_2 - \vartheta$ both become zero with linear rate at the same point and both stay negative until $\lambda_2 - \vartheta$ becomes zero again at some point z_1 with a linear rate. Thus λ_ϑ stays positive and is non-integrable at z_1 which means $\Phi \to \infty$ as $z \to z_1^-$. If $\vartheta > \theta_{max}$, the first zero of $\lambda_2 - \vartheta$ lies in the range where λ_1 is still positive and hence the blow up occurs already at the first zero of $\lambda_2 - \vartheta$. □

In particular, if $\vartheta \notin [\theta_{min}, \theta_{max})$, Φ does not have finite first moment or does not exist globally and is thus not a candidate for a self-similar profile. If Φ is a solution to equation (2.53) on $(0, z_*)$, then by Lemma 2.24 we can extend Φ by zero on (z_*, ∞) and obtain a solution to equation (2.53) on $[0, \infty)$ in $W^{1,1}_{loc}(\mathbb{R}_+)$ with compact support. Note that due to the behavior of the exponent r at the end of the support, the solution becomes smoother as $\vartheta \to \theta_{min}^+$ and is completely smooth at θ_{min}.

The difference between eq. (2.51) and (2.53) is that in (2.53) the ϑ is a parameter that is free to choose while in (2.51) we have the requirement $\vartheta = \theta[\Phi]$. However, we have the following consistency result:

Lemma 2.26. *Every solution Φ to equation (2.53) satisfies $\theta[\Phi] = \vartheta$. In particular every solution to equation (2.53) also solves (2.51).*

Proof. The equation for Φ is equivalent to

$$\frac{2}{1+\beta}\Phi(z) + \frac{1}{1+\beta} z\Phi'(z) = \beta z^{-\beta-1}\Phi(z) - (z^{-\beta} - \vartheta)\Phi'(z).$$

Multiplying this identity with z and integrating, where we integrate the terms that contain Φ' by parts, we arrive at

$$\begin{aligned}\frac{2}{1+\beta}\int_0^\infty z\Phi(z)\,\mathrm{d}z - \frac{2}{1+\beta}\int_0^\infty z\Phi(z)\,\mathrm{d}z = \beta \int_0^\infty z^{-\beta}\Phi(z)\,\mathrm{d}z + (1-\beta)\int_0^\infty z^{-\beta}\Phi(z)\,\mathrm{d}z \\ - \vartheta \int_0^\infty \Phi(z)\,\mathrm{d}z,\end{aligned}$$

which simplifies to $\theta[\Phi] = \vartheta$. □

For Proposition 2.5 it remains to show that solutions to (2.53) actually exist and are unigue up to normalization:

Lemma 2.27. *For every $\vartheta \in [\theta_{min}, \theta_{max})$ there exists a solution Φ to eq. (2.53) on $(0, z_*)$. This solution is unique up to a constant factor.*

Proof. For existence choose some $z_1 \in (0, z_*)$ and $C > 0$. Then there is a local solution $\Phi' = \lambda_\vartheta \Phi$ around z_1 with $\Phi(z_1) = C$ that exists on $(0, z_*)$. By Lemma 2.24 the solution tends to 0 at $z = 0$ and $z = z_*$ with the corresponding asymptotic behavior. For uniqueness, let Φ_1, Φ_2 be two solutions. Then by rescaling we assume that $\lim_{z \to 0^+} z^{-\beta}\Phi_1(z) = \lim_{z \to 0^+} z^{-\beta}\Phi_2(z)$. Define the functions $\Psi_i(z) = z^{-\beta}\Phi_i(z)$. Then the function $W = (\Psi_1 - \Psi_2)^2$ satisfies the equation

$$\begin{aligned} W'(z) &= 2(\lambda_\vartheta(z) - \beta z^{-1})W(z), \\ W(0) &= 0. \end{aligned}$$

In the proof of Lemma 2.24 we established $(\lambda_\vartheta(z) - \beta z^{-1}) = \mathcal{O}(z^{\beta-1})$ as $z \to 0^+$. In particular this function is integrable on $(0, z_*)$, which implies $W = 0$ everywhere. $\square$

3. Self-similar behavior of the exchange-driven growth model with product kernel

The contents of this chapter are joint work with A. Schlichting and were published in [17].

3.1. Introduction and results

3.1.1. The exchange-driven growth model

The exchange-driven growth model describes a broad class of physical processes in which pairs of clusters consisting of an integer number of monomers can grow or shrink only by exchanging a single monomer [7]. The physical motivation behind the growth processes based on this exchange mechanism is quite different from classical aggregation models like the Smoluchowski coagulation equation [56, 24], which explains its recent interest. Moreover, the underlying exchange mechanism is not restricted to physical models but can be applied to social phenomena like migration [39], population dynamics [46], and wealth exchange [36]. It is also found in diverse phenomena at contrasting scales from microscopic level polymerization processes [15], to cloud [35] and galaxy formation mechanisms at massive scales, as well as in statistical physics [41]. Although this process is not necessarily realized by chemical kinematics, it is convenient to be interpreted as a reaction network of the form

$$X_{k-1} + X_l \underset{K(k,l-1)}{\overset{K(l,k-1)}{\rightleftharpoons}} X_k + X_{l-1} \,, \qquad \text{for} \quad k,l \geq 1 \,. \tag{3.1}$$

The clusters of size $k \geq 1$ are denoted by X_k. Additionally, the variable X_0 represents empty volume. The kernel $K(k, l-1)$ encodes the rate of the exchange of a single monomer from a cluster of size k to a cluster of size $l-1$. Here and in the following the notation $k \geq 1$ means $k \in \mathbb{N} = \{1, 2, \dots\}$ and $l \geq 0$ denotes $l \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$. The concentrations of X_k in (3.1) are denoted by $(c_k)_{k\geq 0}$ and satisfy for $k \geq 0$ the reaction rate equation formally obtained from (3.1) by mass-action kinetics

$$\begin{aligned} \dot{c}_k \;=\; & \sum_{l\geq 1} K(l,k-1) c_l c_{k-1} - \sum_{l\geq 1} K(k,l-1) c_k c_{l-1} \\ & - \sum_{l\geq 1} K(l,k) c_l c_k + \sum_{l\geq 1} K(k+1,l-1) c_{k+1} c_{l-1} \,, \qquad \text{for } k \geq 0 \,. \end{aligned} \tag{EDG}$$

It is easy to see that, at least formally, the quantities

$$\sum_{k\geq 0} c_k \qquad \text{and} \qquad \sum_{k\geq 1} k\, c_k$$

are conserved during the evolution. The first sum can be interpreted as the total volume of the system and the second one as mass (or density, if normalized). Without loss of generality by a suitable time-change, the volume is normalized to 1, so throughout we assume that

$$\sum_{k\geq 0} c_k = 1 \qquad \text{and} \qquad \sum_{k\geq 1} k\, c_k = \rho \in [0,\infty). \tag{3.2}$$

With this normalization, equation (EDG) can also be viewed as mean-field limit of an interacting stochastic particle system, where N particles on a complete graph of size L move between sites according to a jump process in which a jump between a site with k particles and a site with l particles occurs with rate $K(k,l)/(N-1)$. Then the statistical description of the population of clusters in the limit $N, L \to \infty$ such that ${}^{N}/_{L} \to \rho$ is given by equation (EDG), where c_k is the fraction of sites with k particles. This coarse-grained limit was rigorously derived in [29].

The mathematical theory of the equation (EDG) itself started with well-posedness results for kernels with at most linear growth in [22, 54], that is $K(k,l) \leq C\,k\,l$. In addition, for nearly symmetric kernels satisfying $K(k,l) = K(l,k)$, the global well-posedness can be extended to kernels satisfying $K(k,l) \leq C(k^\mu l^\nu + k^\nu l^\mu)$ with $\mu, \nu \in [0,2]$ and $\mu + \nu \leq 3$ (cf. [22, Theorem 2]). The long-time behavior of solutions is investigated in [54, 21], where the crucial assumption on the kernels is a detailed balance condition or some suitable monotonicity properties. For kernels satisfying the detailed balance condition, equation (EDG) has many striking similarities to the Becker-Döring equation [5, 3]. In particular, there exist unique equilibrium states ω^ρ with density ρ up to a critical value ρ_c, and a solution c with density ρ converges $c(t) \to \omega^{\min(\rho,\rho_c)}$ as $t \to \infty$, where the convergence is strong if $\rho \leq \rho_c$ and weak if $\rho > \rho_c$. In the latter case, the bulk of the system relaxes to ω^{ρ_c} while the excess density $(\rho - \rho_c)$ condensates in larger and fewer clusters, which is analogous to the classical LSW [47, 59] coarsening picture treated in [49, 55]. In light of these results, it is natural to ask whether the excess density in (EDG) coarsens in a self-similar way.

It is worth mentioning that condensation and self-similar behavior is already present on the level of stochastic particle systems. In zero-range processes [31, 38, 28] and explosive condensation models [58, 13, 23] the coarsening happens with rates satisfying the detailed balance condition and in particular $K(k,0) > 0$ for all $k \in \mathbb{N}$. The attractive interaction between particles causes condensation in those models. Although the zero-range process's kernel is bounded, coarsening and convergence to a self-similar profile is formally described in [28] in the mean-field case. A first rigorous result beyond the mean-field situation is obtained in [6]. They derive an effective process of the multi-condensate phase in the zero-range process on a finite graph with diverging particle density. Contrary to explosive models with unbounded kernels, it is possible to observe even instantaneous gelation within suitable limits. For inclusion processes, one often studies the case $K(k,0) \to 0$ for all $k \in \mathbb{N}$ in the limit of infinite volume or particle density [30, 11, 37] so that the microscopic dynamics are irreducible and non-degenerate. However, the limiting coarsening mechanism is driven by the absorbing boundary with $K(k,0) = 0$ for all $k \in \mathbb{N}$, which is also the case for the kernels we consider.

In this chapter, we provide rigorous results about the coarsening and self-similar behavior of solutions to (EDG) with the specific family of product kernels

$$K(k,l) = K_\lambda(k,l) = a_\lambda(k)a_\lambda(l) \qquad \text{with} \qquad a_\lambda(k) = \begin{cases} k^\lambda, & \lambda > 0, \\ 1 - \delta_{k,0}, & \lambda = 0, \end{cases} \tag{3.3}$$

for all $\lambda \in [0,2)$. These and more general symmetric homogeneous kernels were introduced and investigated in [7]. A crucial property is that $K(k,0) = 0$, which on the level of clusters means that a cluster with no particles cannot regain particles and hence is virtually removed from the system. In particular this violates the aforementioned detailed balance condition. It is easy to see that the only equilibrium is the vacuum state $c_0 = 1,\ c_k = 0,\ k \geq 1$. During the evolution, particles distribute among fewer and larger clusters over time while smaller clusters die out. This means that the driving coarsening mechanism in this case is the loss of volume, in contrast to the detailed balance case, where coarsening is induced by attraction between particles and only affects the excess density.

The symmetry and product form of K simplify the system (EDG) considerably. We introduce the moments for some $\kappa \in [0,\infty)$ by

$$M_\kappa = M_\kappa[c] = \sum_{l \geq 1} l^\kappa c_l. \tag{3.4}$$

Note that we exclude $k = 0$ in the summation, so $M_0[c] = 1 - c_0$ is not conserved and decreases over time. With this definition, the system (EDG) becomes

$$\begin{cases} \dot{c}_0 = M_\lambda[c]\ c_1, \\ \dot{c}_1 = M_\lambda[c]\ \big(-2c_1 + 2^\lambda c_2\big), \\ \dot{c}_k = M_\lambda[c]\ \big((k-1)^\lambda c_{k-1} - 2k^\lambda c_k + (k+1)^\lambda c_{k+1}\big), \quad k \geq 2. \end{cases} \tag{EDG$_\lambda$}$$

The first question regarding to coarsening is the large-time behavior of the average cluster size among living clusters, which plays the role of the characteristic length-scale. Intuitively, this quantity should grow in time. Indeed, by conservation of mass, the average cluster size, denoted $\ell(t)$, is given by

$$\ell(t) = \frac{1}{1 - c_0(t)} \sum_{k=1}^{\infty} k\, c_k(t) = \frac{\rho}{M_0[c]}, \tag{3.5}$$

hence the length-scale of the system is inversely proportional to the volume of living particles, which decreases by equation (EDG$_\lambda$). More specifically, the scaling analysis in [7] predicts that

$$\ell(t) \propto \begin{cases} t^\beta, & \text{if } 0 \leq \lambda < 3/2, \\ \exp(Ct), & \text{if } \lambda = 3/2, \\ (t_{\text{gel}} - t)^\beta, & \text{if } 3/2 < \lambda < 2, \end{cases} \qquad \text{with } \beta = (3-2\lambda)^{-1}, \tag{3.6}$$

and the involved constants C, t_{gel} and the one in $\propto$ depend on the initial data. Hence, coarsening is expected on an algebraic timescale for $0 \le \lambda < 3/2$, transitioning into a gelation regime for $3/2 < \lambda < 2$, where the solution only exists up to the gelation time $t_{\mathrm{gel}} < \infty$ at which all the mass vanishes to infinity. At the transition $\lambda = 3/2$ solutions exist globally and we expect coarsening on an exponential timescale with a non-universal rate C. Our first result confirms these coarsening rates.

Theorem 3.1 (Coarsening rates). *Let $0 \le \lambda < 2$ and set $\beta = (3 - 2\lambda)^{-1}$. Then the following statements hold, with all constants only depending on λ, ρ and moments of the initial data up to order λ:*

1. *If $0 \le \lambda < 3/2$, then every solution c to equation (EDG$_\lambda$) exists globally and there are positive constants C_1, C_2, t_0 such that*

$$C_1 t^\beta \le \ell(t) \le C_2 t^\beta \qquad \text{for all } t \ge t_0.$$

2. *Let $\lambda = 3/2$, then every solution c to equation (EDG$_\lambda$) exists globally and there are positive constants C_1, C_2, K_1, K_2, t_0 such that*

$$K_1 \exp(C_1 t) \le \ell(t) \le K_2 \exp(C_2 t) \qquad \text{for all } t \ge t_0.$$

3. *If $3/2 < \lambda \le 2$, then every solution c to equation (EDG$_\lambda$) exists only locally on a maximal interval $[0, t^*)$ for some $t^* > 0$ and there are positive constants C_1, C_2, t_0 such that*

$$C_1 (t^* - t)^\beta \le \ell(t) \le C_2 (t^* - t)^\beta \qquad \text{for all } t_0 \le t < t^*.$$

Remark 3.2. *In the case $3/2 < \lambda \le 2$ it is easy to see from the proof (see Proposition 3.28) that the blow up time t^* goes to zero as the λ-th moment of the initial data diverges.*

The next question is whether solutions become self-similar as $t \to \infty$, which is formally addressed in [7]. The crucial observation is, that (EDG$_\lambda$) becomes a discrete linear weighted heat equation after a suitable non-autonomous time-change. Considering the corresponding weighted heat equation on the continuum scale (see Section 3.1.3) and formal scaling argument, the calculations in [7] suggest that any solution c with mass $M_1[c] = \rho$ is asymptotically self-similar to a profile $g_\lambda : [0, \infty) \to (0, \infty)$ for a suitable scaling function $s(t) \propto \ell(t)$ of the form (3.6). In mathematical terms, we expect that that following relation holds

$$c_k(t) \propto \rho\, s(t)^{-2} g_\lambda\Big(s(t)^{-1} k\Big) \qquad \text{for } t \gg 1.$$

Hereby, the profile g_λ is explicitly given by

$$g_\lambda(x) = \frac{1}{Z_\lambda} \frac{x^{1-\lambda}}{2-\lambda} \exp\bigg(-\frac{x^{2-\lambda}}{(2-\lambda)^2}\bigg), \tag{3.7}$$

where Z_λ is a normalization constant such that $\int_{[0,\infty)} x\, g_\lambda(x)\,\mathrm{d}x = 1$ and given by

$$Z_\lambda = (2-\lambda)^{\frac{2}{2-\lambda}}\,\Gamma\Big(1+\frac{1}{2-\lambda}\Big). \tag{3.8}$$

The appropriate object for the rigorous analysis of self-similarity is the empirical measure associated to a solution c given by

$$\mu_c(t) = s(t)\sum_{k\geq 1} c_k(t)\delta_{s(t)^{-1}k}. \tag{3.9}$$

The normalization in (3.9) is chosen, such that

$$M_0[c] = 1 - c_0 = s^{-1}(t)\int_0^\infty \mathrm{d}\mu_c \qquad \text{and} \qquad M_1[c] = \int_0^\infty x\,\mathrm{d}\mu_c. \tag{3.10}$$

Self-similar behavior of c for $t\to\infty$ then corresponds to the existence of the limit $\mu_c(t)\to\rho\, g_\lambda$ in a suitable topology, which we define now. From here on, we use the notation $\mathbb{R}_+ = (0,\infty)$ and $\overline{\mathbb{R}}_+ = [0,\infty)$. In addition to the weak convergence of measures in $\mathcal{M}(\overline{\mathbb{R}}_+)$, written as $\mu_n \rightharpoonup \mu$, and defined as

$$\lim_{n\to\infty}\int_{\overline{\mathbb{R}}_+} f(x)\,\mathrm{d}\mu_n(x) = \int_{\overline{\mathbb{R}}_+} f(x)\,\mathrm{d}\mu(x) \qquad \text{for all } f\in C_b^0(\overline{\mathbb{R}}_+),$$

we need two further weak convergence concepts adjusted to the problem setting.

Definition 3.3. *Let $\mathcal{M}_1(\overline{\mathbb{R}}_+)$ be the space of all Borel measures on $\overline{\mathbb{R}}_+$ with finite first moment, that is $\int_0^\infty x\,\mathrm{d}\mu<\infty$ for all $\mu\in\mathcal{M}_1(\overline{\mathbb{R}}_+)$. The space of continuous sublinearly growing functions $\mathcal{C}$ and its subspace $\mathcal{C}_0$ with those vanishing at 0 are defined by*

$$\mathcal{C} = \Big\{f\in C^0(\overline{\mathbb{R}}_+)\colon \lim_{x\to\infty} x^{-1}f(x) = 0\Big\} \qquad \textit{and} \qquad \mathcal{C}_0 = \{f\in\mathcal{C}\colon f(0)=0\}.$$

A family of measures $\mu_n\in\mathcal{M}_1(\overline{\mathbb{R}}_+)$ converges weakly to $\mu\in\mathcal{M}_1(\overline{\mathbb{R}}_+)$ with respect to $\mathcal{C}$, denoted by $\mu_n\rightharpoonup\mu$, if

$$\lim_{n\to\infty}\int_{\overline{\mathbb{R}}_+} f(x)\,\mathrm{d}\mu_n(x) = \int_{\overline{\mathbb{R}}_+} f(x)\,\mathrm{d}\mu(x) \qquad \textit{for all } f\in\mathcal{C}.$$

Likewise, $\mu_n\rightharpoonup\mu$ with respect to $\mathcal{C}_0$, if the above limit holds for all $f\in\mathcal{C}_0$.

With this definition, we prove weak convergence to the self-similar profile with an explicit scaling function except at the transition $\lambda = 3/2$, where the scaling function can be described asymptotically. Furthermore, the weak convergence is with respect to $\mathcal{C}$ for $\lambda\in[0,1)$ and with respect to $\mathcal{C}_0$ if $\lambda\geq 1$ due to technical reason, see Remark 3.5.

Theorem 3.4 (Self-similar behavior)**.** *Let $\rho > 0$.*

1. *For $0 \le \lambda < 3/2$ there exists $C = C(\lambda, \rho) > 0$ and a corresponding scaling function*

$$s(t) = Ct^{\beta} \qquad \text{with } \beta = (3 - 2\lambda)^{-1},$$

such that every global solution c to equation (EDG$_\lambda$) with $M_1[c] = \rho$ converges

$$\begin{cases} \mu_c(t) \rightharpoonup \rho\, g_\lambda & \text{with respect to } \mathcal{C} & \text{if } \lambda \in [0,1) \\ \mu_c(t) \rightharpoonup \rho\, g_\lambda & \text{with respect to } \mathcal{C}_0 & \text{if } \lambda \in [1, 3/2) \end{cases} \qquad \text{as } t \to \infty. \tag{3.11}$$

2. *For $\lambda = 3/2$ there exists a scaling function $s : \overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}_+$ and a constant $C = C(\rho)$ such that for every $\varepsilon > 0$ it holds*

$$\lim_{t\to\infty} \exp(-(C+\varepsilon)t)\, s(t) = 0 \qquad \text{and} \qquad \lim_{t\to\infty} \exp(-(C-\varepsilon)t)\, s(t) = \infty,$$

and every global solution c to equation (EDG$_\lambda$) with $M_1[c] = \rho$ converges

$$\mu_c(t) \rightharpoonup \rho\, g_\lambda \qquad \text{with respect to } \mathcal{C}_0 \qquad \text{as } t \to \infty. \tag{3.12}$$

3. *For $3/2 < \lambda < 2$ and t^* as in Theorem 3.1 (3) there exists $C = C(\lambda, \rho) > 0$ such that for the scaling function*

$$s(t) = C(t^* - t)^{\beta} \qquad \text{with } \beta = (3 - 2\lambda)^{-1},$$

every solution c to equation (EDG$_\lambda$) existing on the finite time interval $[0, t^*)$, converges

$$\mu_c(t) \rightharpoonup \rho\, g_\lambda \qquad \text{with respect to } \mathcal{C}_0 \qquad \text{as } t \to t^*. \tag{3.13}$$

Remark 3.5. *Note that the difference between the weak convergence with respect to $\mathcal{C}$ in comparison to the one with respect to $\mathcal{C}_0$ in Definition 3.3 is that in the latter a Dirac measure at 0 might occur. The reason why we can only prove convergence with respect to $\mathcal{C}_0$ in the case $\lambda \ge 1$ is that the analysis relies on an energy method involving a discrete version (3.19) of the weighted H^1-seminorm*

$$\mathcal{E}_\lambda(f) = \int_0^\infty x^\lambda |f'(x)|^2 \,\mathrm{d}x,$$

for which the corresponding embedding into $C^{0,\frac{1-\lambda}{2}}([0,\infty))$ only holds for $\lambda < 1$, while for $\lambda \ge 1$ the modulus of continuity is only controlled away from $x = 0$, see Lemma 3.24 and Proposition 3.36. We conjecture that the weak convergence in Theorem 3.4 in fact holds with respect to $\mathcal{C}$ for all $\lambda \in [0,2)$. Our results still imply that the total variation of μ_c, and hence the size of the Dirac, is a priori bounded from above in terms of moments of the initial data.

3.1.2. Time-change and tail distribution

The common factor M_λ in (EDG$_\lambda$) is eliminated through the time change

$$\tau(t) = \int_0^t M_\lambda[c](s)\,\mathrm{d}s. \tag{3.14}$$

Since the right-hand-side of (EDG$_\lambda$) never contains c_0 which can simply be obtained from the conservation law (3.10), c_0 is ignored in the following considerations. Consequently, we define $u(\tau(t),k) = c_k(t)$ for $k \in \mathbb{N}$. Since $a_\lambda(0) = 0$ for all $\lambda \in [0,2)$, the value of $u(\tau(t),0)$ is not specified. Nevertheless, it is convenient to set it to zero $u(\tau,0) = 0$ for all $\tau \geq 0$. We see that u solves the equation

$$\frac{\mathrm{d}}{\mathrm{d}\tau}u(\tau,k) = (k-1)^\lambda u(\tau,k-1) - 2k^\lambda u(\tau,k) + (k+1)^\lambda u(\tau,k+1), \quad k \geq 1,$$

which can be written such that it takes the form of a spatially discrete heat equation with Dirichlet boundary condition

$$\begin{cases} \partial_\tau u = \Delta_{\mathbb{N}}(a_\lambda u), & k \geq 1, \\ (a_\lambda u)(\tau,0) = 0, & \tau \geq 0. \end{cases} \tag{DP}$$

The case $\lambda = 0$ can be treated explicitly (see Appendix A.3). Here, the discrete Laplacian $\Delta_{\mathbb{N}}$ is conveniently expressed by the discrete differential operators

$$\partial^- u(k) = u(k) - u(k-1) \quad \text{and} \quad \partial^+ u(k) = u(k+1) - u(k), \quad \text{for } k \geq 1, \tag{3.15}$$

such that it holds $\Delta_{\mathbb{N}} = \partial^-\partial^+$. An elementary calculation shows that the discrete differential operators satisfy a version of the integration by parts formula

$$\sum_{k=a}^{b} \partial^+ u(k)v(k) = u(b+1)v(b) - u(a)v(a-1) - \sum_{k=a}^{b} u(k)\partial^- v(k). \tag{3.16}$$

For brevity, we abuse notation and subsequently write $u = u(t,k)$. If u is a solution of equation (DP) we can directly calculate the evolution of the moments

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}M_0 &= \sum_{k=1}^{\infty} \Delta_{\mathbb{N}}(a_\lambda u) = -u(t,1) \leq 0, \\ \frac{\mathrm{d}}{\mathrm{d}t}M_1 &= \sum_{k=1}^{\infty} k\Delta_{\mathbb{N}}(a_\lambda u) = 0. \end{aligned}$$

The first identity is highlighting the fact, that c_0 is omitted and the second identity is the conservation of total mass. It is more convenient to study not the equation for u, but the one of the tail distribution U associated to u given by

$$U(t,k) = \sum_{l \geq k} u(t,l), \qquad \text{for } k \geq 1. \tag{3.17}$$

Again, the value of $U(t,0)$ is not specified, but by the convention $u(t,0) = 0$, we obtain $U(t,0) = U(t,1)$, which we interpret as Neumann boundary condition. The main motivation is that the evolution operator from (DP) becomes a weighted Laplace operator L_λ in divergence form defined on the Hilbert space $\ell^2(\mathbb{N})$:

$$L_\lambda U(k) = \partial^-(a_\lambda \partial^+ U)(k). \tag{3.18}$$

It is obvious by the integration by parts formula (3.16) that L_λ is symmetric and negative semi-definite with Dirichlet form given by

$$E_\lambda(U,V) = \langle V, -L_\lambda V\rangle_2 = \sum_{k=1}^{\infty} k^\lambda\, \partial^+ U\, \partial^+ V. \tag{3.19}$$

We also write $E_\lambda(U) = E_\lambda(U,U)$. Now, let u be a solution to (DP) (cf. Corollary 3.18 for well-posedness) and U as in (3.17), then U solves the Neumann problem

$$\begin{cases} \partial_t U = L_\lambda U, \\ (a_\lambda \partial^+ U)(t,0) = 0. \end{cases} \tag{NP}$$

Indeed, for $k \geq 1$ we calculate

$$\begin{aligned} \partial_t U(t,k) &= \sum_{l\geq k} \Delta_{\mathbb{N}}(a_\lambda u)(l) = -\partial^-(a_\lambda u)(k), \\ \partial^+ U(t,k) &= \sum_{l\geq k+1} u(t,l) - \sum_{l\geq k} u(t,l) = -u(t,k). \end{aligned}$$

Furthermore, we find that

$$\sum_{k\geq 1} U(t,k) = \sum_{k\geq 1}\sum_{l\geq k} u(t,l) = \sum_{l\geq 1} u(t,l) \sum_{k\leq l} = \sum_{l\geq 1} u(t,l) l = M_1[u],$$

which shows that $M_0[U] = M_1[u] = \text{const}$. A formal dimensional analysis suggests that $k \propto t^\alpha$, where

$$\alpha = \frac{1}{2-\lambda} \in \left[\frac{1}{2}, \infty\right), \tag{3.20}$$

which corresponds for $\lambda = 0$ to the classical parabolic scaling. In the following identities and estimates, we see that occurrences of square-roots in the classical parabolic theory are replaced by the exponent α. At the heart of the analysis of equation (NP) is the following discrete Nash-inequality, which connects the Dirichlet form (3.19) to the L^1 and L^2-norm of U.

Proposition 3.6 (Discrete Nash-inequality)**.** *Let $\lambda \in [0,2)$. Then for all $U \in \ell^2(\mathbb{N})$ with $E_\lambda(U) < \infty$ it holds*

$$\|U\|_2^2 \lesssim \|U\|_1^{\frac{2(2-\lambda)}{3-\lambda}} E_\lambda(U)^{\frac{1}{3-\lambda}}. \tag{DNI}$$

Here and in the following, we use the notation $A \lesssim B$ if there is a numerical constant $C = C(\lambda) > 0$ independent of all other parameters such that $A \le CB$. We write $A \approx B$ if $A \lesssim B$ and $B \lesssim A$.

The above discrete Nash-inequality enables us to obtain the optimal decay rates of the L^2, L^∞ norms and of the Dirichlet energy E_λ for the fundamental solution to equation (NP). These estimates imply scaling-invariant decay and continuity estimates in time and space for general solutions.

Theorem 3.7 (Decay and continuity). *Let U be a solution to* (NP), *then*

$$\|U(t,\cdot)\|_\infty \lesssim \|U_0\|_1 (1+t)^{-\alpha}. \tag{3.21}$$

Moreover, there exist explicit continuous functions (see Lemma 3.24) $\theta_\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ and $\omega_\lambda : [1,\infty) \to \mathbb{R}_+$ with $\omega_\lambda(1) = 0$ such that

$$|U(t,k_2) - U(t,k_1)| \lesssim \|U_0\|_1 t^{-\alpha} \Big|\theta_\lambda(t^{-\alpha}k_2) - \theta_\lambda(t^{-\alpha}k_1)\Big|^{\frac{1}{2}}, \tag{3.22}$$

$$|U(t,k) - U(s,k)| \lesssim \|U_0\|_1 s^{-\alpha} \omega_\lambda\Big({}^{t}\!/_{s}\Big), \tag{3.23}$$

for all $k, k_1, k_2 \in \mathbb{N}$ and all $0 < s \le t$.

The estimate (3.21) translates to the lower coarsening bound from Theorem 3.1, as the length-scale ℓ is inversely proportional to the zero moment. Together with the continuity estimates the decay of solutions also provides the necessary compactness to prove self-similarity in Theorem 3.4.

3.1.3. Continuum equation and scaling solution

The space-continuous analogue of equation (NP) is

$$\begin{cases} \partial_t \varphi = \partial_x (a_\lambda \partial_x \varphi) = \mathcal{L}_\lambda \varphi, & (t,x) \in \mathbb{R}_+^2, \\ a_\lambda \partial_x \varphi|_{x=0} = 0, & t \in \mathbb{R}_+, \\ \varphi(0,\cdot) = \varphi_0, & x \in \mathbb{R}_+. \end{cases} \tag{NP'}$$

where the boundary condition is the natural one for the equation to conserve the zero moment. For $\lambda = 0$, we get the classical Neumann boundary condition $\partial_x \varphi|_{x=0} = 0$ for the heat equation, while for $\lambda > 0$ the coefficient a_λ vanishes at 0 and solutions are in general not smooth up to the boundary. The degeneracy has also the effect, that the boundary condition is only imposed for $\lambda < 1$, that is in the range $\lambda \ge 1$ the equation (NP') is well-posed without any boundary condition (see Remark 3.29). By the conservation of the zero moment, it is natural to look for a (normalized) scaling solution

$$\gamma_\lambda(t,x) = t^{-\alpha} \mathcal{G}_\lambda(t^{-\alpha}x), \tag{3.24}$$

for some profile $\mathcal{G}_\lambda$. Plugging this ansatz into the equation leads by simple calculations as in [7, Ch. III] to

$$\mathcal{G}_\lambda(x) = Z_\lambda^{-1} \exp\Big(-\alpha^2 x^{2-\lambda}\Big), \tag{3.25}$$

where Z_λ is chosen such that $\mathcal{G}_\lambda$ has integral 1 and explicitly given in (3.8). In general a solution to equation (NP') with given initial data can be constructed using the associated fundamental solution $\Psi_\lambda(t,x,y)$ (see Proposition 3.30).

Our definition of solutions in Section 3.3.1 immediately entails that the scaling solution γ_λ from (3.24) with $\mathcal{G}_\lambda$ given in (3.25) solves (NP') in a suitable measure-valued sense. To compare solutions on $\mathbb{N}$ of (NP) with those on $\overline{\mathbb{R}}_+$ of (NP'), we introduce for $\varepsilon > 0$ the following operations between discrete measures on $\mathbb{N}$ and measures on $\overline{\mathbb{R}}_+$

$$\iota_\varepsilon : \mathcal{M}(\mathbb{N}) \to \mathcal{M}\big(\overline{\mathbb{R}}_+\big) \qquad \text{with} \qquad (\iota_\varepsilon U)(x) = \varepsilon^{-\alpha} U(\lfloor \varepsilon^{-\alpha} x \rfloor + 1), \tag{3.26}$$

$$\pi_\varepsilon : \mathcal{M}\big(\overline{\mathbb{R}}_+\big) \to \mathcal{M}(\mathbb{N}) \qquad \text{with} \qquad (\pi_\varepsilon \mu)(k) = \mu\Big([(k-1)\varepsilon^{-\alpha}, k\varepsilon^{-\alpha})\Big). \tag{3.27}$$

Note that we have $\pi_\varepsilon \circ \iota_\varepsilon = \mathrm{id}$, and both operations are adjoint to each other in the sense that for $U \in \mathcal{M}(\mathbb{N})$ and $\mu \in \mathcal{M}(\overline{\mathbb{R}}_+)$ it holds

$$\int_0^\infty \iota_\varepsilon U(x)\,\mathrm{d}\mu(x) = \varepsilon^{-\alpha} \sum_{k=1}^\infty U(k)\pi_\varepsilon\mu(k).$$

As a consequence, the maps are mass-conserving, i.e

$$\int_0^\infty \iota_\varepsilon U(x)\,\mathrm{d}x = \sum_{k=1}^\infty U(k) \qquad \text{and} \qquad \sum_{k=1}^\infty \pi_\varepsilon\mu(k) = \mu(\overline{\mathbb{R}}_+).$$

Now let U_ε be a sequence of solutions to equation (NP) with initial data $U_{0,\varepsilon}$. We define the sequence of functions $\mathcal{U}_\varepsilon$ by

$$\mathcal{U}_\varepsilon(t,x) = (\iota_\varepsilon U_\varepsilon)(\varepsilon^{-1}t, x) = \varepsilon^{-\alpha} U_\varepsilon(\varepsilon^{-1}t, \lfloor \varepsilon^{-\alpha} x \rfloor + 1). \tag{3.28}$$

Then we have the following convergence result.

Theorem 3.8 (Convergence of the tail distribution)**.** *Let $\mathcal{U}_\varepsilon$ as above and assume that $\|U_{0,\varepsilon}\|_1$ is bounded and $\mathcal{U}_\varepsilon(0,\cdot) \rightharpoonup \mu_0$ as $\varepsilon \to 0$ for some $\mu_0 \in \mathcal{M}(\overline{\mathbb{R}}_+)$. Then there exists a unique global-in-time weak solution $\mathcal{U}$ to equation (NP') with initial data μ_0 and it holds*

1. *If $0 \le \lambda < 1$, $\mathcal{U}_\varepsilon \to \mathcal{U}$ locally uniformly on $\mathbb{R}_+ \times \overline{\mathbb{R}}_+$.*
2. *If $1 \le \lambda < 2$, $\mathcal{U}_\varepsilon \to \mathcal{U}$ locally uniformly on $\mathbb{R}_+^2$ and*

$$\sup_{0<\varepsilon\le 1} \mathcal{U}_\varepsilon(t,0) < \infty \qquad \textit{for all } t > 0.$$

The precise definition of weak solutions will be given in Section 3.3. If $U_{0,\varepsilon} = U_0$ for some $U_0 \in \ell^1_+(\mathbb{N})$ with $M_0[U] = \rho$, then it is easy to check that $\mathcal{U}_\varepsilon(0,\cdot) \rightharpoonup \|U_0\|_1\delta_0 = \rho\,\delta_0$. Hence, by applying Theorem 3.8, we get that $\mathcal{U}_\varepsilon$ converges to a multiple of the solution starting from δ_0, which is the scaling solution $\gamma(t,x) = t^{-\alpha}\mathcal{G}_\lambda(t^{-\alpha}x)$ defined in (3.24). In particular for $t = 1$, we have

$$\mathcal{U}_\varepsilon(1,x) = \varepsilon^{-\alpha} U(\varepsilon^{-1}, \lfloor \varepsilon^{-\alpha} x \rfloor + 1) \to \rho\,\gamma(1,x) = \rho\,\mathcal{G}_\lambda(x).$$

Hence, the scaling limit in fact implies long-time behavior after setting $t = \varepsilon^{-1}$.

Corollary 3.9. *Let U be a solution to equation (NP) with $M_0[U] = \rho$. Then the rescaled function $\hat{U}(t,x) = t^\alpha U(t, \lfloor t^\alpha x \rfloor + 1)$ is locally uniformly bounded on $\mathbb{R}_+ \times \overline{\mathbb{R}}_+$ and the following holds:*

1. *If $0 \le \lambda < 1$, $\hat{U}(t,\cdot) \to \rho\, \mathcal{G}_\lambda$ as $t \to \infty$ locally uniformly on $\overline{\mathbb{R}}_+$.*
2. *If $1 \le \lambda < 2$, $\hat{U}(t,\cdot) \to \rho\, \mathcal{G}_\lambda$ as $t \to \infty$ locally uniformly on $\mathbb{R}_+$ and*

$$\limsup_{t\to\infty} \hat{U}(t,0) < \infty.$$

Outline

The rest of this chapter is organized as follows. In Section 3.2 we analyze solutions to equation (NP) and deduce results for equations (DP) and (EDG_λ). The main results which are crucial throughout the chapter are the L^∞-decay and continuity estimates for solutions to (NP) (Section 3.2.3), which utilize the discrete weighted Nash-inequality (DNI) proved in Section 3.2.2. The usefulness of these estimates lies in the fact that they are optimal in terms of the scaling $k \sim t^\alpha$. On the level of equation (DP), the L^∞-decay translates to a decay estimate for the zero moment, which is synonymous with the coarsening rate. Proving further scale-characteristic bounds for the moments (Section 3.2.4) we obtain estimates for the time change τ as in (3.14) which allows to relate the analysis to equation (EDG_λ) and obtain Theorem 3.1.

In Section 3.3 we prove the discrete-to-continuous scaling limit. First, we give the explicit construction of the fundamental solution of the continuum problem (NP'), which is possible via a suitable change of variables relating the evolution to the explicitly analyzable Bessel process. Decay and regularity properties of solutions can be read off from the explicit fundamental solution. Moreover, it allows us to define a sensible notion of weak solution for equation (NP') in terms of the adjoint equation, which includes the scaling solution γ_λ and has a built-in uniqueness property. The proof of Theorem 3.8 is given in Section 3.3.2 and relies on a replacement estimate for the defect between the discrete operator L_λ and the continuous operator $\mathcal{L}_\lambda$ (Section 3.3.3) which yields that the rescaled discrete solutions $\mathcal{U}_\varepsilon$ are approximate weak solutions to equation (NP'). Here, the technical part is due to the degeneracy of the equation, which implies that the test functions are not smooth at $x = 0$ (Section 3.3.4). Using this approximation property and the compactness inherited from the scale-invariant decay and continuity estimates, one can pass to the limit and obtain Theorem 3.8.

To prove Theorem 3.4, we relate the scaling-limit to the long-time behavior of the empirical measures associated with solutions to equation (DP) in Section 3.4. In particular, the scaling limit implies precise asymptotics for the moments which translate to asymptotics for the time change τ and allow us to obtain the self-similar behavior for solutions to (EDG_λ) with explicit scaling function.

3.2. Analysis of the discrete equations

3.2.1. Well-posedness

Before analyzing properties of solutions, we first collect some well-posedness results. For this we specify the notions of solutions for all three equations. We define suitable weighted $\ell^1(\mathbb{N})$-spaces by setting for $\mu \geq 0$

$$X_\mu(\mathbb{N}) = \{u \in \ell^\infty(\mathbb{N})\colon \|u\|_{X_\mu} < \infty\}, \qquad \text{with} \qquad \|u\|_{X_\mu} = M_\mu\big[|u|\big],$$
$$X_\mu^+(\mathbb{N}) = \{u \in X_\mu\colon u \geq 0\},$$

and define $X_\mu(\mathbb{N}_0), X_\mu^+(\mathbb{N}_0)$ in the obvious way. We consider all of the above spaces as Banach spaces equipped with the norm $\|\cdot\|_{X_\mu}$. When there is no danger of confusion, we just write X_μ, X_μ^+.

Definition 3.10 (Solutions to (EDG_λ))**.** *Let $\lambda \in [0,2)$, $c^{(0)} \in X^+_{\max(1,\lambda)}(\mathbb{N}_0)$ and $T \in (0,\infty]$. Then $c = c_k(t)\colon [0,T) \to X^+_{\max(1,\lambda)}(\mathbb{N}_0)$ is a solution to equation (EDG_λ) with initial data $c^{(0)}$ and kernel given in (3.3) provided that:*

1. *It holds $t \mapsto \|c(t,\cdot)\|_{X_{\max(1,\lambda)}} \in L^\infty_{\mathrm{loc}}([0,T))$.*
2. *For every $k \geq 0$ holds $t \mapsto c_k(t) \in C^0([0,T))$ and*
$$c_k(t) = c_k^{(0)} + \int_0^t \mathrm{EDG}_\lambda[c](s,k)\,\mathrm{d}s,$$
where $\mathrm{EDG}_\lambda[c]$ denotes the right-hand side in (EDG_λ).

Definition 3.11 (Solutions to (DP))**.** *Let $\lambda \in [0,2)$, $\mu \geq 1$, $u_0 \in X_\mu(\mathbb{N})$ and $T \in (0,\infty]$. Then $u = u(t,k)\colon [0,T) \to X_\mu$ is a solution to equation (DP) in X_μ with initial data u_0 if the following holds:*

1. *It holds $t \mapsto \|u(t,\cdot)\|_{X_\mu} \in L^\infty_{\mathrm{loc}}([0,T))$.*
2. *For every $k \in \mathbb{N}$ holds $t \mapsto u(t,k) \in C^0([0,T))$ and*
$$u(t,k) = u_0(k) + \int_0^t \Delta_{\mathbb{N}}(a_\lambda u)(s,k)\,\mathrm{d}s.$$

Definition 3.12 (Solutions to (NP))**.** *Let $\lambda \in [0,2)$, $\mu \geq 0$, $U_0 \in X_\mu(\mathbb{N})$ and $T \in (0,\infty]$. Then $U = U(t,k)\colon [0,T) \to X_\mu$ is a solution to equation (NP) in X_μ with initial data U_0 if the following holds:*

1. *It holds $t \mapsto \|U(t,\cdot)\|_{X_\mu} \in L^\infty_{\mathrm{loc}}([0,T))$.*
2. *For every $k \in \mathbb{N}$ holds $t \mapsto U(t,k) \in C^0([0,T))$ and*
$$U(t,k) = U_0(k) + \int_0^t L_\lambda U(s,k)\,\mathrm{d}s. \tag{3.29}$$

In all three cases we call solutions *global solutions* in the case $T = \infty$ and *local solutions* in the case $T < \infty$. Note that the integral formulations for u and U imply that solutions are indeed smooth and (DP), respectively (NP) hold pointwise, while a solution to (EDG_λ) is a priori only Lipschitz continuous. We start with the well-posedness of (NP) and then deduce corresponding results for the other two equations, see Corollaries 3.18 and 3.19.

Proposition 3.13. *For $\mu \geq 0$ and $U_0 \in X_\mu(\mathbb{N})$ exists a unique global solution U to equation (NP) in X_μ with initial data U_0 given by the representation formula*

$$U(t,k) = \sum_{l=1}^{\infty} \Phi(t,k,l)\, U_0(l),$$

where $\Phi = \Phi(t,k,l)$ is the fundamental solution, i.e $\Phi(\cdot,\cdot,l)$ is the solution to equation (NP) in X_μ for all $\mu \geq 0$ with initial data $\Phi(0,k,l) = \delta_{kl}$.

The proof of Proposition 3.13 is split into several auxiliary results. First, we provide a technical Lemma, which in its full scope is not needed in the existence proof, but plays a crucial role in the derivation of moment estimates later.

Lemma 3.14. *1. For all $\mu \geq 0$ and $\lambda \in [0,2)$, it holds*

$$\frac{L_\lambda(k^\mu)}{k^{\mu+\lambda-2}} \to \mu(\mu+\lambda-1), \qquad \text{as } k \to \infty.$$

2. For all $\mu > 0$ and $\lambda \in [1,2)$, there exists a positive constant $C = C(\mu,\lambda) \geq 1$ such that

$$C^{-1}k^{\mu+\lambda-2} \leq L_\lambda(k^\mu) \leq Ck^{\mu+\lambda-2}, \qquad \text{for all } k \in \mathbb{N}.$$

Proof. We calculate

$$\begin{aligned} L_\lambda(k^\mu) &= k^\lambda\Big((k+1)^\mu - k^\mu\Big) - (k-1)^\lambda\Big(k^\mu - (k-1)^\mu\Big) \\ &= k^{\mu+\lambda}\Big((1+k^{-1})^\mu - 1 + (1-k^{-1})^\lambda\Big((1-k^{-1})^\mu - 1\Big)\Big) = k^{\mu+\lambda} f_{\mu,\lambda}(k^{-1}), \end{aligned}$$

where $f_{\mu,\lambda}(x) = (1+x)^\mu - 1 + (1-x)^\lambda((1-x)^\mu - 1)$. To show the first statement, a simple calculation gives that

$$f_{\mu,\lambda}(0) = 0 = f'_{\mu,\lambda} \qquad \text{and} \qquad f''_{\mu,\lambda}(0) = 2\mu(\mu+\lambda-1),$$

hence $f_{\mu,\lambda}(x)x^{-2} \to \mu(\mu+\lambda-1)$ as $x \to 0$, which gives $L_\lambda(k^\mu)k^{2-\lambda-\mu} \to \mu(\mu+\lambda-1)$ as $k \to \infty$. The second statement also follows from the asymptotic behavior of $f_{\mu,\lambda}$ if we can show in addition that $f_{\mu,\lambda} > 0$ on $(0,1]$. For that end, we view $f_{\mu,\lambda}$ as a function of two variables $f_\lambda(\mu, x) = f_{\mu,\lambda}(x)$ with parameter λ. Note that we have $0 = f_\lambda(0, x) = f_\lambda(\mu, 0)$. We calculate the partial derivative for $\mu > 0$, $x \in (0, 1)$

$$\begin{aligned}\partial_\mu f_\lambda(\mu, x) &= \log(1+x)(1+x)^\mu + (1-x)^\lambda \log(1-x)(1-x)^\mu \\ &\geq \log(1+x)(1+x)^\mu + (1-x)\log(1-x)(1+x)^\mu \\ &= (1+x)^\mu(\log(1+x) + (1-x)\log(1-x)) = (1+x)^\mu g(x),\end{aligned}$$

where the lower bound follows from the fact that $\log(1-x) \leq 0$, $(1-x)^\lambda \leq (1-x)$ for $\lambda \geq 1$ and $(1-x)^\mu \leq (1+x)^\mu$ for $\mu > 0$. If we can show that $g(x) > 0$ for all $x \in (0, 1)$, then it follows from $f_\lambda(0, x) = 0$ that $f_\lambda(\mu, x) > 0$ for all $\mu > 0$, $x \in (0, 1]$. Calculating derivatives of g, we have

$$\begin{aligned}g'(x) &= (1+x)^{-1} - \log(1-x) - 1, \\ g''(x) &= -(1+x)^{-2} + (1-x)^{-1}, \\ g'''(x) &= 2(1+x)^{-3} + (1-x)^{-2}.\end{aligned}$$

Thus we have $g(0) = g'(0) = g''(0) = 0$ and $g''' > 0$ on $[0, 1)$, hence $g > 0$ on $[0, 1)$. □

Next, we prove a first existence result. To that end we introduce the following notation: For a function $f \colon \overline{\mathbb{R}}_+ \times \mathbb{N} \to \mathbb{R}$ we write $f \in C_c^\infty(\overline{\mathbb{R}}_+ \times \mathbb{N})$ if f satisfies the following properties:

1. There exists $N \in \mathbb{N}$ and $T \in \mathbb{R}$ such that $f(t, k) = 0$ if $t \geq T$ or $k \geq N$.

2. For every $k \in \mathbb{N}$ the map $t \mapsto f(t, k)$ is smooth.

Lemma 3.15. *Let $\mu \geq 0$, $U_0 \in X_\mu^+(\mathbb{N})$ and $f \in C_c^\infty(\overline{\mathbb{R}}_+ \times \mathbb{N})$ with $f \geq 0$. Then there exists a global solution U in X_μ^+ with initial data U_0 to the inhomogeneous equation*

$$\partial_t U - L_\lambda U = f, \tag{3.30}$$

in the sense of Definition 3.12 with (3.29) replaced by $U(t, k) = U_0(k) + \int_0^t (L_\lambda U + f)(s, k)\,\mathrm{d}s$.

Proof. For existence of solutions we apply standard regularization and truncation techniques. For $m > 0$ define $a_\lambda^{(m)}(k)$ to be

$$a_\lambda^{(m)}(k) = \begin{cases} k^\lambda, & k \leq m, \\ m^\lambda, & k > m. \end{cases}$$

Then $a_\lambda^{(m)}$ is bounded from above and below on $[1, \infty)$ and the corresponding elliptic operator is $L_\lambda^{(m)} = \partial^-(a_\lambda^{(m)}\partial^+)$. By standard arguments, there exists a unique non-negative solution $U^{(m)}$ to equation

$$\partial_t U^{(m)} - L_\lambda^{(m)} U^{(m)} = f$$

with initial data U_0 that satisfies $\sup_{0\le t\le T} M_\mu\Big[U^{(m)}(t,\cdot)\Big] < \infty$ for all $T \ge 0$ and $\mu \ge 0$ such that $M_\mu[U_0] < \infty$. Let $N \in \mathbb{N}$, then by using the discrete integration by parts (3.16), we arrive at

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\sum_{l=1}^{N} k^\mu U^{(m)} &= \sum_{l=1}^{N} k^\mu \partial_t U^{(m)} = \sum_{l=1}^{N} k^\mu L_\lambda^{(m)} U^{(m)} + \sum_{l=1}^{N} k^\mu f \\ &= \sum_{l=1}^{N} L_\lambda^{(m)}(k^\mu) U^{(m)} + \sum_{l=1}^{N} k^\mu f + R(t,N),\end{aligned}$$

where $R(t,N) = (N+1)^\mu a_\lambda^{(m)}(N)\partial^+ U^{(m)}(t,N) - \partial^+(k^\mu)(N) a_\lambda^{(m)}(N) U^{(m)}(t,N+1)$ are the boundary terms from the discrete integration by parts. Now for $k \le m$ we have $L_\lambda^{(m)}(k^\mu) = L_\lambda(k^\mu) \lesssim k^{\mu+\lambda-2} \lesssim k^\mu$ by Lemma 3.14, while for $k > m$ we have

$$L_\lambda^{(m)}(k^\mu) = m^\lambda \partial^- \partial^+(k^\mu) \lesssim m^\lambda k^{\mu-2} \le k^{\mu+\lambda-2} \lesssim k^\mu,$$

again by Lemma 3.14 and the fact that $L_0 = \partial^-\partial^+$. Thus we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{l=1}^{N} k^\mu U^{(m)} \lesssim M_\mu[U^{(m)}] + M_\mu[f] + R(t,N).$$

Next, we note that for fixed m, $R(t,N) \le C(m) M_\mu[U^{(m)}]$, hence R is bounded on each compact time interval. Moreover, because $M_\mu[U^{(m)}] < \infty$, we have $R(t,N) \to 0$ as $N \to \infty$ for every $t \ge 0$. Thus by integrating in time and letting $N \to \infty$, we obtain

$$M_\mu[U^{(m)}](t) \lesssim M_\mu[U_0] + \int_0^t M_\mu[f](s)\,\mathrm{d}s + \int_0^t M_\mu[U^{(m)}](s)\,\mathrm{d}s.$$

We conclude with Gronwall's lemma that $\sup_{0\le t\le T} M_\mu[U^{(m)}](t) \le C(T)$ independent of m, because of $f \in C_c^\infty(\overline{\mathbb{R}}_+ \times \mathbb{N})$. Next, we establish compactness of the sequence $U^{(m)}$. Because the above calculation holds in particular for M_0, we have that $U^{(m)}(t,k)$ is uniformly bounded on each compact time interval. We also have

$$\Big|L_\lambda^{(m)} U^{(m)}\Big|(k) \lesssim k^\lambda \|U^{(m)}\|_\infty \le k^\lambda M_0[U^{(m)}].$$

Hence, for each $k \in \mathbb{N}$ the time derivative of $U^{(m)}(t,k)$ is uniformly bounded on every bounded time interval. By using Arzela-Ascoli's theorem and standard diagonal arguments it is easy to see that for a subsequence we have $U^{(m)}(t,k) \to U(t,k)$ as $m \to \infty$ for every $k \in \mathbb{N}$ locally uniformly in time. Passing to the limit in the time integrated equation

$$U^{(m)}(t,k) = U_0(k) + \int_0^t \Big(L_\lambda^{(m)} U^{(m)} + f\Big)(s,k)\,\mathrm{d}s,$$

it follows that U is a solution to equation (3.30) with initial data U_0. Furthermore, by Fatou's lemma all moment bounds of $U^{(m)}$ carry over to U, which finishes the proof. $\square$

Corollary 3.16. *For $U_0 \in X_0(\mathbb{N})$ and any $T > 0$ exists at most one solution U to equation* (NP) *on $[0, T)$ in X_0 with initial data U_0.*

Proof. Because of linearity it suffices to show that every solution U in X_0 with initial data 0 is equal to 0. For that end, let $T > 0$, $l \in \mathbb{N}$, $\eta \in C_c^\infty((0, T))$ with $\eta \geq 0$ and let $f(t, k) = \eta(t)\delta_{lk}$. Then by Lemma 3.15 there exists a non-negative solution to the backwards equation

$$\partial_t V + L_\lambda V = -f,$$

on the interval $[0, T]$ with terminal data $V(T, \cdot) = 0$ and $\sup_{0 \leq t \leq T} M_\mu[V(t, \cdot)] < \infty$ for all $\mu \geq 0$. Multiplying equation (NP) for U with V, taking sums and integrating in time we obtain

$$\begin{aligned} 0 &= \int_0^T \sum_{k=1}^{\infty} V\partial_t U - V L_\lambda U \, \mathrm{d}t = -\int_0^T \sum_{k=1}^{\infty} (\partial_t V + L_\lambda V) U \, \mathrm{d}t \\ &= \int_0^T \sum_{k=1}^{\infty} f U \, \mathrm{d}t = \int_0^T \eta(t) U(t, l) \, \mathrm{d}t. \end{aligned}$$

Here we used integration by parts in time and space, which is easily justified using that U is a solution in X_0 and the moment bounds on V. Since $\eta(t)$ and l were arbitrary, we conclude that $U = 0$. $\square$

Based on the well-posedness result for (NP) from Lemma 3.15 and Corollary 3.16, we have a Green function representation for the solutions.

Corollary 3.17. *There exists a function $\Phi : \overline{\mathbb{R}}_+ \times \mathbb{N} \times \mathbb{N} \to \overline{\mathbb{R}}_+$ with the following properties*

1. *For all $\mu \geq 0$, $l \in \mathbb{N}$, $\Phi(\cdot, \cdot, l)$ is the unique global solution to equation* (NP) *in X_μ^+ with initial data $\Phi(0, k, l) = \delta_{kl}$.*

2. *For all $t \geq 0$, $k, l \in \mathbb{N}$ it holds $\Phi(t, k, l) = \Phi(t, l, k)$.*

3. *For all $t \geq 0$, $l \in \mathbb{N}$ it holds $M_0[\Phi(t, \cdot, l)] = 1$.*

4. *For all $\mu \geq 0$, $U_0 \in X_\mu(\mathbb{N})$ the function*

$$U(t, k) = \sum_{l=1}^{\infty} \Phi(t, k, l) U_0(l)$$

is the unique global solution to equation (NP) *in X_μ with initial data U_0 and $M_0[U(t, \cdot)] = M_0[U_0]$ for all $t \geq 0$.*

Proof. The existence and uniqueness of Φ follows directly from Lemma 3.15 and Corollary 3.16. The symmetry of Φ follows by using the same parabolic approximation as in the proof of Lemma 3.15, where symmetry is clear for the approximating functions and hence carries over to the limit. The fourth property is a direct consequence of Lemma 3.15. Indeed, since all moments of Φ are finite, the following calculations involving discrete integration by parts and differentiating under the sum can be rigorously justified to conclude that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\sum_{l=1}^{\infty}\Phi(t,k,l)U_0(l) &= \sum_{k=1}^{\infty}\partial_t\Phi(t,k,l)U_0(l)\\ &= \sum_{k=1}^{\infty}L_\lambda\Phi(t,k,l)U_0(l) = L_\lambda\left(\sum_{l=1}^{\infty}\Phi(t,k,l)U_0(l)\right),\end{aligned}$$

which proves the representation formula. Finally, from Lemma 3.14 we get the bound $L_\lambda(k^\mu) \le Ck^\mu$, which gives the bound

$$\frac{\mathrm{d}}{\mathrm{d}t}M_\mu[\Phi(t,\cdot,l)] = \sum_{k=1}^{\infty}L_\lambda(k^\mu)\Phi(t,\cdot,l) \le CM_\mu[\Phi(t,\cdot,l)].$$

Hence, we get $\sup_{0\le t\le T}M_\mu[\Phi(t,\cdot,l)] \lesssim C(T)l^\mu$. Then from the representation formula follows

$$M_\mu[|U(t,\cdot)|] \le \sum_{l=1}^{\infty}M_\mu[\Phi(t,\cdot,l)]|U_0(l)| \le C(T)M_\mu[|U_0|],$$

which finishes the proof. $\square$

As a consequence of the well-posedness result for equation (NP) we also obtain a well-posedness result for equation (DP), since both equations are linked by taking the discrete derivative, respectively anti-derivative.

Corollary 3.18. *Let $\mu \ge \max(1,\lambda)$, $u_0 \in X_\mu^+(\mathbb{N})$. Then there exists a unique global solution u to equation* (DP) *in X_μ^+ with initial data u_0. Furthermore, this solution satisfies*

$$\frac{\mathrm{d}}{\mathrm{d}t}M_0[u(t,\cdot)] = -u(t,1) \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}t}M_1[u(t,\cdot)] = 0.$$

Proof. Let U_0 denote the tail distribution associated to u_0. Then there exists a unique global solution to equation (NP) with initial data U_0 and $M_0[U(t,\cdot)] = M_0[U_0]$ for all $t \ge 0$. Then it is easily checked that $u(t,k) = -\partial^+U(t,k)$ is a solution to equation (DP) with initial data u_0 and $M_1[u(t,\cdot)] = M_0[U(t,\cdot)] = M_0[U_0] = M_1[u_0]$ by Corollary 3.17, and $\frac{\mathrm{d}}{\mathrm{d}t}M_0[u] = \frac{\mathrm{d}}{\mathrm{d}t}U(t,1) = \partial^+U(t,1) = -u(t,1)$. The bound for the higher moments also follows from Corollary 3.17, observing that $M_\mu[u]$ is comparable to $M_{\mu-1}[U]$, hence u is a solution in X_μ. The non-negativity of the solution u easily follows from the comparison principle for the discrete Laplacian, showing that $u(t,k)=0$ implies $\partial_t u(t,k) \ge 0$. For uniqueness, let u, v be two solutions to equation (DP) in X_μ with the same initial data. Then their tail distributions U, V are solutions to equation (NP) in $X_{\mu-1}$. Indeed, for $k, N \in \mathbb{N}$, $N > k$ we calculate

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{l=k}^{N}u(t,k) = \sum_{l=k}^{N}\Delta_{\mathbb{N}}(a_\lambda u)(t,k) = \partial^+(a_\lambda u)(t,N) - \partial^-(a_\lambda u)(t,k).$$

Hence, we have the representation

$$\sum_{l=k}^{N} u(t,k) = \sum_{l=k}^{N} u(0,k) + \int_0^t \partial^+(a_\lambda u)(s,N) - \partial^-(a_\lambda u)(s,k)\,\mathrm{d}s.$$

Now $\partial^+(a_\lambda u)(s,N) \le \sup_{0\le s\le t} M_\mu[u] < \infty$ and $\partial^+(a_\lambda u)(s,N) \to 0$ as $N \to \infty$ for all $s \in [0,t]$ because u is a solution in X_μ. Hence letting $N \to \infty$ in the above equality yields

$$U(t,k) = U(0,k) - \int_0^t \partial^-(a_\lambda u)(s,k)\,\mathrm{d}s = U_0(k) + \int_0^t \partial^-(a_\lambda \partial^+ U)(s,k)\,\mathrm{d}s,$$

which, together with the fact that $M_\mu[u]$ and $M_{\mu-1}[U]$ are comparable, shows that U is a solution to equation (NP) in $X_{\mu-1}$. We conclude that if u and v have the same initial data, then $U = V$ due to uniqueness for equation (NP) and thus $u = v$. □

By classical means it is straightforward to prove well-posedness of the equation (DP) also in $X_\mu^+(\mathbb{N})$ for $\mu \in [0,1]$. However, in our applications to (EDG$_\lambda$), the μ-moment with $\mu \ge \max\{1,\lambda\}$ is naturally appearing and we obtain from the above result the local well-posedness of (EDG$_\lambda$) for the full relevant range $\lambda \in [0,2)$.

Corollary 3.19. *Let $\lambda \in [0,2)$, $c^{(0)} \in X^+_{\max(1,\lambda)}(\mathbb{N}_0)$. Then there exists $T > 0$ and a solution c to equation (EDG$_\lambda$) on $[0,T)$ with initial data $c^{(0)}$. Furthermore, the following statements hold:*

1. *Any solution c to equation (EDG$_\lambda$) on a finite interval $[0,T)$ can be extended if*

$$\sup_{0\le t<T} M_\lambda[c(t,\cdot)] < \infty.$$

2. *Any solution c to equation (EDG$_\lambda$) on $[0,T)$ conserves for all $t \in [0,T)$*

$$\sum_{k=0}^{\infty} c_k(t) = \sum_{k=0}^{\infty} c_k(0) \qquad \text{and} \qquad \sum_{k=0}^{\infty} k c_k(t) = \sum_{k=0}^{\infty} k c_k(0).$$

Proof. Let u be the global solution to equation (DP) with initial data $u_0 = c^{(0)}$. Then by Corollary 3.18 we have that $M_{\max(1,\lambda)}[u]$ is bounded on each finite time interval, which allows to define

$$s : [0,\infty) \to [0,s^*) \qquad \text{with} \qquad s(t) = \int_0^t \frac{1}{M_\lambda[u](r)}\,\mathrm{d}r,$$

for some $s^* \in (0,\infty]$. By setting $c_k(s(t)) = u(t,k)$ for $k \ge 1$ and $c_0 = 1 - M_0[c]$, we obtain by straightforward calculus that $c_k(s)$ is a solution to equation (EDG$_\lambda$) on the time interval $[0,s^*)$. To show the statement regarding extension of solutions, if c

is a solution to equation (EDG$_\lambda$) on $[0,T)$ with $\sup_{0\le t<T} M_{\max(1,\lambda)}[c(t,\cdot)] < \infty$, then by equation (EDG$_\lambda$) we get that c_k is uniformly Lipschitz continuous on $[0,T)$ and hence there exists c_k^* such that $\lim_{t\to T} c_k(t) = c_k^*$. By Fatou's lemma we also have that $M_{\max(1,\lambda)}[c^*] < \infty$, hence $c^* \in X^+_{\max(1,\lambda)}$ and we can extend the solution by solving (EDG$_\lambda$) locally with initial data c^*. For the last statement, we recall that the time change

$$\tau(t) = \int_0^t M_\lambda[c(s,\cdot)]\,\mathrm{d}s$$

yields a (local) solution $u(\tau(t),k) = c_k(t)$ to equation (DP) in $X_{\max(1,\lambda)}$, and then by uniqueness for u the desired result follows from the identities for $M_0[u], M_1[u]$ from Corollary 3.18. □

3.2.2. Discrete Nash inequality: Proof of Proposition 3.6

The goal of this section is to prove the discrete Nash-type interpolation inequality in Proposition 3.6. We first give a proof of the continuous version of the Nash-inequality and then use it to prove the discrete version. As in the discrete case (3.19), we define the Dirichlet form for $f,g \in L^2(\overline{\mathbb{R}}_+)$ by

$$\mathcal{E}_\lambda(f,g) = \int_0^\infty |x|^\lambda f'(x)g'(x)\,\mathrm{d}x. \tag{3.31}$$

Proposition 3.20 (Continuous Nash-inequality)**.** *Let $\lambda \in [0,2)$. Then for all $f \in L^2(\overline{\mathbb{R}}_+)$ with $\mathcal{E}_\lambda(f) < \infty$ it holds*

$$\|f\|_2^2 \lesssim \|f\|_1^{\frac{2(2-\lambda)}{3-\lambda}} \mathcal{E}_\lambda(f)^{\frac{1}{3-\lambda}}. \tag{CNI}$$

Taking Proposition 3.20 for granted, we can now reduce the discrete Nash-inequality to the continuous case by considering the piecewise linear interpolation of the discrete function U.

Proof of Proposition 3.6. We define the function f by

$$f(x) = \begin{cases} U(1), & 0 \le x \le 1, \\ U(k) + \partial^+ U(k)(x-k), & k \le x \le k+1. \end{cases}$$

Then we calculate

$$\|f\|_1 = \left|U(1)\right| + \sum_{k=1}^\infty \int_0^1 \left(\left|U(k)\right|(1-x) + \left|U(k+1)\right|x\right)\mathrm{d}x \le 2\|U\|_1.$$

Similarly, we get

$$\mathcal{E}_\lambda(f) = \sum_{k=1}^{\infty} |\partial^+ U(k)|^2 \int_k^{k+1} x^\lambda \, \mathrm{d}x \lesssim E_\lambda(U),$$

and

$$\begin{aligned} \|f\|_2^2 &= U(1)^2 + \sum_{k=1}^{\infty} \int_0^1 \big(U(k) + \partial^+ U(k) x\big)^2 \, \mathrm{d}x \\ &= U(1)^2 + \frac{1}{3} \sum_{k=1}^{\infty} \big(U(k)^2 + U(k+1)^2 + U(k)U(k+1)\big) \geq \frac{1}{3}\|U\|_2^2. \end{aligned}$$

Thus applying (CNI) to the function f yields the desired inequality for U. □

Hence, it remains to proof Proposition 3.20. We generalize the argument given in [10, Section 4.4] due to [12]. For doing so, we have to adapt two ingredients of the proof to cover the weighted Dirichlet form (3.31). First, we need the following weighted version of the Pólya–Szegő rearrangement inequality.

Lemma 3.21 (Weighted Pólya–Szegő)**.** *Let $\lambda \in [0, 2)$. Then for all non-negative $f \in H^1_{\mathrm{loc}}(\overline{\mathbb{R}}_+)$ with $\mathcal{E}_\lambda(f) < \infty$ holds*

$$\mathcal{E}_\lambda(f^*) \leq \mathcal{E}_\lambda(f), \tag{3.32}$$

where f^* is the non-increasing rearrangement of f.

Proof. For the proof, let $k = \frac{\lambda}{2} \in [0, 1)$. Let $\Omega \subset \overline{\mathbb{R}}_+$ arbitrary and $\Omega^* = [0, |\Omega|)$ be the interval from 0 to $|\Omega|$, then it holds

$$\int_{\partial\Omega^*} |x|^k \, \mathcal{H}^0(\mathrm{d}x) \leq \int_{\partial\Omega} |x|^k \, \mathcal{H}^0(\mathrm{d}x). \tag{3.33}$$

Indeed, for the proof one can argue that it is enough to consider intervals $\Omega = (x, x + r)$ for $x \in \overline{\mathbb{R}}_+$ and $r > 0$ (see [2, Theorem 6.1] on how to reduce to this statement). Then, the desired inequality becomes the obvious statement

$$r^k \leq |x|^k + |x + r|^k.$$

For the rest of the proof, we can follow exactly along the same lines as in [2, Theorem 8.1] with the only difference, that now the isoperimetric inequality (3.33) is used. □

The next ingredient is a weighted Poincaré inequality.

Lemma 3.22 (Weighted Poincaré inequality)**.** *For any $\lambda \in [0, 2)$ exists $C_{\mathrm{PI}}(\lambda) \in (0, \infty)$ such that for any $R > 0$ and any $f \in H^1_{\mathrm{loc}}(\overline{\mathbb{R}}_+)$ with $\int_0^R f \, \mathrm{d}x = 0$ holds*

$$\int_0^R |f|^2 \, \mathrm{d}x \leq R^{2-\lambda} C_{\mathrm{PI}}(\lambda) \int_0^R |f'|^2 |x|^\lambda \, \mathrm{d}x. \tag{3.34}$$

Moreover, the constant $C_{\mathrm{PI}}(\lambda)$ is bounded by

$$C_{\mathrm{PI}}(\lambda) \leq \frac{1}{2(2-\lambda)(4-\lambda)}.$$

Proof. Rescaling reduces (3.34) to the inequality

$$\int_0^1 |f|^2 \,\mathrm{d}x \leq C_{\mathrm{PI}}(\lambda) \int_0^1 |f'|^2 |x|^\lambda \,\mathrm{d}x.$$

The above inequality follows from an argument by [14]. Let $g : [0,1] \to \mathbb{R}$ be a monotone increasing absolutely continuous function. Then, it holds

$$\begin{aligned}
2\int_0^1 |f|^2 \,\mathrm{d}x &= \int_0^1 \int_0^1 (f(x) - f(y))^2 \,\mathrm{d}x \,\mathrm{d}y \\
&= \int_0^1 \int_0^1 \left(\int_x^y \frac{f'(\xi)}{\sqrt{g'(\xi)}} \sqrt{g'(\xi)} \,\mathrm{d}\xi \right)^2 \mathrm{d}x \,\mathrm{d}y \\
&\leq \int_0^1 \int_0^1 \int_x^y \frac{|f'(\xi)|^2}{g'(\xi)} \,\mathrm{d}\xi \int_x^y g'(\xi) \,\mathrm{d}\xi \,\mathrm{d}x \,\mathrm{d}y \\
&= \int_0^1 \frac{|f'(\xi)|^2 |\xi|^\lambda}{g'(\xi)|\xi|^\lambda} \int_0^\xi \int_\xi^1 (g(y) - g(x)) \,\mathrm{d}y \,\mathrm{d}x \,\mathrm{d}\xi \\
&\leq \sup_{\xi \in [0,1]} \left(\frac{1}{g'(\xi)|\xi|^\lambda} \int_0^\xi \int_\xi^1 (g(y) - g(x)) \,\mathrm{d}y \,\mathrm{d}x \right) \int_0^1 |f'(\xi)|^2 |\xi|^\lambda \,\mathrm{d}\xi
\end{aligned}$$

Hence for any choice of g, where the sup in ξ is finite, the first term provides an upper bound on $C_{\mathrm{PI}}(\lambda)$. We choose $g(\xi) = \frac{\xi^r - 1}{r}$ for some $r > 0$ yet to be determined and obtain the upper bound

$$C_{\mathrm{PI}}(\lambda) \leq \frac{1}{2} \sup_{\xi \in [0,1]} \left(\xi^{1-r-\lambda} \frac{\xi(1 - \xi^r)}{r(1+r)} \right). \tag{3.35}$$

We choose $r = 1 - \frac{\lambda}{2}$ and note that with this choice

$$\xi^{2-r-\lambda} \left(1 - \xi^r\right) = \xi^{1-\frac{\lambda}{2}} \left(1 - \xi^{1-\frac{\lambda}{2}}\right) \leq \frac{1}{4}.$$

Hence, we obtain the bound

$$C_{\mathrm{PI}}(\lambda) \leq \frac{1}{8} \frac{1}{\left(1 - \frac{\lambda}{2}\right)\left(2 - \frac{\lambda}{2}\right)} = \frac{1}{2(2-\lambda)(4-\lambda)}.$$

□

Now, the proof of Proposition 3.20 follows along the same lines as in [10, Section 4.4].

Proof of Proposition 3.20. We can assume without loss of generality that f is non-negative and denote with f^* its non-increasing rearrangement. Then, we have $\|f^*\|_2 = \|f\|_2$ by Cavalieri's principle and thanks to Lemma 3.21 also $\mathcal{E}_\lambda(f^*) \leq \mathcal{E}_\lambda(f)$. So, we can consider non-increasing non-negative functions. For any $R > 0$ let $f_R = f\mathbb{1}_{[0,R)}$. Since f is non-increasing, it holds

$$f - f_R \leq f(R) \leq \bar{f}_R = \frac{\|f_R\|_1}{R}.$$

Taking the L^2-norm of the above inequality gives

$$\|f - f_R\|_2^2 \leq \bar{f}_R \|f - f_R\|_1 = \frac{\|f_R\|_1}{R}\|f - f_R\|_1. \tag{3.36}$$

Likewise, we can write

$$\|f_R\|_2^2 = \left\|f_R - \bar{f}_R\right\|_2^2 + \left\|\bar{f}_R\,\mathbb{1}_{[0,R)}\right\|_2^2.$$

Applying the weighted Poincaré inequality from Lemma 3.22 to the first term results in the estimate

$$\|f_R\|_2^2 \leq R^{2-\lambda} C_{\mathrm{PI}}(\lambda)\mathcal{E}_\lambda(f) + \frac{\|f_R\|_1^2}{R}, \tag{3.37}$$

where we used that $\mathcal{E}_\lambda(f_R) \leq \mathcal{E}_\lambda(f)$. Now, we write $\|f\|_2^2 \leq \|f_R\|_2^2 + \|f - f_R\|_2^2$ and apply the two estimates (3.36) and (3.37) to arrive at

$$\|f\|_2^2 \leq R^{2-\lambda} C_{\mathrm{PI}}(\lambda)\mathcal{E}_\lambda(f) + \frac{\|f_R\|_1}{R}(\|f_R\|_1 + \|f - f_R\|_1) \leq R^{2-\lambda} C_{\mathrm{PI}}(\lambda)\mathcal{E}_\lambda(f) + \frac{\|f\|_1^2}{R}.$$

The choice

$$R^* = \left(\frac{\|f\|_1^2}{C_{\mathrm{PI}}(\lambda)\mathcal{E}_\lambda(f)}\right)^{\frac{1}{3-\lambda}},$$

yields the claimed estimate (CNI). □

3.2.3. Decay and continuity: Proof of Theorem 3.7

Recall that every solution U to equation (NP) can be represented by

$$U(t,k) = \sum_{l=1}^{\infty} \Phi(t,k,l) U_0(l), \tag{3.38}$$

where Φ is the fundamental solution, see Proposition 3.13. By the classical arguments from [48] and the discrete Nash inequality (DNI), we obtain the decay of the Green function.

Lemma 3.23. *Let $\Phi : \overline{\mathbb{R}}_+ \times \mathbb{N} \times \mathbb{N}$ be the fundamental solution of (NP) from Proposition 3.13, then it holds*

$$\|\Phi(t,\cdot,l)\|_2 \lesssim (1+t)^{-\frac{\alpha}{2}} \qquad \textit{and} \qquad \|\Phi(t,\cdot,l)\|_\infty \lesssim (1+t)^{-\alpha}. \tag{3.39}$$

Proof. For convenience we use the notation $\Phi(t) = \Phi(t,\cdot,l)$ for some fixed l during the proof. We define $f(t) = \|\Phi(t)\|_2^2$ and calculate

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = -2E_\lambda(\Phi(t)).$$

Then by the Nash-type inequality (DNI) we estimate

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) \lesssim -f(t)^{3-\lambda},$$

where we used that all fundamental solutions have unit mass. Integrating this differential inequality and using $f(0) = 1$ we get the desired inequality for $\|\Phi\|_2$. To strengthen this estimate to a uniform bound (with faster decay), we apply the representation formula (3.38) and obtain

$$\Phi(2t,k,l) = \sum_{m=1}^{\infty} \Phi(t,k,m)\Phi(t,m,l) \leq \|\Phi(t)\|_2^2,$$

where we used that $\Phi(t,k,l) = \Phi(t,l,k)$ by symmetry of the operator L_λ. $\square$

The representation formula (3.38) for general solutions then directly implies the L^∞-decay estimate in Theorem 3.7. Next we analyze the temporal decay of the Dirichlet form E_λ of solutions. The identity $\frac{\mathrm{d}}{\mathrm{d}t}\|U\|_2^2 = -2E_\lambda(U)$ and the above L^2 estimate suggest an estimate of the form $E_\lambda(U) \lesssim t^{-(\alpha+1)}$. The next lemma shows that this is indeed the case and in particular we have a Nash-continuity estimate.

Lemma 3.24. *Let $\Phi : \overline{\mathbb{R}}_+ \times \mathbb{N} \times \mathbb{N}$ be the fundamental solution of* (NP) *from Proposition 3.13, then for all $0 < s < t$ and $k_1, k_2, l \in \mathbb{N}$ it holds*

$$E_\lambda(\Phi(t,\cdot,l)) \lesssim t^{-(\alpha+1)}, \tag{3.40}$$

$$|\Phi(t,k_2,l) - \Phi(t,k_1,l)| \lesssim t^{-\alpha}\left|\theta_\lambda(t^{-\alpha}k_2) - \theta_\lambda(t^{-\alpha}k_1)\right|^{\frac{1}{2}}, \tag{3.41}$$

$$|\Phi(t,k,l) - \Phi(s,k,l)| \lesssim s^{-\alpha}\omega_\lambda({}^t\!/\!_s), \tag{3.42}$$

where

$$\theta_\lambda(x) = \begin{cases} \frac{1}{1-\lambda}x^{1-\lambda}, & \lambda \neq 1, \\ \log(x), & \lambda = 1, \end{cases} \tag{3.43}$$

$$\omega_\lambda(r) = \begin{cases} \frac{2}{|1-\alpha|}\left|(r - {}^1\!/\!_2)^{\frac{1-\alpha}{2}} - ({}^1\!/\!_2)^{\frac{1-\alpha}{2}}\right|, & \lambda \neq 1, \\ \log(2r-1), & \lambda = 1. \end{cases} \tag{3.44}$$

Proof. The Cauchy-Schwarz inequality gives $|\langle U, -L_\lambda U\rangle_2| \leq \|U\|_2\|L_\lambda U\|_2$, hence with $f(t) = E_\lambda(\Phi(t,\cdot,l)) = \langle \Phi(t,\cdot,l), -L_\lambda\Phi(t,\cdot,l)\rangle_2$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = -2\langle L_\lambda\Phi(t,\cdot,l), L_\lambda\Phi(t,\cdot,l)\rangle \leq -2\|\Phi(t,\cdot,l)\|_2^{-2} f^2 \lesssim -t^\alpha f(t)^2,$$

where we used the L^2 decay estimate on Φ. Integrating this differential inequality yields

$$f(t) \lesssim \frac{1}{f(0)^{-1} + t^{\alpha+1}} \le t^{-(\alpha+1)}.$$

For the second statement the (discrete) fundamental theorem of calculus and Cauchy-Schwarz inequality imply

$$\begin{aligned}|\Phi(t,k_2,l) - \Phi(t,k_1,l)| &\le \sum_{m=k_1}^{k_2} |\partial^+\Phi(t,m,l)| \le E_\lambda(\Phi(t,\cdot,l))^{\frac{1}{2}} \left(\sum_{m=k_1}^{k_2} m^{-\lambda}\right)^{\frac{1}{2}} \\ &\lesssim t^{-\frac{\alpha+1}{2}} |\theta_\lambda(k_2) - \theta_\lambda(k_1)|^{\frac{1}{2}} = t^{-\alpha} \Big|\theta_\lambda(t^{-\alpha}k_2) - \theta_\lambda(t^{-\alpha}k_1)\Big|^{\frac{1}{2}}.\end{aligned}$$

For the last statement we use for $0 \le t_0 < t$ the representation

$$\Phi(t,k,l) = \sum_{m=1}^{\infty} \Phi(t-t_0,k,m)\Phi(t_0,m,l).$$

In particular for $0 < t_0 < s < t$ we have

$$\begin{aligned}|\partial_t\Phi(t,k,l)| &= \left|\sum_{m=1}^{\infty} \partial_t\Phi(t-t_0,k,m)\Phi(t_0,m,l)\right| \\ &= \left|\sum_{m=1}^{\infty} L_{\lambda,m}\Phi(t-t_0,m,k)\Phi(t_0,m,l)\right| \\ &\le E_\lambda(\Phi(t-t_0,\cdot,k))^{\frac{1}{2}} E_\lambda(\Phi(t_0,\cdot,l))^{\frac{1}{2}} \lesssim (t-t_0)^{-\frac{\alpha+1}{2}} t_0^{-\frac{\alpha+1}{2}},\end{aligned}$$

hence

$$|\Phi(t,k,l) - \Phi(s,k,l)| \le \int_s^t |\partial_r\Phi(r,k,l)| \,\mathrm{d}r \le t_0^{-\frac{\alpha+1}{2}} \int_s^t (r-t_0)^{-\frac{\alpha+1}{2}} \,\mathrm{d}r.$$

Choosing $t_0 = s/2$ and evaluating the integral on the right-hand-side we arrive at

$$|\Phi(t,k,l) - \Phi(s,k,l)| \lesssim \frac{2}{|1-\alpha|} s^{-\frac{\alpha+1}{2}} \left|\left(t - \frac{s}{2}\right)^{\frac{1-\alpha}{2}} - \left(\frac{s}{2}\right)^{\frac{1-\alpha}{2}}\right| = s^{-\alpha}\omega_\lambda\Big(\frac{t}{s}\Big).$$

□

Again the continuity estimates for general solutions to equation (NP) in the second part in Theorem 3.7 follow from the representation (3.38).

3.2.4. Moment estimates

We want to estimate the moments $M_\mu[\Phi(t,\cdot,l)]$ of the fundamental solution of equation (NP) from above and below optimally in terms of scaling.

Lemma 3.25. *Let $\Phi : \overline{\mathbb{R}}_+ \times \mathbb{N} \times \mathbb{N}$ be the fundamental solution of (NP) from Proposition 3.13, then for some $C = C(\lambda, \mu) > 0$ the following moment bounds hold:*

$$\begin{aligned}
&\text{for } \mu > 0: && M_\mu[\Phi(t,\cdot,l)] \le \left(l^{\frac{1}{\alpha}} + Ct\right)^{\alpha\mu};\\
&\text{for } \lambda \ge 1 \text{ and } \mu > 0: && M_\mu[\Phi(t,\cdot,l)] \ge \left(l^{\frac{1}{\alpha}} + Ct\right)^{\alpha\mu};\\
&\text{for } \mu < 0: && M_\mu[\Phi(t,\cdot,l)] \ge \left(l^{\frac{1}{\alpha}} + Ct\right)^{\alpha\mu}.
\end{aligned}$$

Proof. For the proof, C always denotes a constant that may depend on λ and exponents μ and ν. The first estimate is obtained for $\mu \ge 2-\lambda$. Taking the time derivative, applying Jensen's inequality with the power $0 \le \frac{\mu+\lambda-2}{\mu} \le 1$, and using Lemma 3.14 to estimate the term $L_\lambda(k^\mu)$, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} M_\mu[\Phi(t,\cdot,l)] &= \sum_{k=1}^{\infty} L_\lambda(k^\mu)\Phi(t,k,l) \le C \sum_{k=1}^{\infty} k^{\mu+\lambda-2}\Phi(t,k,l)\\
&\le C\left(\sum_{k=1}^{\infty} k^\mu \Phi(t,k,l)\right)^{\frac{\mu+\lambda-2}{\mu}} = C M_\mu[\Phi(t,\cdot,l)]^{\frac{\mu+\lambda-2}{\mu}}.
\end{aligned}$$

By using $M_\lambda[\Phi(0,\cdot,l)] = l^\mu$, the above differential inequality is integrated to

$$M_\mu[\Phi(t,\cdot,l)] \le \left(l^{\frac{1}{\alpha}} + Ct\right)^{\alpha\mu}.$$

Then, for any $0 < \nu < \mu$ we apply again Jensen's inequality to arrive at

$$M_\nu[\Phi] \le M_\mu[\Phi]^{\frac{\nu}{\mu}} \le \left(l^{\frac{1}{\alpha}} + Ct\right)^{\alpha\nu},$$

which shows that the above upper estimate holds in fact for any $\mu > 0$. Next we derive a lower bound in the case $\lambda \ge 1$, $0 < \mu < 2-\lambda$. Indeed, a similar calculation as above yields

$$\frac{\mathrm{d}}{\mathrm{d}t} M_\mu[\Phi(t,\cdot,l)] = \sum_{k=1}^{\infty} L_\lambda(k^\mu)\Phi(t,k,l) \ge C \sum_{k=1}^{\infty} k^{\mu+\lambda-2}\Phi(t,k,l) \ge C M_\mu[\Phi(t,\cdot,l)]^{\frac{\mu+\lambda-2}{\mu}},$$

by the second statement of Lemma 3.14 and the fact that $\frac{\mu+\lambda-2}{\mu} < 0$. This is then integrated as above and yields the inequality

$$M_\mu[\Phi(t,\cdot,l)] \ge \left(l^{\frac{1}{\alpha}} + Ct\right)^{\alpha\mu},$$

and by applying Jensen's inequality this inequality holds for all $\mu > 0$. Also note that the above estimate holds for all $\mu < 0$. Indeed, for $\mu < 0$ we apply Jensen's inequality again to obtain

$$M_\mu[\Phi] \ge M_1[\Phi]^\mu \ge \left(l^{\frac{1}{\alpha}} + Ct\right)^{\alpha\mu}. \qquad \square$$

Next we prove a general interpolation inequality for moments.

Lemma 3.26. *For $u \in \ell^\infty_+(\mathbb{N})$ with $M_1[u] < \infty$ and every $\mu \in (0,1)$ holds*

$$M_\mu[u] \leq 2M_0[u]^{1-\mu} M_1[u]^\mu.$$

Proof. For any $N \in \mathbb{N}$, we have the estimate

$$M_\mu[u] = \sum_{k=1}^{\infty} k^\mu u(t,k) = \sum_{k=1}^{N} k^\mu u(t,k) + \sum_{k=N+1}^{\infty} k^\mu u(t,k) \leq N^\mu M_0[u] + (N+1)^{\mu-1} M_1[u].$$

Now, we choose N to be the largest natural number such that $N \leq M_1[u]M_0[u]^{-1}$, which implies also $N \geq M_1[u]M_0[u]^{-1} - 1$ and thus the estimate

$$\begin{aligned} N^\mu M_0[u] + (N+1)^{\mu-1} M_1[u] &\leq (M_1[u]M_0[u]^{-1})^\mu M_0[u] + (M_1[u]M_0[u]^{-1})^{\mu-1} M_1[u] \\ &= 2M_0[u]^{1-\mu} M_1[u]^\mu. \end{aligned}$$

□

The estimates from Lemma 3.25 and Lemma 3.26 yield various moment bounds for solutions to equation (DP).

Proposition 3.27. *Any solution u to the equation (DP) in $X^+_{\max(1,\lambda)}$ with $M_1[u] = \rho$ satisfies the moment bounds:*

1. *For $0 < \mu < 1$, there exist constants $C_1 = C_1(\lambda, \mu) > 0$ and $C_2 = C_2(\lambda) > 0$ such that*

$$C_1 \rho^{-\frac{\mu}{1-\mu}} M_\mu[u_0]^{\frac{1}{1-\mu}} (1 \vee t)^{-\alpha} \leq M_0[u(t,\cdot)] \leq C_2\, \rho\, t^{-\alpha}.$$

2. *For $0 < \mu < 1$, there exist constants $C_1 = C_1(\lambda, \mu) > 0$ and $C_2 = C_2(\lambda) > 0$ such that*

$$C_1 M_\mu[u_0] (1 \vee t)^{\alpha(\mu-1)} \leq M_\mu[u(t,\cdot)] \leq C_2\, \rho\, t^{\alpha(\mu-1)}.$$

3. *For $\mu > 1$, there exist constants $C_1 = C_1(\lambda, \mu) \geq 0$ and $C_2 = C_2(\lambda, \mu) > 0$ such that*

$$C_1\, \rho\, t^{\alpha(\mu-1)} \leq M_\mu[u(t,\cdot)] \leq C_2 M_\mu[u_0] (1 \vee t)^{\alpha(\mu-1)}.$$

   Furthermore, C_1 is strictly positive for $\lambda \in [1,2)$.

Proof. We start with the second statement. Note that, up to constants that depend only on μ, $M_\mu[u]$ is comparable to $M_{\mu-1}[U]$, where U is the tail distribution corresponding to u and thus a solution to equation (NP). Then the third inequality from Lemma 3.25 and the representation formula (3.38) yield

$$M_{\mu-1}[U(t,\cdot)] = \sum_{l=1}^{\infty} M_{\mu-1}[\Phi(t,\cdot,l)] U_0(l) \geq \sum_{l=1}^{\infty} \left(l^{1/\alpha} + Ct\right)^{\alpha(\mu-1)} U_0(l). \tag{3.45}$$

Next, for $t \leq 1$ we have

$$\left(l^{1/\alpha} + Ct\right)^{\alpha(\mu-1)} \geq \left(l^{1/\alpha} + C\right)^{\alpha(\mu-1)} \geq l^{\mu-1}(1+C)^{\alpha(\mu-1)},$$

while for $t \geq 1$ we estimate

$$\left(l^{1/\alpha} + Ct\right)^{\alpha(\mu-1)} \geq t^{\alpha(\mu-1)}\left(l^{1/\alpha} + C\right)^{\alpha(\mu-1)} \geq t^{\alpha(\mu-1)} l^{\mu-1}(1+C)^{\alpha(\mu-1)}.$$

Thus, for $t \geq 0$ we have the estimate $\left(l^{1/\alpha} + Ct\right)^{\alpha(\mu-1)} \geq l^{\mu-1}(1+C)^{\alpha(\mu-1)}(1 \vee t)^{\alpha(\mu-1)}$. Plugging this estimate into the representation formula (3.45) gives the lower bound in statement *(2)*. Next, we note that the upper bound from the statement *(1)* immediately follows from the fact that $M_0[u(t,\cdot)] = U(t,0)$ and (3.21). This enables us to prove the upper bound in the statement *(2)* by interpolation. Indeed, by Lemma 3.26 we have

$$M_\mu[u] \leq 2M_0[u]^{1-\mu} M_1[u]^\mu = 2M_0[u]^{1-\mu}\rho^\mu \leq C\,\rho\, t^{\alpha(\mu-1)}.$$

Using the interpolation inequality in the other direction and the lower bound obtained for $M_\mu[u]$ in (3.45), we have for every $\mu \in (0,1)$ that

$$M_0[u]^{1-\mu} \geq \frac{1}{2}\rho^{-\mu} M_\mu[u] \geq \frac{1}{2}\rho^{-\mu} C M_\mu[u_0](1 \vee t)^{\alpha(\mu-1)},$$

which implies the lower bound in statement *(1)*. We turn to the proof of statement *(3)*. In the case $\lambda \geq 1$, the lower bound follows immediately from the second inequality in Lemma 3.25 with $\left(l^{1/\alpha} + Ct\right)^{\alpha(\lambda-1)} \geq Ct^{\alpha(\lambda-1)}$ and the representation formula, whereas the upper bound is proved along the same lines as the lower bound in the case $0 < \mu < 1$, making use of the first inequality from Lemma 3.25. ☐

3.2.5. Coarsening rates: Proof of Theorem 3.1

Recall that equation (DP) and equation (EDG_λ) are linked by the time change τ defined in (3.14), where the function $u(\tau, k)$ defined by $u(\tau(t), k) = c_k(t)$ for $k \geq 1$ is a solution to equation (DP) if $c_k(t)$ is a solution to the system (EDG_λ). Then the moment estimates from above imply the following estimates on τ, from which Theorem 3.1 easily follows.

Proposition 3.28. *The time change τ in (3.14) satisfies for any $0 \leq \lambda < 2$ and $\beta = (3-2\lambda)^{-1}$ the following bounds, with all constants only depending on λ, ρ and $M_\lambda[c^{(0)}]$:*

1. *Let $0 \leq \lambda < 3/2$, then every solution c to equation (EDG_λ) exists globally and there are positive constants C_1, C_2, t_0 such that*

$$C_1 t^{\frac{\beta}{\alpha}} \leq \tau(t) \leq C_2 t^{\frac{\beta}{\alpha}} \qquad \text{for all } t \geq t_0.$$

2. *Let $\lambda = 3/2$, then every solution c to equation (EDG_λ) exists globally and there are positive constants C_1, C_2, K_1, K_2, t_0 such that*

$$K_1 \exp(C_1 t) \leq \tau(t) \leq K_2 \exp(C_2 t) \qquad \text{for all } t \geq t_0.$$

3. *Let $3/2 < \lambda \leq 2$, then every solution c to equation (EDG$_\lambda$) exists only locally on a maximal interval $[0, t^*)$ for some $t^* > 0$ and there are positive constants C_1, C_2, t_0 such that*

$$C_1(t^* - t)^{\frac{\beta}{\alpha}} \leq \tau(t) \leq C_2(t^* - t)^{\frac{\beta}{\alpha}} \qquad \textit{for all } t_0 \leq t \leq t^*.$$

Proof. By construction we have $\dot{\tau} = M_\lambda[u(\tau, \cdot)]$, where u is a solution to equation (DP) on $\operatorname{Im}(\tau)$. Because of Corollary 3.18 we can assume without loss of generality that u is a global solution, even if τ is a bounded function. Then plugging the bounds from Proposition 3.27 with $\mu = \lambda$ into the differential equation for τ one easily sees that $\tau(t)$ remains locally bounded ($\lambda \leq 3/2$) or blows up in finite time ($\lambda > 3/2$), see calculations below. Corollary 3.19 then implies that in the first case solutions can be extended globally, while in the second case $M_\lambda[c]$ blows up in finite time. Next, the lower moment bounds imply in any case there exists some $t_0 > 0$, depending only on λ, ρ and $M_\lambda[c_0]$, such that $\tau(t) \geq 1$ for $t \geq t_0$, so we have differential inequalities

$$C_1 \tau^{\alpha(\lambda-1)} \leq \dot{\tau} \leq C_2 \tau^{\alpha(\lambda-1)}. \tag{3.46}$$

We first consider the case $\lambda < 3/2$ in which $\alpha(\lambda - 1) < 1$. Dividing (3.46) by $\tau^{\alpha(\lambda-1)}$ and integrating from t_0 to t yields

$$\Big(\tau(t_0)^{\frac{\alpha}{\beta}} + \tfrac{\alpha}{\beta} C_1(t - t_0)\Big)^{\frac{\beta}{\alpha}} \leq \tau(t) \leq \Big(\tau(t_0)^{\frac{\alpha}{\beta}} + \tfrac{\alpha}{\beta} C_2(t - t_0)\Big)^{\frac{\beta}{\alpha}}. \tag{3.47}$$

It is easy to see that $\tau(t_0)$ can also be estimated from above and below in terms of λ, ρ and $M_\lambda[c^{(0)}]$, hence after adjusting t_0 the desired inequality for τ holds. In the case $\lambda = 3/2$ we have $\alpha(\lambda - 1) = 1$, and hence integrating the differential inequality yields

$$\tau(t_0) \exp\Big(C_1(t - t_0)\Big) \leq \tau(t) \leq \tau(t_0) \exp\Big(C_2(t - t_0)\Big),$$

which leads to the second statement. For the third statement, we have to consider (3.47), but with β negative in this case, which shows that τ has to blow up. The behavior at the blowup time follows after dividing the differential inequality for τ by $\tau^{\alpha(\lambda-1)}$ and integrating from t to t^* for $t_0 < t < t^*$ to arrive at

$$\Big(-\tfrac{\alpha}{\beta} C_2(t^* - t)\Big)^{\frac{\beta}{\alpha}} \leq \tau(t) \leq \Big(-\tfrac{\alpha}{\beta} C_1(t^* - t)\Big)^{\frac{\beta}{\alpha}}.$$

□

Proof of Theorem 3.1. The first statement of Proposition 3.27 shows that $M_0[u(t, \cdot)]$ is of order $t^{-\alpha}$. Hence, the average cluster size $\ell(t)$ defined in (3.5) translates to the time rescaled moment $\rho / M_0[u(\tau(t), \cdot)]$ and becomes $\rho\, \tau(t)^\alpha$. With this, Theorem 3.1 is a direct consequence of Proposition 3.28. □

3.3. Scaling limit from discrete to continuum

3.3.1. Solutions to the continuum equation

First, we give the explicit construction of the fundamental solution of the problem (NP'). We emphasize that in this subsection the value of λ can be taken in the range $\lambda \in (-\infty, 2)$. We make a change of variables that transforms the operator $\mathcal{L}_\lambda$ in (NP') into the generator of the Bessel process, see [45]. For this we define the new variable

$$z(x) = \int_0^x \frac{1}{\sqrt{a_\lambda(y)}}\,\mathrm{d}y = \frac{2}{2-\lambda} x^{1-\frac{\lambda}{2}}, \qquad \text{whence} \qquad x(z) = \left(\frac{2-\lambda}{2} z\right)^{\frac{1}{1-\frac{\lambda}{2}}}.$$

Then if $\varphi(t,x)$ is a solution to equation (NP'), the function $\tilde\varphi$ defined by $\tilde\varphi(2t, z(x)) = \varphi(t,x)$ solves the equation

$$\partial_t \tilde\varphi = \tfrac{1}{2}\partial_z^2 \tilde\varphi + \tilde a_\lambda \partial_z \tilde\varphi \qquad \text{and} \qquad \partial_z \tilde\varphi|_{z=0} = 0, \tag{TNP}$$

where

$$\tilde a_\lambda(z(x)) = \frac{a'_\lambda(x)}{4\sqrt{a_\lambda(x)}} = \frac{\lambda}{4} x(z)^{\frac{\lambda}{2}-1} = \frac{c_\lambda}{z} \qquad \text{with} \qquad c_\lambda = \frac{\lambda}{2(2-\lambda)} \in \left(-\frac{1}{2}, \infty\right).$$

Hence, the equation (TNP) becomes the generator of the reflected Bessel process [45, p. 10] of dimension $2c_\lambda + 1$. By comparison with [45, Chapter 3], the fundamental solution is explicitly given by

$$\tilde\Psi_{c_\lambda}(t,z,y) = \frac{y^{2c_\lambda}}{t^{c_\lambda + \frac{1}{2}}} \exp\left(-\frac{z^2+y^2}{2t}\right) h_{c_\lambda}\left(\frac{zy}{t}\right).$$

Here, h_ν is an entire function that can be expressed in terms of the modified Bessel function of the first kind $I_\nu(z) = z^\nu h_{\nu+\frac{1}{2}}(z)$. We have $c_\lambda + \frac{1}{2} = \alpha$ and $2c_\lambda = \lambda\alpha$, which allows to rewrite $\tilde\Psi$ as

$$\tilde\Psi(t,z,y) = \frac{y^{\lambda\alpha}}{t^\alpha} \exp\left(-\frac{z^2+y^2}{2t}\right) h_{c_\lambda}\left(\frac{zy}{t}\right). \tag{3.48}$$

Remark 3.29. *As noted in [45, Section 3], the fundamental solution (3.48) to (TNP) with Neumann (reflecting) boundary condition agrees in the range $c_\lambda \geq 1/2$ to the one with Dirichlet (absorbing) boundary conditions, which has the stochastic interpretation that both boundary conditions are in this case non-effective since the process cannot reach 0 in finite time [45, Proposition 1]. The range $c_\lambda \in \left(-1/2, 1/2\right)$ translates to $\lambda < 1$, whereas $c_\lambda \geq 1/2$ is the range $\lambda \in [1,2)$.*

Next we want to transform back to the equation (NP'). Because $\tilde\Psi_{c_\lambda}(t,\cdot,y) \to \delta_y$ for $t \to 0$, we arrive for all smooth f at the identity

$$\begin{aligned}\int_0^\infty \tilde\Psi_{c_\lambda}(t, z(x), y) f(x)\,\mathrm{d}x &= \int_0^\infty \frac{1}{z'(x(z))} \tilde\Psi_{c_\lambda}(t,z,y) f(x(z))\,\mathrm{d}z \\ &\to \frac{1}{z'(x(y))} f(x(y)) = \left(\frac{2-\lambda}{2}\right)^{\lambda\alpha} y^{\lambda\alpha} f(x(y)), \quad \text{as } t \to 0.\end{aligned}$$

Since we want the fundamental solution Ψ_λ for equation (NP') to converge to a Dirac mass as $t \to 0$, we transform the equation back, normalize accordingly and end up with the definition

$$\begin{aligned}\Psi_\lambda(t,x,y) &= (2\alpha)^{\lambda\alpha} z(y)^{-\lambda\alpha} \tilde{\Psi}_{c_\lambda}(2t, z(x), z(y)) \\ &= \left(\frac{2}{2-\lambda}\right)^{\lambda\alpha} \frac{1}{(2t)^\alpha} \exp\left(-\frac{z(x)^2 + z(y)^2}{4t}\right) h_{c_\lambda}\left(\frac{z(x)z(y)}{2t}\right). \end{aligned} \tag{3.49}$$

By consulting [45, (14)], we see that

$$h_{c_\lambda}(0) = \frac{1}{2^{c_\lambda - 1/2}\Gamma(c_\lambda + 1/2)} = \frac{1}{2^{(\lambda-1)\alpha}\Gamma(\alpha)}$$

and can rewrite the normalization constant (3.8) of the scaling profile (3.25) as

$$Z_\lambda = \alpha^{-2\alpha}\Gamma(\alpha+1) = \alpha^{-2\alpha+1}\Gamma(\alpha) = \alpha^{\lambda\alpha}\Gamma(\alpha).$$

Hence, for $y = 0$ we arrive at the scaling solution (3.24). Also, by definition, $\Psi_\lambda(0,\cdot,y) = \delta_y$ and $\Psi(\cdot,\cdot,y)$ is a solution to equation (NP'). In this explicit form it is easy to verify basic properties of the fundamental solution.

Proposition 3.30 (Fundamental solution)**.** *For every $\lambda \in [0,2)$ the function Ψ_λ defined by (3.49) has the following properties:*

1. $\Psi_\lambda \in C^\infty(\mathbb{R}_+^3) \cap C^0(\mathbb{R}_+ \times \overline{\mathbb{R}}_+^2)$.
2. $\Psi_\lambda(t,x,y) = \Psi_\lambda(t,y,x)$.
3. *For every $y \in \overline{\mathbb{R}}_+$ holds $\partial_t \Psi_\lambda(t,\cdot,y) - \mathcal{L}_\lambda \Psi_\lambda(t,\cdot,y) = 0$ and $a_\lambda \partial_x \Psi_\lambda(t,\cdot,y)|_{x=0} = 0$.*
4. *It holds the normalization property*
$$\int_{\mathbb{R}_+} \Psi_\lambda(t,x,y)\,\mathrm{d}x = 1.$$
5. *It holds $\Psi_\lambda(t,\cdot,y) \rightharpoonup \delta_y$ in $\mathcal{M}(\overline{\mathbb{R}}_+)$ as $t \to 0$.*
6. *For all $k \geq 0$ it holds $\mathcal{L}_{\lambda,1}^{(k)}\Psi_\lambda(t,x,y) = \mathcal{L}_{\lambda,2}^{(k)}\Psi_\lambda(t,x,y)$, where $\mathcal{L}_{\lambda,i}^{(k)}$ denotes the k-fold composition of $\mathcal{L}_\lambda$ applied to the i-th spatial variable for $i = 1,2$.*

Proof. Properties *(1)–(5)* are easily verified based on the above calculations. The last property is implied by the properties *(1)–(3)*. Indeed, property 3 states that

$$\begin{aligned}\partial_t \Psi_\lambda(t,x,y) &= \mathcal{L}_\lambda \Psi_\lambda(t,\cdot,y)|_x = a_\lambda(x)\partial_1^2 \Psi_\lambda(t,x,y) + a_\lambda'(x)\partial_1 \Psi_\lambda(t,x,y) \\ &= \mathcal{L}_{\lambda,1}\Psi_\lambda(t,x,y).\end{aligned}$$

Using the symmetry of Ψ_λ we also have

$$\begin{aligned}\partial_t \Psi_\lambda(t,x,y) &= \partial_t \Psi_\lambda(t,y,x) = \mathcal{L}_\lambda \Psi_\lambda(t,\cdot,x)|_y = \mathcal{L}_\lambda \Psi_\lambda(t,x,\cdot)|_y \\ &= a_\lambda(y)\partial_2^2 \Psi_\lambda(t,x,y) + a_\lambda'(y)\partial_2 \Psi_\lambda(t,x,y) = \mathcal{L}_{\lambda,2}\Psi_\lambda(t,x,y),\end{aligned}$$

which implies the statement for $k=1$. The rest of the statement follows easily by induction, since the function $\mathcal{L}_{\lambda,1}\Psi_\lambda(t,x,y)$ is also symmetric and solves the same equation as Ψ_λ. □

Proposition 3.30 motivates to define for $g \in C_c^\infty(\mathbb{R}_+)$ the time-evolution $\mathcal{S}_\lambda(t)g$ by using the fundamental solution as integral kernel, i.e

$$\mathcal{S}_\lambda(t)g = \int_{\mathbb{R}_+} \Psi_\lambda(t,\cdot,y)g(y)\,\mathrm{d}y. \tag{3.50}$$

For this, we deduce the following properties.

Corollary 3.31. *For any $g \in C_c^\infty(\mathbb{R}_+)$, $\mathcal{S}_\lambda(t)g$ from (3.50) is a solution of equation (NP') with initial data g. Furthermore, the following estimates hold:*

1. *For all $p \in [1,\infty]$, $k \geq 0$ and $t \geq 0$ it holds*

$$\|\mathcal{L}_\lambda^{(k)}\mathcal{S}_\lambda(t)g\|_p \leq \|\mathcal{L}_\lambda^{(k)}g\|_p.$$

2. *For all $\nu \geq 0$, $k \geq 0$ and $t \geq 0$ it holds*

$$\|x^\nu \mathcal{L}_\lambda^{(k)}\mathcal{S}_\lambda(t)g\|_\infty \lesssim \begin{cases} (1 \vee t)^{\alpha(\nu-1)}\|\mathcal{L}_\lambda^{(k)}g\|_1 + \|\mathcal{L}_\lambda^{(k)}g\|_\infty + \|x^\nu\mathcal{L}_\lambda^{(k)}g\|_\infty, & \text{if } \nu < 1, \\ t^{\alpha(\nu-1)}\|\mathcal{L}_\lambda^{(k)}g\|_1 + \|x^\nu\mathcal{L}_\lambda^{(k)}g\|_\infty, & \text{if } \nu \geq 1. \end{cases}$$

Proof. Using the properties of Proposition 3.30 and the fact that $g \in C_c^\infty(\mathbb{R}_+)$ it is easy to prove that $\mathcal{S}_\lambda(t)g$ is a solution to equation (NP') with initial data g. Also, since

$$\begin{aligned}\mathcal{L}_\lambda^{(k)}\mathcal{S}_\lambda(t)g &= \int_{\mathbb{R}_+} \mathcal{L}_{\lambda,1}^{(k)}\Psi_\lambda(t,x,y)g(y)\,\mathrm{d}y = \int_{\mathbb{R}_+} \mathcal{L}_{\lambda,2}^{(k)}\Psi_\lambda(t,x,y)g(y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}_+} \Psi_\lambda(t,x,y)\mathcal{L}_\lambda^{(k)}g(y)\,\mathrm{d}y = \mathcal{S}_\lambda(t)\mathcal{L}_\lambda^{(k)}g,\end{aligned}$$

and $\mathcal{L}_\lambda^{(k)}g \in C_c^\infty(\mathbb{R})$ if $g \in C_c^\infty(\mathbb{R}_+)$, it suffices to prove all desired inequalities for $k=0$. For the first inequality we have

$$\begin{aligned}\|\mathcal{S}_\lambda(t)g\|_p^p &= \int_{\mathbb{R}_+}\left|\int_{\mathbb{R}_+}\Psi_\lambda(t,x,y)g(y)\,\mathrm{d}y\right|^p\mathrm{d}x \leq \int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\Psi_\lambda(t,x,y)|g(y)|^p\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_{\mathbb{R}_+}|g(y)|^p\int_{\mathbb{R}_+}\Psi_\lambda(t,x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}_+}|g(y)|^p\,\mathrm{d}y.\end{aligned}$$

Here we used Jensen's inequality with respect to the probability measure $\Psi_\lambda(t,x,\cdot)$. For the second inequality we split the integral

$$|x^\nu \mathcal{S}_\lambda(t)g(x)| \le \int_{\mathbb{R}_+} x^\nu \Psi_\lambda(t,x,y)|g(y)|\,\mathrm{d}y \le \int_0^{rx} x^\nu \Psi_\lambda(t,x,y)|g(y)|\,\mathrm{d}y + r^{-\nu}\|x^\nu g\|_\infty.$$

Using the explicit form (3.49) and the asymptotics of Bessel functions, one can show that for some $r>0$ small enough we have for $y \le rx$ the bound

$$\Psi_\lambda(t,x,y) \le Z_\lambda^{-1} t^{-\alpha} \exp(-c_r t^{-1} x^{2-\lambda}) =: t^{-\alpha} F_\lambda(t^{-\alpha}x).$$

Now a simple calculation shows that the maximum of the function $x \to x^\nu F_\lambda(t^{-\alpha}x)$ is attained at x of order t^α, hence for $y \le rx$ the estimate $x^\nu \Psi_\lambda(t,x,y) \lesssim t^{\alpha(\nu-1)}$ holds. This directly implies the desired estimate in the case $\nu \ge 1$. If $\nu < 1$, we want an estimate that does not blow up at $t=0$. Here we use that $\|x^\nu f\|_\infty \le \|f\|_\infty + \sup_{x\ge 1}|x^\nu f|$. On $[1,\infty)$, the function $x \to x^\nu F_\lambda(t^{-\alpha}x)$ attains its maximum at x of order $1 \vee t^\alpha$, and finally

$$\sup_{t\in\mathbb{R}_+} t^{-\alpha} F_\lambda(t^{-\alpha}) < \infty.$$

$\square$

To analyze the relation between the discrete and the continuous model, we have to work with a weak formulation of equation (NP'), which is based on the adjoint equation (3.51). Thus, we define for $f \in C_c^\infty((0,T)\times\mathbb{R}_+)$ the solution operator $\mathcal{T}_\lambda(t)f$ by

$$\mathcal{T}_\lambda(t)f = \int_0^t \mathcal{S}_\lambda(t-s)f(s,\cdot)\,\mathrm{d}s.$$

Note that $\varphi(t,\cdot) = \mathcal{T}_\lambda(t)f$ is a solution to the inhomogeneous equation

$$\begin{cases} \partial_t\varphi - \mathcal{L}_\lambda\varphi = f, & \text{on } \mathbb{R}_+\times\mathbb{R}_+ \\ a_\lambda \partial_x\varphi|_{x=0} = 0, & \text{on } \mathbb{R}_+, \\ \varphi(0,\cdot) = 0, & \text{on } \mathbb{R}_+. \end{cases} \tag{3.51}$$

Therewith, the definition of weak solutions reads as follows.

Definition 3.32. *For $T>0$, a family of measures $\{\mu_t\}_{t\in[0,T)} \subset \mathcal{M}(\overline{\mathbb{R}}_+)$ is a weak solution to equation (NP') on $[0,T)$ with initial data $\mu_0 \in \mathcal{M}(\overline{\mathbb{R}}_+)$ if for all $f \in C_c^\infty((0,T)\times\mathbb{R}_+)$ it holds*

$$\int_0^T \int_{\overline{\mathbb{R}}_+} f(T-t,x)\,\mathrm{d}\mu_t(x)\,\mathrm{d}t = \int_{\overline{\mathbb{R}}_+} (\mathcal{T}_\lambda(T)f)(x)\,\mathrm{d}\mu_0(x). \tag{3.52}$$

Equivalently, since $\mathcal{T}_\lambda(t)f$ solves the inhomogeneous equation (3.51), μ_t is a weak solution if and only if

$$\int_0^T \int_{\overline{\mathbb{R}}_+} (\partial_t \varphi + \mathcal{L}_\lambda \varphi) \, \mathrm{d}\mu_t(x) \, \mathrm{d}t = - \int_{\overline{\mathbb{R}}_+} \varphi(0,x) \, \mathrm{d}\mu_0(x), \tag{3.53}$$

for all $\varphi(t,\cdot) = \mathcal{T}_\lambda(T-t)f$ with $f \in C_c^\infty((0,T) \times \mathbb{R}_+)$. The definition (3.53) looks more like standard weak formulations of PDE. However, we specify the test function class φ only in terms of the image of the adjoint operator on smooth functions. The reason for this is that (3.52) automatically implies that weak solutions are unique as distributions on $(0,T) \times \mathbb{R}_+$, which is needed to identify the limit of a sequence of approximate solutions (see next subsection). By Corollary 3.31, the class of test functions has good regularity and decay properties. With these, it is easy to verify that the scaling solution γ_λ (3.24) with $\mathcal{G}_\lambda$ given in (3.25) solves (NP') in the weak sense with initial data δ_0.

We close this subsection with the observation that the scaling solution γ_λ is indeed also attractive for all solutions in relative entropy, as in the classical result for the heat equation with $\lambda = 0$.

Remark 3.33. *Let $\mu(t)$ be a solution to equation (NP') starting from some $\mu^{(0)}$ with mass $\rho > 0$. Then the relative entropy of $\mu(t)$ with respect to $\rho\,\gamma_\lambda(t)$ is dissipated, which follows from the simple calculation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}\Big(\mu(t) \mid \rho\,\gamma_\lambda(t)\Big) = -\mathcal{I}_\lambda\Big(\mu(t) \mid \rho\,\gamma_\lambda(t)\Big) = -\int a_\lambda \partial_x \log \frac{\mu(t)}{\rho\,\gamma_\lambda(t)} \partial_x \frac{\mu(t)}{\rho\,\gamma_\lambda(t)} \, \mathrm{d}\gamma_\lambda(t), \tag{3.54}$$

once sufficient regularity is established for any $t > 0$. In Appendix A.1 we prove that the following weighted logarithmic Sobolev inequality holds: For all $\lambda \in [0,2)$, there exists $C_{\mathrm{LSI}} = C_{\mathrm{LSI}}(\lambda)$ such that for any $t > 0$ and any measure $\mu \in \mathcal{M}(\mathbb{R}_+)$ with mass $\rho > 0$ and $\mathcal{H}(\mu \mid \rho\,\gamma_\lambda(t)) < \infty$ the inequality

$$\mathcal{H}\Big(\mu \mid \rho\,\gamma_\lambda(t)\Big) \le 4 C_{\mathrm{LSI}}\, t\, \mathcal{I}_\lambda\Big(\mu \mid \rho\,\gamma_\lambda(t)\Big) \tag{3.55}$$

holds. Hence, once sufficient regularity for solutions to (NP') is established, it immediately follows that those converge to the self-similar profile $\rho\,\gamma_\lambda(t)$ in relative entropy and hence also in $L^1(\overline{\mathbb{R}}_+)$ by the Pinsker inequality.

The argument suggests that solutions to the discrete equation (NP) also get close to the continuum equation (NP') in the limit $t \to \infty$. The weighted logarithmic Sobolev inequality suggests that the entropy method after [60] might be applicable as well. However, here we opted for a more classical approach based on the Nash inequality (Proposition 3.6) and the resulting Nash continuity estimates (Theorem 3.7).

3.3.2. Strategy and proof of Theorem 3.8

Let U_ε be a sequence of solutions to equation (NP) with $\sup_{0<\varepsilon\le 1} \|U_{0,\varepsilon}\|_1 < \infty$ and $\mathcal{U}_\varepsilon$ be the associated sequence of approximate solutions as in (3.28). To see that $\mathcal{U}_\varepsilon$ converges to a solution to equation (NP'), let $T > 0$ and $\varphi(t,\cdot) = \mathcal{T}_\lambda(T-t)f$, $f \in C_c^\infty((0,T)\times\mathbb{R}_+)$. Multiplying $\mathcal{U}_\varepsilon$ with $\partial_t\varphi$ and integrating over space-time, we get

$$\begin{aligned}\int_0^T\int_0^\infty \mathcal{U}_\varepsilon \partial_t\varphi \,\mathrm{d}x\,\mathrm{d}t &= \varepsilon^{-\alpha}\int_0^T \sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t,k)\pi_\varepsilon\partial_t\varphi(t,k)\,\mathrm{d}t \\ &= \varepsilon^{-\alpha}\int_0^T \sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t,k)\partial_t\pi_\varepsilon\varphi(t,k)\,\mathrm{d}t.\end{aligned}$$

Integrating by parts in time, using the equation for U and the symmetry of L_λ, we arrive at

$$\begin{aligned}\int_0^T\int_0^\infty \mathcal{U}_\varepsilon \partial_t\varphi \,\mathrm{d}x\,\mathrm{d}t = &-\int_0^\infty \mathcal{U}_\varepsilon(0,x)\varphi(0,x)\,\mathrm{d}x \\ &- \varepsilon^{-\alpha}\int_0^T \sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t,k)\varepsilon^{-1}L_\lambda\pi_\varepsilon\varphi(t,k)\,\mathrm{d}t,\end{aligned} \tag{3.56}$$

with π_ε as in (3.27). To relate the above identity to the weak formulation of equation (3.53), we need to express the last line in terms of $\mathcal{L}_\lambda\varphi$. Adding and subtracting the term $\pi_\varepsilon\mathcal{L}_\lambda\varphi$ in the last summation and using the identity

$$\varepsilon^{-\alpha}\sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t,k)\pi_\varepsilon\mathcal{L}_\lambda\varphi(t,k) = \int_0^\infty \mathcal{U}_\varepsilon\,\mathcal{L}_\lambda\varphi\,\mathrm{d}x,$$

we get

$$\varepsilon^{-\alpha}\sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t,k)\varepsilon^{-1}L_\lambda\pi_\varepsilon\varphi(t,k) = \int_0^\infty \mathcal{U}_\varepsilon\,\mathcal{L}_\lambda\varphi\,\mathrm{d}x + \varepsilon^{-\alpha}\sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t,k)\mathcal{R}_\varepsilon(\varphi,k),$$

where

$$\mathcal{R}_\varepsilon(\varphi,k) = \varepsilon^{-1}L_\lambda\pi_\varepsilon\varphi(k) - \pi_\varepsilon\mathcal{L}_\lambda\varphi(k) \tag{3.57}$$

denotes the defect between the discrete and continuous operator. The crucial ingredient for the proof is the following estimate on the defect which shows that the rescaled discrete operator can be replaced with the continuous operator on functions that are regular enough.

Lemma 3.34 (Replacement lemma)**.** *Let $\varphi \in C^0(\overline{\mathbb{R}}_+) \cap C^3(\mathbb{R}_+)$ with the following properties:*

1. *The map $x \mapsto a_\lambda(x)\partial_x\varphi(x)$ is Lipschitz-continuous on $\overline{\mathbb{R}}_+$.*
2. *The boundary condition $a_\lambda\partial_x\varphi|_{x=0} = 0$ is satisfied.*

Then the following estimates hold:

$$\|\varepsilon^{-1}L_\lambda\pi_\varepsilon\varphi\|_\infty \lesssim \|\mathcal{L}_\lambda\varphi\|_\infty\varepsilon^\alpha, \tag{3.58}$$

$$|\mathcal{R}_\varepsilon(\varphi,k)| \lesssim \varepsilon^\alpha\int_{(k-2)\varepsilon^\alpha}^{(k+1)\varepsilon^\alpha}\Big(x^{\lambda-1}|\partial_x^2\varphi| + x^\lambda|\partial_x^3\varphi|\Big)\,\mathrm{d}x, \qquad \textit{for } k\geq 3. \tag{3.59}$$

The above result together with the previous calculations yields that rescaled solutions of the discrete problem (NP) are approximate solutions of the continuous equation (NP').

Proposition 3.35 (Approximate weak solutions). *Let U_ε and $\mathcal{U}_\varepsilon$ be as above. Then for $\varphi(t,\cdot) = \mathcal{T}_\lambda(T-t)f$ with $f \in C_c^\infty((0,T)\times\mathbb{R}_+)$ it holds*

$$\int_0^T \int_0^\infty \mathcal{U}_\varepsilon \left(\partial_t\varphi + \mathcal{L}_\lambda\varphi\right) \mathrm{d}x\, \mathrm{d}t = -\int_0^\infty \mathcal{U}_\varepsilon(0,x)\varphi(0,x)\, \mathrm{d}x + \mathrm{o}(1) \qquad \text{as } \varepsilon \to 0,$$

where the terms in $\mathrm{o}(1)$ *depend on* T, f *and the bound on* $U_{0,\varepsilon}$*.*

The full rigorous proof of Proposition 3.35 is given at the end of this section. To make use of Lemma 3.34 we need regularity estimates for φ, as the error term $\mathcal{R}_\varepsilon$ contains derivatives up to third order. However, because of the degeneracy of a_λ the higher derivatives blow up at 0. This can be resolved by introducing a small-scale boundary region at 0 where the error term vanishes in the limit thanks to the uniform bound (3.21) on $\mathcal{U}_\varepsilon$ from Theorem 3.7.

To pass to the limit in the approximate weak formulation we have to establish compactness in a suitable topology. The scale-invariant estimates from Section 3.2.3 in fact imply boundedness and equicontinuity on compact sets, which yields compactness with respect to (local) uniform convergence.

Proposition 3.36 (Compactness). *Let U_ε and $\mathcal{U}_\varepsilon$ be as above. Then for all $x, y \in \overline{\mathbb{R}}_+$ and $0 < s < t$ it holds*

$$\begin{aligned}
\|\mathcal{U}_\varepsilon(t,\cdot)\|_\infty &\lesssim \|U_{0,\varepsilon}\|_1 t^{-\alpha},\\
|\mathcal{U}_\varepsilon(t,x) - \mathcal{U}_\varepsilon(t,y)| &\lesssim t^{-\alpha}\|U_{0,\varepsilon}\|_1 \left(|\theta_\lambda(t^{-\alpha}x) - \theta_\lambda(t^{-\alpha}y)|^{\frac{1}{2}} + \Xi_{\lambda,\varepsilon}(t,x,y)\right),\\
|\mathcal{U}_\varepsilon(t,x) - \mathcal{U}_\varepsilon(s,x)| &\lesssim s^{-\alpha}\|U_{0,\varepsilon}\|_1 \omega_\lambda(t,s),
\end{aligned}$$

where $\theta_\lambda, \omega_\lambda$ are as in Lemma 3.24 and $\Xi_{\lambda,\varepsilon} \to 0$ as $\varepsilon \to 0$ locally uniformly on $\mathbb{R}_+ \times \overline{\mathbb{R}}_+^2$ in the case $\lambda < 1$ and locally uniformly on $\mathbb{R}_+^3$ in the case $\lambda \geq 1$.

Proof of Proposition 3.36. This result is an easy consequence of Theorem 3.7. The L^∞ bound and continuity estimate for the time variable are immediate from (3.21), respectively (3.23) in Theorem 3.7. For continuity in space we apply (3.22) with $x_\varepsilon = \varepsilon^\alpha(\lfloor\varepsilon^{-\alpha}x\rfloor + 1)$, $y_\varepsilon = \varepsilon^\alpha(\lfloor\varepsilon^{-\alpha}y\rfloor + 1)$ and obtain

$$\begin{aligned}
|\mathcal{U}_\varepsilon(t,x) - \mathcal{U}_\varepsilon(t,y)| &\lesssim \|U_0\|_1 t^{-\alpha} \left|\theta_\lambda(t^{-\alpha}x_\varepsilon) - \theta_\lambda(t^{-\alpha}y_\varepsilon)\right|^{\frac{1}{2}}\\
&\leq \|U_0\|_1 t^{-\alpha} \left|\theta_\lambda(t^{-\alpha}x) - \theta_\lambda(t^{-\alpha}y)\right|^{\frac{1}{2}}\\
&\quad + \|U_0\|_1 t^{-\alpha}\left(|\theta_\lambda(t^{-\alpha}x) - \theta_\lambda(t^{-\alpha}x_\varepsilon)|^{\frac{1}{2}} + |\theta_\lambda(t^{-\alpha}y) - \theta_\lambda(t^{-\alpha}y_\varepsilon)|^{\frac{1}{2}}\right).
\end{aligned}$$

Note that we have $|x - x_\varepsilon| \lesssim \varepsilon^\alpha$, $|y - y_\varepsilon| \lesssim \varepsilon^\alpha$. Thus in the case $0 < \lambda < 1$ the Hölder continuity of θ_λ from (3.43) implies

$$|\theta_\lambda(t^{-\alpha}x) - \theta_\lambda(t^{-\alpha}x_\varepsilon)| \lesssim \theta_\lambda(t^{-\alpha}\varepsilon^\alpha),$$

where the right-hand side does not depend on x, whereas for $1 \leq \lambda < 2$ we have

$$|\theta_\lambda(t^{-\alpha}x) - \theta_\lambda(t^{-\alpha}x_\varepsilon)| \lesssim \left| \int_{t^{-\alpha}x}^{t^{-\alpha}x_\varepsilon} \theta'_\lambda(\xi)\,\mathrm{d}\xi \right|,$$

which goes to zero locally uniformly for $t, x > 0$. The same line of reasoning applies to y and y_ε, which finishes the proof. □

Taking the above statements for granted, the convergence result for $\mathcal{U}_\varepsilon$ easily follows.

Proof of Theorem 3.8. It is easy to check that by Proposition 3.36 the sequence $\mathcal{U}_\varepsilon$ satisfies the assumptions of the Arzela-Ascoli Theorem for discontinuous functions (cf. Proposition A.2) on each compact subset of $\mathbb{R}_+ \times \overline{\mathbb{R}}_+$ in the case $0 \leq \lambda < 1$, respectively $\mathbb{R}_+^2$ in the case $1 \leq \lambda < 2$. Thus by exhaustion with compact sets and a diagonal argument each sequence $\varepsilon \to 0$ has a subsequence (not relabeled) such that $\mathcal{U}_\varepsilon \to \mathcal{U}$ locally uniformly for some function $\mathcal{U} \in C^0(\mathbb{R}_+ \times \overline{\mathbb{R}}_+)$, respectively $C^0(\mathbb{R}_+^2)$. To identify the limit, let $T > 0$ and $f \in C_c^\infty((0,T) \times \mathbb{R}_+)$. Then by Proposition 3.35, applied with $\varphi(t,\cdot) = \mathcal{T}_\lambda(T-t)f$, we have that

$$\int_0^T \int_{\mathbb{R}_+} f(T-t,x)\mathcal{U}_\varepsilon(t,x)\,\mathrm{d}t\,\mathrm{d}x = \int_{\mathbb{R}_+} \mathcal{T}_\lambda(T)f(x)\mathcal{U}_\varepsilon(0,x)\,\mathrm{d}x + \mathrm{o}(1), \qquad \text{as } \varepsilon \to 0.$$

Letting $\varepsilon \to 0$ and using that $\mathcal{U}_\varepsilon(0,\cdot) \rightharpoonup \mu_0$ we arrive at

$$\int_0^T \int_{\mathbb{R}_+} f(T-t,x)\mathcal{U}(t,x)\,\mathrm{d}t\,\mathrm{d}x = \int_{\mathbb{R}_+} \mathcal{T}_\lambda(T)f(x)\,\mathrm{d}\mu_0(x).$$

Thus $\mathcal{U}(t,x)$ is a weak solution of equation (NP') with initial data μ_0 and because of continuity it is unique on $\mathbb{R}_+ \times \overline{\mathbb{R}}_+$, respectively $\mathbb{R}_+^2$, which in turn implies that the convergence holds for every sequence $\varepsilon \to 0$. Thus in the case $0 \leq \lambda < 1$, the limit is completely characterized, whereas for $1 \leq \lambda < 2$ we cannot identify the limit at $x = 0$ but only have boundedness of $\mathcal{U}_\varepsilon(t,0)$ for $t > 0$ by Proposition 3.36. □

It remains to prove Lemma 3.34 and Proposition 3.35, which is done in the next two subsections.

3.3.3. Replacement lemma

We split the proof of Lemma 3.34 into several steps.

Lemma 3.37. *Let φ be as in Lemma 3.34. Then, it holds*

$$|\partial_x\varphi|(x) \leq \|\mathcal{L}_\lambda\varphi\|_\infty x^{1-\lambda}, \tag{3.60}$$
$$|\partial_x^2\varphi(x)| \leq 2\|\mathcal{L}_\lambda\varphi\|_\infty x^{-\lambda}. \tag{3.61}$$

Proof. Because $a_\lambda \partial_x \varphi$ is Lipschitz and equal to zero at the boundary, we have

$$|a_\lambda(x)\partial_x\varphi(x)| \le \|\partial_x(a_\lambda\partial_x\varphi)\|_\infty x,$$

which gives the first statement after dividing by a_λ. This estimate then directly implies that $a'_\lambda\partial_x\varphi$ is bounded by $\|\partial_x(a_\lambda\partial_x\varphi)\|_\infty$, and by Leibniz rule

$$|a_\lambda(x)\partial_x^2\varphi(x)| \le |\partial_x(a_\lambda\partial_x\varphi)(x)| + |a'_\lambda(x)\partial_x\varphi(x)| \le 2\|\partial_x(a_\lambda\partial_x\varphi)\|_\infty.$$

□

Lemma 3.38. *Let φ be as in Lemma 3.34. Then, it holds*

$$\|\varepsilon^{-1}L_\lambda\pi_\varepsilon\varphi\|_\infty \lesssim \|\mathcal{L}_\lambda\varphi\|_\infty\varepsilon^\alpha.$$

Proof. First we consider the case $k \ge 3$. Writing out the term we get

$$\begin{aligned}
\varepsilon^{-1}L_\lambda\pi_\varepsilon\varphi(k) &= \varepsilon^{-1}\Big(a_\lambda(k)\partial^+\pi_\varepsilon\varphi(k) - a_\lambda(k-1)\partial^+\pi_\varepsilon\varphi(k-1)\Big)\\
&= \varepsilon^{-2\alpha}\Big(a_\lambda(k\varepsilon^\alpha)\partial^+\pi_\varepsilon\varphi(k) - a_\lambda((k-1)\varepsilon^\alpha)\partial^+\pi_\varepsilon\varphi(k-1)\Big)\\
&= \varepsilon^{-2\alpha}\Big(a_\lambda(k\varepsilon^\alpha) - a_\lambda((k-1)\varepsilon^\alpha)\Big)\partial^+\pi_\varepsilon\varphi(k-1)\\
&\quad + \varepsilon^{-2\alpha}a_\lambda(k\varepsilon^\alpha)\Big(\partial^+\pi_\varepsilon\varphi(k) - \partial^+\pi_\varepsilon\varphi(k-1)\Big)\\
&= \mathrm{I} + \mathrm{II},
\end{aligned}$$

and one further calculates

$$\begin{aligned}
\mathrm{I} &= \varepsilon^{-2\alpha}\Big(a_\lambda(k\varepsilon^\alpha) - a_\lambda((k-1)\varepsilon^\alpha)\Big)\int_{(k-1)\varepsilon^\alpha}^{k\varepsilon^\alpha}(\varphi(x) - \varphi(x-\varepsilon^\alpha))\,\mathrm{d}x,\\
\mathrm{II} &= \varepsilon^{-2\alpha}a_\lambda(k\varepsilon^\alpha)\int_{(k-1)\varepsilon^\alpha}^{k\varepsilon^\alpha}\Big(\varphi(x+\varepsilon^\alpha) - 2\varphi(x) + \varphi(x-\varepsilon^\alpha)\Big)\,\mathrm{d}x.
\end{aligned}$$

To estimate I, we note that in the case $0 < \lambda < 1$ the mean value theorem and estimate (3.60) imply the bound

$$\begin{aligned}
|a_\lambda(k\varepsilon^\alpha) - a_\lambda((k-1)\varepsilon^\alpha)| &\lesssim \varepsilon^\alpha((k-1)\varepsilon^\alpha)^{\lambda-1},\\
|\varphi(x) - \varphi(x-\varepsilon^\alpha)| &\lesssim \|\mathcal{L}_\lambda\varphi\|_\infty\varepsilon^\alpha(k\varepsilon^\alpha)^{1-\lambda},
\end{aligned}$$

for any $k \ge 2$ and $x \in [(k-1)\varepsilon^\alpha, k\varepsilon^\alpha)$, while for $1 \le \lambda < 2$ we have the estimate

$$\begin{aligned}
|a_\lambda(k\varepsilon^\alpha) - a_\lambda((k-1)\varepsilon^\alpha)| &\lesssim \varepsilon^\alpha(k\varepsilon^\alpha)^{\lambda-1},\\
|\varphi(x) - \varphi(x-\varepsilon^\alpha)| &\lesssim \|\mathcal{L}_\lambda\varphi\|_\infty\varepsilon^\alpha((k-2)\varepsilon^\alpha)^{1-\lambda},
\end{aligned}$$

for $k \ge 3$. In both cases we get the desired estimate for I. For the second term we apply a similar argument. Here, the estimate (3.61) and Taylor expansion imply for $k \ge 3$ that

$$|\varphi(x+\varepsilon^\alpha) - 2\varphi(x) + \varphi(x-\varepsilon^\alpha)| \lesssim \|\mathcal{L}_\lambda\varphi\|_\infty\varepsilon^{2\alpha}((k-2)\varepsilon^\alpha)^{-\lambda},$$

which then gives the correct estimate for II. In the remaining cases $k \in \{1,2\}$ we have

$$|a_\lambda(k\varepsilon^\alpha)| \lesssim \varepsilon^{\alpha\lambda}, \qquad \text{and} \qquad |\varphi(x) - \varphi(x \pm \varepsilon^\alpha)| \lesssim \|\mathcal{L}_\lambda \varphi\|_\infty \varepsilon^{\alpha(2-\lambda)},$$

where in the case $1 \le \lambda < 2$ we use that (3.60) implies Hölder continuity with exponent $2-\lambda$. Hence we have

$$|\varepsilon^{-1} a_\lambda(k) \partial^+ \pi_\varepsilon \varphi(k)| = |\varepsilon^{-2\alpha} a_\lambda(\varepsilon^\alpha k) \partial^+ \pi_\varepsilon \varphi(k)| \lesssim \|\mathcal{L}_\lambda \varphi\|_\infty \varepsilon^\alpha,$$

which finishes the proof. □

Lemma 3.39 (Taylor expansion)**.** *Let φ be as in Lemma 3.34. Then for $\varepsilon > 0$ and every $m \ge 0$ it holds*

$$\pi_\varepsilon \varphi(\cdot \pm \varepsilon^\alpha) = \sum_{l=0}^{m} \frac{(\pm\varepsilon)^{l\alpha}}{l!} \pi_\varepsilon \partial_x^l \varphi + R_m(\varphi, \pm\varepsilon),$$

with

$$|R_m(\varphi, \pm\varepsilon)(k)| \le \frac{\varepsilon^{(m+1)\alpha}}{(m+1)!} \int_{I^\varepsilon_\pm(k)} |\partial_x^{m+1} \varphi(x)| \,\mathrm{d}x,$$

and

$$I^\varepsilon_\sigma(k) = \begin{cases} [(k-1)\varepsilon^\alpha, (k+1)\varepsilon^\alpha), & \textit{if } \sigma = +\varepsilon; \\ [(k-2)\varepsilon^\alpha, k\varepsilon^\alpha), & \textit{if } \sigma = -\varepsilon. \end{cases}$$

Proof. The statement follows directly by standard Taylor expansion, where we use the integral representation for the residual term

$$\begin{aligned} \pi_\varepsilon \varphi(\cdot + \varepsilon^\alpha)(k) &= \int_{(k-1)\varepsilon^\alpha}^{k\varepsilon^\alpha} \varphi(x + \varepsilon^\alpha) \,\mathrm{d}x \\ &= \int_{(k-1)\varepsilon^\alpha}^{k\varepsilon^\alpha} \sum_{l=0}^{m} \frac{\varepsilon^{l\alpha}}{l!} \partial_x^l \varphi(x) \,\mathrm{d}x + \frac{1}{(m+1)!} \int_{(k-1)\varepsilon^\alpha}^{k\varepsilon^\alpha} \int_x^{x+\varepsilon^\alpha} (x + \varepsilon^\alpha - s)^m \partial_x^{m+1} \varphi(s) \,\mathrm{d}s \,\mathrm{d}x \\ &= \sum_{l=0}^{m} \frac{\varepsilon^{l\alpha}}{l!} \pi_\varepsilon \partial_x^l \varphi + R_m(\varphi, \varepsilon). \end{aligned}$$

We then calculate

$$\begin{aligned} |R_m(\varphi, \varepsilon)(k)| &\le \frac{\varepsilon^{m\alpha}}{(m+1)!} \int_{(k-1)\varepsilon^\alpha}^{k\varepsilon^\alpha} \int_x^{x+\varepsilon^\alpha} |\partial_x^{m+1} \varphi(s)| \,\mathrm{d}s \,\mathrm{d}x \\ &\le \frac{\varepsilon^{m\alpha}}{(m+1)!} \int_{(k-1)\varepsilon^\alpha}^{k\varepsilon^\alpha} \int_{(k-1)\varepsilon^\alpha}^{(k+1)\varepsilon^\alpha} |\partial_x^{m+1} \varphi(s)| \,\mathrm{d}s \,\mathrm{d}x \\ &= \frac{\varepsilon^{(m+1)\alpha}}{(m+1)!} \int_{(k-1)\varepsilon^\alpha}^{(k+1)\varepsilon^\alpha} |\partial_x^{m+1} \varphi(s)| \,\mathrm{d}s. \end{aligned}$$

The calculation for $\varphi(\cdot - \varepsilon^\alpha)$ works similarly. □

With this preparation we can now prove Lemma 3.34.

Proof of Lemma 3.34. Lemma 3.38 proves the statement $\|\varepsilon^{-1}L_\lambda \pi_\varepsilon \varphi\|_\infty \lesssim \|\mathcal{L}_\lambda \varphi\|_\infty \varepsilon^\alpha$. Thus it remains to bound the difference $\mathcal{R}(k) = \varepsilon^{-1}L_\lambda \pi_\varepsilon \varphi(k) - \pi_\varepsilon \mathcal{L}_\lambda \varphi(k)$ for $k \geq 3$. By the fundamental theorem of calculus we have

$$\begin{aligned}\pi_\varepsilon \mathcal{L}_\lambda \varphi(k) &= \int_{(k-1)\varepsilon^\alpha}^{k\varepsilon^\alpha} \partial_x (a_\lambda \partial_x \varphi)(x)\, \mathrm{d}x \\ &= a_\lambda(k\varepsilon^\alpha)\partial_x \varphi(k\varepsilon^\alpha) - a_\lambda((k-1)\varepsilon^\alpha)\partial_x\varphi((k-1)\varepsilon^\alpha),\end{aligned}$$

By using that

$$\partial_x \varphi(k\varepsilon^\alpha) - \partial_x \varphi((k-1)\varepsilon^\alpha) = \pi_\varepsilon \partial_x^2 \varphi(k),$$

we can split the error terms into

$$\mathcal{R}(k) = \mathcal{R}_1(k) + \mathcal{R}_2(k),$$

where

$$\begin{aligned}\mathcal{R}_1(k) &= (a_\lambda(k\varepsilon^\alpha) - a_\lambda((k-1)\varepsilon^\alpha))\left(\varepsilon^{-2\alpha}\partial^+ \pi_\varepsilon \varphi(k-1) - \partial_x \varphi((k-1)\varepsilon^\alpha)\right), \\ \mathcal{R}_2(k) &= a_\lambda(k\varepsilon^\alpha)\left(\varepsilon^{-2\alpha}(\partial^+\pi_\varepsilon\varphi(k) - \partial^+\pi_\varepsilon\varphi(k-1)) - \pi_\varepsilon \partial_x^2 \varphi(k)\right).\end{aligned}$$

For $\mathcal{R}_1(k)$ we use the Taylor expansion from Lemma 3.39 to first order ($m = 1$)

$$\begin{aligned}\partial^+ \pi_\varepsilon \varphi(k-1) &= \pi_\varepsilon \varphi(k) - \pi_\varepsilon \varphi(\cdot - \varepsilon^\alpha)(k) \\ &= \varepsilon^\alpha \pi_\varepsilon \partial_x \varphi(k) + R_1(\varphi, -\varepsilon)(k) \\ &= \varepsilon^\alpha (\varphi(k\varepsilon^\alpha) - \varphi((k-1)\varepsilon^\alpha)) + R_1(\varphi, -\varepsilon)(k).\end{aligned}$$

Hence, we can estimate the first order commutator by writing

$$\begin{aligned}&\varepsilon^{-2\alpha}\partial^+\pi_\varepsilon\varphi(k-1) - \partial_x\varphi((k-1)\varepsilon^\alpha) \\ &\quad = \varepsilon^{-\alpha}\Big(\varphi(k\varepsilon^\alpha) - \varphi((k-1)\varepsilon^\alpha)\Big) - \partial_x \varphi((k-1)\varepsilon^\alpha) + \varepsilon^{-2\alpha}R_1(\varphi, -\varepsilon)(k) \\ &\quad = -\frac{\varepsilon^{-\alpha}}{2}\int_{(k-1)\varepsilon^\alpha}^{k\varepsilon^\alpha} (k\varepsilon^\alpha - s)\partial_x^2 \varphi(s)\, \mathrm{d}s + \varepsilon^{-2\alpha} R_1(\varphi, -\varepsilon)(k),\end{aligned}$$

which yields with the bound from Lemma 3.39

$$|\mathcal{R}_1(k)| \lesssim |a_\lambda(k\varepsilon^\alpha) - a_\lambda((k-1)\varepsilon^\alpha)| \int_{(k-2)\varepsilon^\alpha}^{k\varepsilon^\alpha} |\partial_x^2 \varphi(x)|\, \mathrm{d}x.$$

If $\lambda = 0$, the error term $\mathcal{R}_1$ vanishes. Otherwise we have

$$|a_\lambda(k\varepsilon^\alpha) - a_\lambda((k-1)\varepsilon^\alpha)| \leq \begin{cases} \varepsilon^\alpha a'_\lambda((k-1)\varepsilon^\alpha), & 0 < \lambda < 1, \\ \varepsilon^\alpha a'_\lambda(k\varepsilon^\alpha), & \lambda \geq 1, \end{cases}$$

which implies

$$|\mathcal{R}_1(k)| \lesssim \varepsilon^\alpha \int_{(k-2)\varepsilon^\alpha}^{k\varepsilon^\alpha} x^{\lambda-1} |\partial_x^2 \varphi(x)| \, \mathrm{d}x.$$

For the above estimate we used the fact that $k \geq 3$. For the second error term $\mathcal{R}_2$ we have to expand to second order

$$\begin{aligned} \partial^+ \pi_\varepsilon \varphi(k) - \partial^+ \pi_\varepsilon \varphi(k-1) &= \pi_\varepsilon \Big(\varphi(\cdot + \varepsilon^\alpha) \Big)(k) - 2\pi_\varepsilon \varphi(k) + \pi_\varepsilon \Big(\varphi(\cdot - \varepsilon^\alpha) \Big)(k) \\ &= \varepsilon^{2\alpha} \pi_\varepsilon \partial_x^2 \varphi(k) + R_2(\varphi, \varepsilon)(k) - R_2(\varphi, -\varepsilon)(k). \end{aligned}$$

Hence, by the same argument as before, we obtain

$$|\mathcal{R}_2(k)| \lesssim \varepsilon^\alpha a_\lambda(k\varepsilon^\alpha) \int_{(k-2)\varepsilon^\alpha}^{(k+1)\varepsilon^\alpha} |\partial_x^3 \varphi(x)| \, \mathrm{d}x \lesssim \varepsilon^\alpha \int_{(k-2)\varepsilon^\alpha}^{(k+1)\varepsilon^\alpha} x^\lambda |\partial_x^3 \varphi(x)| \, \mathrm{d}x.$$

□

3.3.4. Approximate weak solutions

The estimates from Corollary 3.31 allow us to control powers of $\mathcal{L}_\lambda$ of solutions to equation (NP'). The error term in the replacement Lemma however is not of this form. Hence we first need an interpolation inequality for the operator $\mathcal{L}_\lambda$.

Lemma 3.40. *Let $\varphi \in C^2(\mathbb{R}_+)$. Then it holds*

$$\|a_\lambda \partial_x \varphi\|_\infty \lesssim \left(\|a_\lambda \varphi\|_\infty + \|\partial_x a_\lambda \varphi\|_1 \right)^{\frac{1}{2}} \|\mathcal{L}_\lambda \varphi\|_\infty^{\frac{1}{2}}.$$

Proof. Define the function

$$\psi(x) = \int_0^x a_\lambda(y) \partial_y \varphi(y) \, \mathrm{d}y.$$

Then we have $\partial_x \psi = a_\lambda \partial_x \varphi$, $\partial_x^2 \psi = \mathcal{L}_\lambda \varphi$ and by standard interpolation, the inequality

$$\|a_\lambda \partial_x \varphi\|_\infty = \|\partial_x \psi\|_\infty \lesssim \|\psi\|_\infty^{\frac{1}{2}} \|\partial_x^2 \psi\|_\infty^{\frac{1}{2}} = \|\psi\|_\infty^{\frac{1}{2}} \|\mathcal{L}_\lambda \varphi\|_\infty^{\frac{1}{2}}$$

holds. Integrating by parts we have

$$\psi(x) = a_\lambda(x)\varphi(x) - a_\lambda \varphi(0) - \int_0^x \partial_y a_\lambda(y) \varphi(y) \, \mathrm{d}y,$$

which implies $\|\psi\|_\infty \lesssim \|a_\lambda \varphi\|_\infty + \|\partial_x a_\lambda \varphi\|_1$. □

With this we can prove that $\mathcal{U}_\varepsilon$ is approximately a weak solution of equation (NP').

Proof of Proposition 3.35. Let $\varphi(t,\cdot) = \mathcal{T}_\lambda(T-t)f$, $f \in C_c^\infty((0,T)\times\mathbb{R}_+)$. By (3.56) we have

$$\int_0^T \int_0^\infty \mathcal{U}_\varepsilon(\partial_t\varphi + \mathcal{L}_\lambda\varphi)\,\mathrm{d}x\,\mathrm{d}t$$
$$= -\int_0^\infty \mathcal{U}_\varepsilon(0,x)\varphi(0,x)\,\mathrm{d}x + \varepsilon^{-\alpha}\int_0^T \sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t,k)\Big(\pi_\varepsilon\mathcal{L}_\lambda\varphi(t,k) - \varepsilon^{-1}L_\lambda\pi_\varepsilon\varphi(t,k)\Big)\,\mathrm{d}t,$$

hence we have to estimate the term

$$R_\varepsilon = \int_0^T \varepsilon^{-\alpha}\sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t,k)|\mathcal{R}_\varepsilon|(t,k)\,\mathrm{d}t, \tag{3.62}$$

where $|\mathcal{R}_\varepsilon| = |\pi_\varepsilon\mathcal{L}_\lambda\varphi - \varepsilon^{-1}L_\lambda\pi_\varepsilon\varphi|$. Let $\sigma(\varepsilon)$ be a non-negative increasing function with $\lim_{\varepsilon\to 0}\sigma(\varepsilon) = 0$ and $\theta(\varepsilon)$ a non-negative decreasing function with $\lim_{\varepsilon\to 0}\varepsilon^\alpha\theta(\varepsilon) = 0$. Then we first split the integration into two regions $[0,\sigma(\varepsilon)]$ and $(\sigma(\varepsilon),\infty)$. In the first region we use the first statement from Lemma 3.34, which implies $|\mathcal{R}_\varepsilon| \lesssim \varepsilon^\alpha\|\mathcal{L}_\lambda\varphi\|_\infty$, to estimate

$$\int_0^{\sigma(\varepsilon)} \varepsilon^{-\alpha}\sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t,k)|\mathcal{R}_\varepsilon|(t,k)\,\mathrm{d}t \lesssim \|\mathcal{L}_\lambda\varphi\|_\infty \int_0^{\sigma(\varepsilon)} \sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t,k)\,\mathrm{d}t$$
$$= \|U_{0,\varepsilon}\|_1\|\mathcal{L}_\lambda\varphi\|_\infty\sigma(\varepsilon).$$

For the second integral, we split the summation into three regions

$$\sum_{1\le k\le\theta(\varepsilon)} + \sum_{\theta(\varepsilon)<k\lesssim\varepsilon^{-\alpha}} + \sum_{k\gtrsim\varepsilon^{-\alpha}} = \mathrm{I} + \mathrm{II} + \mathrm{III}.$$

In the first region we apply the estimate (3.21) from Theorem 3.7 for U that yields $U_\varepsilon(\varepsilon^{-1}t,k) \lesssim \|U_{0,\varepsilon}\|_1(\sigma(\varepsilon)\varepsilon^{-1})^{-\alpha}$, since $t \ge \sigma(\varepsilon)$, and the estimate for $\mathcal{R}_\varepsilon$ from above to obtain

$$\mathrm{I} \lesssim \|U_{0,\varepsilon}\|_1\|\mathcal{L}_\lambda\varphi\|_\infty\varepsilon^\alpha\sigma(\varepsilon)^{-\alpha}\theta(\varepsilon).$$

For the other two sums we use the estimate from Lemma 3.34 that yields

$$|\mathcal{R}_\varepsilon| \lesssim \varepsilon^\alpha\int_{(k-2)\varepsilon^\alpha}^{(k+1)\varepsilon^\alpha} x^{\lambda-1}|\partial_x^2\varphi| + x^\lambda|\partial_x^3\varphi|\,\mathrm{d}x.$$

For the second order term we can apply Lemma 3.37 to conclude $x^{\lambda-1}|\partial_x^2\varphi| \lesssim \|\mathcal{L}_\lambda\varphi\|_\infty x^{-1}$. Next we apply Lemma 3.40 to the function $\mathcal{L}_\lambda\varphi$ to obtain

$$\|a_\lambda\partial_x\mathcal{L}_\lambda\varphi\|_\infty \lesssim (\|a_\lambda\mathcal{L}_\lambda\varphi\|_\infty + \|\partial_x a_\lambda\mathcal{L}_\lambda\varphi\|_1)^{\frac{1}{2}}\,\|\mathcal{L}_\lambda^2\varphi\|_\infty^{\frac{1}{2}}.$$

We calculate the term on the left

$$a_\lambda\partial_x\mathcal{L}_\lambda\varphi = a_\lambda\partial_x^2(a_\lambda\partial_x\varphi) = a_\lambda(a_\lambda\partial_x^3\varphi + 2\partial_x a_\lambda\partial_x^2\varphi + \partial_x^2 a_\lambda\partial_x\varphi).$$

By using Lemma 3.37 it holds

$$\left|2\partial_x a_\lambda \partial_x^2\varphi + \partial_x^2 a_\lambda \partial_x\varphi\right|(x) \lesssim \|\mathcal{L}_\lambda\varphi\|_\infty x^{-1},$$

which then implies the bound

$$|a_\lambda \partial_x^3\varphi|(x) \lesssim (\|a_\lambda \mathcal{L}_\lambda\varphi\|_\infty + \|\partial_x a_\lambda \mathcal{L}_\lambda\varphi\|_1)^{\frac{1}{2}} \|\mathcal{L}_\lambda^2\varphi\|_\infty^{\frac{1}{2}} x^{-\lambda} + \|\mathcal{L}_\lambda\varphi\|_\infty x^{-1-\lambda}.$$

In particular the negative powers in the last expression are bounded for $x \gtrsim 1$, thus we can conclude

$$\begin{aligned}\mathrm{III} &\lesssim \|U_{0,\varepsilon}\|_1((\|a_\lambda \mathcal{L}_\lambda\varphi\|_\infty + \|\partial_x a_\lambda \mathcal{L}_\lambda\varphi\|_1)^{\frac{1}{2}} \|\mathcal{L}_\lambda^2\varphi\|_\infty^{\frac{1}{2}} + \|\mathcal{L}_\lambda\varphi\|_\infty)\varepsilon^\alpha \\ &\leq \|U_{0,\varepsilon}\|_1 C_1(\varphi, T)\varepsilon^\alpha,\end{aligned}$$

where

$$C_1(\varphi, T) = \sup_{t\in[0,T]} ((\|a_\lambda \mathcal{L}_\lambda\varphi\|_\infty + \|\partial_x a_\lambda \mathcal{L}_\lambda\varphi\|_1)^{\frac{1}{2}} \|\mathcal{L}_\lambda^2\varphi\|_\infty^{\frac{1}{2}} + \|\mathcal{L}_\lambda\varphi\|_\infty).$$

To estimate the term II we use again the L^∞ estimate for U_ε and get

$$\begin{aligned}\mathrm{II} &\lesssim \|U_{0,\varepsilon}\|_1 C_1(\varphi, T)\varepsilon^\alpha \sigma(\varepsilon)^{-\alpha} \int_{\varepsilon^\alpha\theta(\varepsilon)}^1 x^{-1-\lambda}\,\mathrm{d}x \lesssim \|U_{0,\varepsilon}\|_1 C_1(\varphi)\varepsilon^\alpha\sigma(\varepsilon)^{-\alpha}(\varepsilon^\alpha\theta(\varepsilon))^{-\lambda} \\ &= \|U_{0,\varepsilon}\|_1 C_1(\varphi, T)\varepsilon^{\alpha(1-\lambda)}\sigma(\varepsilon)^{-\alpha}\theta(\varepsilon)^{-\lambda}.\end{aligned}$$

In summary we obtain the following estimate for the full error term in (3.62)

$$R_\varepsilon \lesssim \|U_{0,\varepsilon}\|_1 C_1(\varphi, T)\Big(\sigma(\varepsilon) + T\Big(\varepsilon^\alpha\sigma(\varepsilon)^{-\alpha}\theta(\varepsilon) + \varepsilon^{\alpha(1-\lambda)}\sigma(\varepsilon)^{-\alpha}\theta(\varepsilon)^{-\lambda} + \varepsilon^\alpha\Big)\Big).$$

To make the right-hand side converge to zero we need to choose appropriate functions σ and θ. We make the ansatz $\sigma(\varepsilon) = \varepsilon^a$, $\theta(\varepsilon) = \varepsilon^{-\alpha+b}$ for some $a, b > 0$. This leads to the requirements

$$-\alpha a + b > 0 \qquad \text{and} \qquad \alpha(1-a) - \lambda b > 0,$$

which are satisfied for some a, b small enough with $b > \alpha a$. Finally, we have to check that we have good control of the norms of $\mathcal{L}_\lambda\varphi$ involved in the quantity $C_1(\varphi, T)$. This follows easily from Corollary 3.31 and the explicit formula for φ, since we have

$$\begin{aligned}\|\mathcal{L}_\lambda\varphi(t,\cdot)\|_\infty = \|\mathcal{L}_\lambda\mathcal{T}_\lambda(T-t)f\|_\infty &\leq \int_0^{T-t} \|\mathcal{L}_\lambda\mathcal{S}_\lambda(T-t-s)f(s,\cdot)\|_\infty\,\mathrm{d}s \\ &\leq T\|\mathcal{L}_\lambda f\|_{L^\infty(\mathbb{R}_+^2)},\end{aligned}$$

and similarly for the other norms. Hence we conclude that

$$\varepsilon^{-\alpha}\int_0^T \sum_{k=1}^\infty U_\varepsilon(\varepsilon^{-1}t, k)|\mathcal{R}_\varepsilon|(t,k)\,\mathrm{d}t \leq \|U_{0,\varepsilon}\|_1 C(T, f)\varepsilon^r,$$

for some exponent $r > 0$, finishing the proof. □

3.4. Convergence to self-similarity

3.4.1. Convergence of the empirical measure and moments

Let u be a solution to equation (DP). In this subsection we apply Corollary 3.9 to the tail distribution of u to extract statements regarding weak convergence and convergence of moments. To that end, let $\sigma\colon \overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}_+$ and define the empirical measure associated to u and σ by

$$\mu(t) = \sigma(t) \sum_{k=1}^{\infty} u(t,k)\delta_{\sigma(t)^{-1}k}. \tag{3.63}$$

Then with the notion of weak convergence in Definition 3.3 we have the following result.

Proposition 3.41 (Weak convergence of the empirical measure)**.** *Let u be a solution to equation* (DP) *with $M_1[u] = \rho$, $\sigma\colon \overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}_+$ with the property $\lim_{t\to\infty} t^{-\alpha}\sigma(t) = 1$, μ the associated empirical measure as above and g_λ as in* (3.7)*. Then for $0 \le \lambda < 1$, $\mu(t) \rightharpoonup \rho\, g_\lambda$ with respect to $\mathcal{C}$ as $t \to \infty$, whereas for $1 \le \lambda < 2$, $\mu(t) \rightharpoonup \rho\, g_\lambda$ with respect to $\mathcal{C}_0$.*

Proof. We first consider the case $0 \le \lambda < 1$. Then for $x \ge 0$ we have

$$\mu\big(t,(x,\infty)\big) = \sigma(t) \sum_{k=\lfloor\sigma(t)x\rfloor+1}^{\infty} u(t,k) = t^{-\alpha}\sigma(t)\hat{U}(t,x(t)),$$

with $x(t) = t^{-\alpha}\sigma(t)x$ and $\hat{U}$ as in Corollary 3.9. In particular $\mu(t)$ is bounded in total variation. By the assumption on σ we have $x(t) \to x$ as $t \to \infty$. Because the convergence in Corollary 3.9 is uniform and the limit is a continuous function, this implies that $\hat{U}(t,x(t)) \to \rho\,\mathcal{G}_\lambda(x)$ as $t \to \infty$. Thus we conclude

$$\mu\big(t,(x,\infty)\big) \to \rho\,\mathcal{G}_\lambda(x) = \rho\int_x^\infty g_\lambda(y)\,\mathrm{d}y, \qquad \text{as } t \to \infty,$$

and, more generally,

$$\mu\big(t,(a,b]\big) = \hat{U}(t,a) - \hat{U}(t,b) \to \rho\big(\mathcal{G}_\lambda(a) - \mathcal{G}_\lambda(b)\big) = \rho\int_a^b g_\lambda(x)\,\mathrm{d}x, \qquad \text{as } t \to \infty,$$

for $0 \le a < b < \infty$. Note that in the case $a = 0$ the integral over $[a,b]$ coincides with the integral over $(a,b]$ since $\mu(t)(\{0\}) = 0$. By linearity and tightness of the measure μ (the first moment is constant in time) we conclude that

$$\int_0^\infty \chi(x)\mu(t,x) \to \int_0^\infty \chi(x)\rho(x)\,\mathrm{d}x, \qquad \text{as } t \to \infty,$$

for all functions $\chi = \sum_{k=0}^\infty \theta_k \mathbf{1}_{I_k}$, $I_0 = [0,a_1]$, $I_k = (a_k,a_{k+1}]$, $\theta_k \le C$, $a_k < a_{k+1}$, $a_k \to \infty$. Then by approximation (Corollary 3.9 implies that μ is uniformly bounded) the above convergence holds for bounded continuous functions, and using the bound on the first moment of μ the convergence is also extended to the class of functions $\mathcal{C}$. The argument in the case $1 \leq \lambda < 2$ works in the same way, except that the convergence on the level of characteristic functions only holds for functions with support outside of 0, and thus we can only approximate continuous functions vanishing at 0. $\square$

Corollary 3.42. *Let $u = u(t,k)$ be a solution to equation* (DP) *with $M_0[u](0) = 1, M_1[u] = \rho$. Then for every $\nu \in (0,1]$ we have*

$$\lim_{t\to\infty} t^{\alpha(1-\nu)} M_\nu[u](t) = \rho \int_0^\infty x^\nu g_\lambda(x)\,\mathrm{d}x,$$

and in the case $0 \leq \lambda < 1$ the above identity also holds for $\nu = 0$.

Proof. For $\nu < 1$ this follows from Proposition 3.41 with $\sigma(t) = t^\alpha$, applying the weak convergence to the test function $f(x) = x^\nu$, because then

$$t^{\alpha(1-\nu)} M_\nu[u] = \int_0^\infty x^\nu \,\mathrm{d}\mu(t,x) \to \rho \int_0^\infty x^\nu g_\lambda(x)\,\mathrm{d}x, \qquad \text{as } t \to \infty.$$

The statement for $\nu = 1$ follows directly from conservation of the first moment. $\square$

The next goal is to show that a result similar to Corollary 3.42 holds for higher moments in the case $\lambda \geq 1$. The main idea is that differentiating a high moment in time gives a lower moment so we can bootstrap estimates from lower to higher moments.

Lemma 3.43. *Let $\lambda \geq 1$, $\nu > 1$ and u be a solution to equation* (DP) *with $M_1[u] = \rho$ and $M_\nu[u_0] < \infty$. Then there exists an explicit positive constant $C = C(\nu, \lambda, \rho)$ such that*

$$\lim_{t\to\infty} t^{\alpha(1-\nu)} M_\lambda[u] = C.$$

Proof. We show that if the statement holds for $\nu + \lambda - 2$ with constant C, then it holds for ν with constant $\frac{\nu}{\alpha} C$. To that end we use that

$$\frac{\mathrm{d}}{\mathrm{d}t} M_\nu[u] = \sum_{k=1}^\infty k^\lambda \Delta_{\mathbb{N}}(k^\nu) u(t,k).$$

While clear on a formal level, the integration by parts here poses a potential problem, since the boundary term contains large powers of k. However, since the equation is linear this can be resolved by proving the above identity for the fundamental solution to equation (DP) and using a representation similar to (3.38). Next, by Lemma 3.14 we have

$$\frac{k^\lambda \Delta_{\mathbb{N}}(k^\nu)}{k^{\nu+\lambda-2}} \to \nu(\nu-1).$$

Hence, for $\varepsilon > 0$ there exists k_0 such that for all $k \geq k_0$ it holds

$$\Big(\nu(\nu-1)-\varepsilon\Big)k^{\nu+\lambda-2} \leq k^\lambda \Delta_{\mathbb{N}}(k^\nu) \leq \Big(\nu(\nu-1)+\varepsilon\Big)k^{\nu+\lambda-2},$$

Therewith, we can estimate

$$\begin{aligned}
\sum_{k=1}^{\infty} k^\lambda \Delta_{\mathbb{N}}(k^\nu)u(t,k) &\geq \sum_{k=1}^{k_0} k^\lambda \Delta_{\mathbb{N}}(k^\nu)u(t,k) + (\nu(\nu-1)-\varepsilon) \sum_{k=k_0+1}^{\infty} k^{\nu+\lambda-2}u(t,k) \\
&= \sum_{k=1}^{k_0} (k^\lambda \Delta_{\mathbb{N}}(k^\nu) - (\nu(\nu-1)-\varepsilon)k^{\nu+\lambda-2})u(t,k) \\
&\quad + (\nu(\nu-1)-\varepsilon)M_{\nu+\lambda-2}[u].
\end{aligned}$$

In the finite sum we estimate $u(t,k) \leq M_0[u] \leq Ct^{-\alpha}$, while using the assumption for $M_{\nu+\lambda-2}[u]$ to conclude that

$$\begin{aligned}
\liminf_{t\to\infty} t^{\alpha(1-\nu)+1} \sum_{k=1}^{\infty} k^\lambda \Delta_{\mathbb{N}}(k^\nu)u(t,k) &\geq \liminf_{t\to\infty} t^{\alpha(1-\nu)+1}(\nu(\nu-1)-\varepsilon)M_{\nu+\lambda-2}[u] \\
&= (\nu(\nu-1)-\varepsilon)C.
\end{aligned}$$

Note that the finite sum vanishes in the limit since $t^{-\alpha} \cdot t^{\alpha(1-\nu)+1} = t^{1-\alpha\nu} \to 0$, because $\alpha \geq 1, \nu > 1$. By an analogous computation we also have the upper bound

$$\limsup_{t\to\infty} t^{\alpha(1-\nu)+1} \sum_{k=1}^{\infty} k^\lambda \Delta_{\mathbb{N}}(k^\nu)u(t,k) \leq (\nu(\nu-1)+\varepsilon)C.$$

To compute the limit of $t^{\alpha(1-\nu)}M_\lambda[u]$, the lower and upper bound from above imply that there exists $t_0 > 0$ such that for all $t \geq t_0$ it holds

$$(\nu(\nu-1)C-\varepsilon)t^{\alpha(\nu-1)-1} \leq \frac{\mathrm{d}}{\mathrm{d}t}M_\nu[u] \leq (\nu(\nu-1)C+\varepsilon)t^{\alpha(\nu-1)-1}.$$

Integrating these inequalities from t_0 to t and letting $t \to \infty$ we obtain that

$$\frac{(\nu(\nu-1)C-\varepsilon)}{\alpha(\nu-1)} \leq \liminf_{t\to\infty} t^{\alpha(1-\nu)}M_\lambda[u] \leq \liminf_{t\to\infty} t^{\alpha(1-\nu)}M_\lambda[u] \leq \frac{(\nu(\nu-1)C+\varepsilon)}{\alpha(\nu-1)},$$

which shows that the limit exists and is equal to $\frac{\nu}{\alpha}C$. The full statement of the Lemma is then easily obtained by induction. Let I_n for $n \geq 0$ be defined as

$$I_n = \Big(n(2-\lambda), (n+1)(2-\lambda)\Big].$$

Then let n_0 be the smallest n such that $I_n \cap (1,\infty) \neq \emptyset$. Then for $\nu \in I_{n_0}$ with $\nu \leq 1$ there is nothing to show because Corollary 3.42 applies while for $\nu \in I_{n_0} \cap (1,\infty)$ we have $0 < \nu + \lambda - 2 \leq 1$ by construction. Hence by Corollary 3.42 the desired limit holds for $M_{\nu+\lambda-2}$, and hence by the above considerations also for $M_\nu[u]$. Thus the statement holds for all $\nu \in I_{n_0}$. Then for all $n > n_0$ we have $\nu > 1$ and $\nu + \lambda - 2 \in I_{n-1}$ by construction, enabling the inductive argument. This finishes the proof. $\square$

Since our main interest for the rest of this section is in the quantity $M_\lambda[u]$, we summarize our findings in the following Corollary.

Corollary 3.44. *Let $u = u(t,k)$ be a solution to equation (DP) with $M_0[u](0) = 1, M_1[u] = \rho$. Then there exists a constant $C = C(\lambda, \rho)$ such that*

$$\lim_{t\to\infty} t^{\alpha(1-\lambda)} M_\lambda[u] = C.$$

3.4.2. Self-similar behavior: Proof of Theorem 3.4

Recall that equation (DP) and equation (EDG_λ) are linked by the time change τ defined in (3.14). Therewith, the function $u(\tau, k)$ defined by $u(\tau(t), k) = c_k(t)$ for $k \geq 1$ is a solution to equation (DP) if $c_k(t)$ is a solution to the system (EDG_λ). Then the asymptotic behavior of the moments of u implies the following result.

Proposition 3.45. *Let $0 \leq \lambda < 2$ and set $\beta = (3-2\lambda)^{-1}$. Then for every $\lambda \in [0,2)$ and $\rho \in (0,\infty)$ there exists $C = C(\lambda,\rho) > 0$ such that the following statements hold:*

1. *If $0 \leq \lambda < 3/2$ and c is a global solution to equation (EDG_λ) with $M_1[c] = \rho$, then*

$$\lim_{t\to\infty} t^{-\frac{\beta}{\alpha}} \tau(t) = C.$$

2. *If $\lambda = 3/2$ and c is a global solution to equation (EDG_λ) with $M_1[c] = \rho$, then for every $0 < \varepsilon < C$ it holds*

$$\lim_{t\to\infty} \exp(-(C+\varepsilon)t)\tau(t) = 0, \quad \lim_{t\to\infty} \exp(-(C-\varepsilon)t)\tau(t) = \infty.$$

3. *If $3/2 < \lambda \leq 2$ and c is a solution to equation (EDG_λ) with blow-up time t^*, then it holds*

$$\lim_{t\to t^*} (t^* - t)^{-\frac{\beta}{\alpha}} \tau(t) = C.$$

Proof. Because $\tau(t) \to \infty$ and Corollary 3.44 we have that for every small $\varepsilon > 0$ there exists $t_0 > 0$ such that

$$(C-\varepsilon)\tau(t)^{\alpha(\lambda-1)} \leq M_\lambda[u(\tau(t),\cdot)] \leq (C+\varepsilon)\tau(t)^{\alpha(\lambda-1)},$$

for $t \geq t_0$, where C is as in Corollary 3.44. Using these refined bounds in the differential equation for τ, we obtain

$$(C-\varepsilon)\tau^{\alpha(\lambda-1)} \leq \dot{\tau} \leq (C+\varepsilon)\tau^{\alpha(\lambda-1)}, \tag{3.64}$$

for $t \geq t_0$. Dividing by $\tau^{\alpha(\lambda-1)}$ and integrating from t_0 to t then yields

$$\left(\tau(t_0)^{\frac{\alpha}{\beta}} + \frac{\alpha}{\beta}(C-\varepsilon)(t-t_0)\right)^{\frac{\beta}{\alpha}} \leq \tau(t) \leq \left(\tau(t_0)^{\frac{\alpha}{\beta}} + \frac{\alpha}{\beta}(C+\varepsilon)(t-t_0)\right)^{\frac{\beta}{\alpha}},$$

and passing to the limit $t \to \infty$ we get

$$\left(\frac{\alpha}{\beta}(C-\varepsilon)\right)^{\frac{\beta}{\alpha}} \le \liminf_{t\to\infty} t^{-\frac{\beta}{\alpha}}\tau(t) \le \limsup_{t\to\infty} t^{-\frac{\beta}{\alpha}}\tau(t) \le \left(\frac{\alpha}{\beta}(C+\varepsilon)\right)^{\frac{\beta}{\alpha}},$$

which gives the desired statement with constant $\left(\frac{\alpha}{\beta}C\right)^{\frac{\beta}{\alpha}}$ after letting $\varepsilon \to 0$. In the case $\lambda = 3/2$ we have $\alpha(\lambda-1) = 1$ and the inequality (3.64) gives

$$\tau(t_0)\exp((C-\varepsilon)t) \le \tau(t) \le \tau(t_0)\exp((C+\varepsilon)t),$$

which yields the second statement. For the third statement, let t^* denote the blow-up time of τ. Then dividing the inequalities (3.64) by $\tau^{\alpha(\lambda-1)}$ and integrating from t to t^* for $t_0 < t < t^*$ we get

$$\left(-\frac{\alpha}{\beta}(C+\varepsilon)(t^*-t)\right)^{\frac{\beta}{\alpha}} \le \tau(t) \le \left(-\frac{\alpha}{\beta}(C-\varepsilon)(t^*-t)\right)^{\frac{\beta}{\alpha}},$$

which implies the third statement. □

With these preparations we can prove Theorem 3.4. Recall that for a given solution c of equation (EDG$_\lambda$) and a scaling function $s\colon \overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}_+$ the corresponding empirical measure is given by

$$\mu_c(t) = s(t)\sum_{k=1}^{\infty} c_k(t)\delta_{s(t)^{-1}k}.$$

Proof of Theorem 3.4. We start with the case $0 \le \lambda < \frac{3}{2}$. Let $u(\tau(t),k) = c_k(t)$, define the function σ by $\sigma(\tau(t)) = s(t)$ and rewrite the empirical measure in terms of u and σ as

$$\mu(t) = \sigma(\tau(t))\sum_{k=1}^{\infty} u(\tau(t),k)\delta_{\sigma(\tau(t))^{-1}k}.$$

Note that σ satisfies

$$\tau(t)^{-\alpha}\sigma(\tau(t)) = \tau(t)^{-\alpha}s(t) = C^{-1}\left(t^{-\frac{\beta}{\alpha}}\tau(t)\right) \to C^{-1}C = 1,$$

as $t \to \infty$ by Proposition 3.45. Hence Proposition 3.41 applies and the desired convergence result follows. In the case $\lambda = \frac{3}{2}$ we simply take the scaling function $s(t) = \tau(t)^\alpha$, then the statement follows immediately from Proposition 3.45 and Proposition 3.41. The case $\frac{3}{2} < \lambda < 2$ follows along the same lines as in the case $0 \le \lambda < \frac{3}{2}$. □

A. Supplementary results

A.1. A weighted logarithmic Sobolev inequality

We introduce the relative entropy and Fisher information for any test function $f : \mathbb{R}_+ \to \mathbb{R}_+$ with respect to the self-similar profile from (3.24) by

$$\operatorname{Ent}_{\gamma_\lambda}(f) = \int f \log f \, \mathrm{d}\gamma_\lambda \qquad \text{and} \qquad \mathcal{E}_{\gamma_\lambda}(f, f) = \int |x|^\lambda |f'|^2 \, \mathrm{d}\gamma_\lambda. \tag{A.1}$$

Then, we have the following result.

Lemma A.1 (Weighted log-Sobolev inequality)**.** *For any $\lambda \in [0, 2]$ exists $C_{\mathrm{LSI}}(\lambda)$ such that the measure $\gamma_\lambda(\cdot) = \gamma_\lambda(1, \cdot)$ from (3.24) satisfies for all $f : \mathbb{R}_+ \to \mathbb{R}_+$ with $\mathcal{E}_{\gamma_\lambda}(f, f) < \infty$ the logarithmic Sobolev inequality*

$$\operatorname{Ent}_{\gamma_\lambda}(f^2) \le C_{\mathrm{LSI}} \mathcal{E}_{\gamma_\lambda}(f, f). \tag{A.2}$$

Proof. We are going to apply [4, Theorem 3], which generalizes the result from [8] to an applicable form for the present situation. By setting $\mu = Z_\lambda \gamma_\lambda$ and $\mathrm{d}\nu(x) = Z_\lambda |x|^\lambda \gamma_\lambda$, we have to show that

$$\begin{aligned} B_- &= \sup_{x<1} \ \mu([0, x]) \log\left(1 + \frac{e^2}{\mu([0, x])}\right) \int_x^1 \frac{1}{\nu(x)} \, \mathrm{d}x < \infty; \\ B_+ &= \sup_{x>1} \ \mu([x, \infty)) \log\left(1 + \frac{e^2}{\mu([x, \infty))}\right) \int_1^x \frac{1}{\nu(x)} \, \mathrm{d}x < \infty. \end{aligned}$$

Then, we have that $C_{\mathrm{LSI}}(\lambda) \le 4 \max\{B_-, B_+\}$, where we use that the particular choice of the median in the proof of [4, Theorem 3] does not enter the upper bound.

Let us first consider B_+, for which we show that asymptotically for $x \to \infty$ it is equivalent to

$$\mu([x, \infty)) \simeq (2 - \lambda) x^{\lambda - 1} \exp(-\alpha^2 x^{2-\lambda}) \quad \text{and} \quad \int_1^x \frac{1}{\nu(x)} \, \mathrm{d}x \simeq (2 - \lambda) x^{-1} \exp(\alpha^2 x^{2-\lambda}).$$

Therewith, the claim follows directly by plugging the above identities into the definition of B_+. Because both sides are strictly positive and go to 0, respectively ∞, as $x \to \infty$, it suffices to show that the derivatives are asymptotically comparable by L'Hospital

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x} \mu([x, \infty)) &= - \exp(-\alpha^2 x^{2-\lambda}), \\ \frac{\mathrm{d}}{\mathrm{d}x} \int_0^x \frac{1}{\nu(x)} \, \mathrm{d}x &= x^{-\lambda} \exp(\alpha^2 x^{2-\lambda}), \end{aligned}$$

while

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big((2-\lambda)x^{\lambda-1}\exp(-\alpha^2x^{2-\lambda})\Big) = (2-\lambda)(\lambda-1)x^{\lambda-2}\exp(-\alpha^2x^{2-\lambda}) - \exp(-\alpha^2x^{2-\lambda}),$$
$$\frac{\mathrm{d}}{\mathrm{d}x}\Big((2-\lambda)x^{-1}\exp(\alpha^2x^{2-\lambda})\Big) = -(2-\lambda)x^{-2}\exp(\alpha^2x^{2-\lambda}) + x^{-\lambda}\exp(\alpha^2x^{2-\lambda}),$$

which gives the correct asymptotic, since $\lambda < 2$. Similar arguments show that for $x \ll 1$ it holds

$$\mu([0,x]) \simeq x \quad \text{and} \quad \int_x^1 \frac{1}{\nu(x)}\,\mathrm{d}x \simeq \begin{cases} C_\lambda - \frac{x^{1-\lambda}}{1-\lambda}, & \text{for } \lambda \in [0,1); \\ -\log x, & \text{for } \lambda = 1; \\ \frac{x^{-(\lambda-1)}}{\lambda-1}, & \text{for } \lambda \in (1,2]. \end{cases} \tag{A.3}$$

From here, we also deduce that $B_- < \infty$ proving the claim by an application of [4, Theorem 3]. □

Proof of (3.55). By rescaling it is enough to prove (3.55) for $t = 1$ and we drop the subscript 1 for now. By setting $f(x)^2 = \frac{\mathrm{d}\mu}{\rho\,\mathrm{d}\gamma}$, we see that (3.55) is equivalent to

$$\mathrm{Ent}_{\rho\gamma_\lambda}(f^2) = \rho\,\mathrm{Ent}_{\gamma_\lambda}(f^2) \le 4\,C_{\mathrm{LSI}}\,\rho\mathcal{E}_{\gamma_\lambda}(f,f) = 4\,C_{\mathrm{LSI}}\mathcal{E}_{\rho\gamma_\lambda}(f,f),$$

with Ent_γ and $\mathcal{E}_\gamma$ as in (A.1). Hence, the logarithmic Sobolev inequality (3.55) follows from Lemma A.1. □

A.2. Arzela-Ascoli Theorem for discontinuous functions

Proposition A.2. *Let (X,d) be a compact separable metric space and $f_n\colon X \to \mathbb{R}$ a sequence of functions with the following properties:*

1. *For all $x \in X$, $f_n(x)$ is a bounded sequence.*
2. *For each $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ and $\delta > 0$ such that*
$$|f_n(x) - f_n(y)| \le \varepsilon,$$
for all $x, y \in X$ with $d(x,y) \le \delta$ and $n \ge n_0$.

Then there exists a continuous function $f\colon X \to \mathbb{R}$ and a subsequence $n_k \to \infty$ such that $f_{n_k} \to f$ uniformly.

Proof. Let $Z \subset X$ be a countable dense subset. Then by the first property, $f_n(z)$ is a bounded sequence for all $z \in Z$, and by Bolzano-Weierstrass and a diagonal argument there exists a subsequence (not relabelled) such that $f_n(z)$ is a Cauchy sequence for all $z \in Z$. Next we show that this implies that $f_n(x)$ is a Cauchy sequence for all $x \in X$. Indeed, let $x \in X$ and $\varepsilon > 0$. By the second property there exists a n_0 and $\delta > 0$ with $|f_n(x) - f_n(y)| \le \varepsilon$ for all $x, y \in X$ with $d(x,y) \le \delta$ and $n \ge n_0$. Then by density there exists $z \in Z$ with $d(x,z) \le \delta$ and because $f_n(z)$ is a Cauchy sequence there exists $n_1 \ge n_0$ with $|f_n(z) - f_m(z)| \le \varepsilon$ for $n, m \ge n_1$. For such n, m we have by triangle inequality

$$|f_n(x) - f_m(x)| \le |f_n(z) - f_m(z)| + |f_n(z) - f_n(x)| + |f_m(z) - f_m(x)| \le 3\varepsilon.$$

This shows that $f_n(x)$ is a Cauchy-sequence and thus has a limit $f(x) = \lim_{n\to\infty} f_n(x)$. The continuity of f easily follows from the second property of f_n by letting $n \to \infty$. To show uniform convergence let $\varepsilon > 0$ and choose δ, n_0 according to the second property. Then by compactness we can cover X with finitely many δ-balls around some points x_k, $k = 1, .., N$ for some $N \in \mathbb{N}$. Then by convergence there exists $n_1 \ge n_0$ such that $|f_n(x_k) - f(x_k)| \le \varepsilon$ for all $k = 1, .., N$. Because by construction for every $x \in X$ there exists k with $d(x, x_k) \le \delta$ we have by triangle inequality for all $n \ge n_1$

$$|f_n(x) - f(x)| \le |f_n(x_k) - f(x_k)| + |f_n(x_k) - f_n(x)| + |f(x_k) - f(x)| \le 3\varepsilon,$$

where the third term is also smaller than ε by letting $n \to \infty$ in the second property. This shows uniform convergence. □

A.3. The discrete heat equation with absorbing boundary

In the special case $\lambda = 0$ we can analyse equation (DP) directly, since the fundamental solution is explicit. The equation then becomes the discrete heat equation with absorbing boundary

$$\begin{cases} \partial_t u = \Delta_{\mathbb{N}} u, & \text{on } \mathbb{N}, \\ u(t,0) = 0, & \text{for } t \ge 0, \\ u(0,k) = c_k, & \text{for } k \ge 1. \end{cases} \tag{A.4}$$

The fundamental solution ψ of (A.4) is obtained from the one φ of the whole lattice $\mathbb{Z}$ by reflection and satisfies $\psi(t,k,l) = \varphi(t,k-l) - \varphi(t,k+l)$. The fundamental solution to the discrete heat equation on $\mathbb{Z}$ is obtained as $\varphi(t,k) = \exp(2t) I_k(2t)$ by using Fourier series and an integral representation formula for the modified Bessel's I_k of the first kind (cf. Lemma 1.29). Hence, the solution u to (A.4) is given by

$$u(t,k) = \sum_{l\ge1} \psi(t,k,l) c_l = \sum_{l\ge1} \Big(\varphi(t,k+l) - \varphi(t,k-l)\Big) c_l. \tag{A.5}$$

Since I_ν is defined for real values ν, we regard $x \mapsto \psi(t,x,l)$, and consequentially also $x \mapsto u(t,x)$ as a function of the continuous variable $x \in \overline{\mathbb{R}}_+$. In this form, we can prove the following result.

Proposition A.3. *Every solution u to (A.4) with $M_1[c] = \rho$ satisfies*

$$u(t, x\sqrt{t}) \simeq \frac{\rho\, x}{t\sqrt{4\pi}} \exp\left(-\frac{x^2}{4}\right) \qquad \text{as } t \to \infty.$$

Here, $a(t) \simeq b(t)$ as $t \to \infty$ denotes asymptotic equivalence $\lim_{t\to\infty} {}^{a(t)}/_{b(t)} = 1$.

To prove the above result, we first derive another formula for ψ.

Lemma A.4. *For $(t,x,l) \in \overline{\mathbb{R}}_+ \times \overline{\mathbb{R}}_+ \times \mathbb{N}$, it holds*

$$\psi(t,x,l) = \frac{1}{t}\sum_{m=1}^{l}(x-l+2m-1)\varphi(t,x-l+2m-1). \tag{A.6}$$

Proof. The modified Bessel functions of the first kind satisfy the relation

$$I_{x-1}(t) - I_{x+1}(t) = \frac{2x}{t} I_k(t),$$

which by inserting in (A.5) yields

$$\varphi(t,x-1) - \varphi(t,x+1) = \frac{x}{t}\varphi(t,k).$$

Then, we expand $\psi(t,x,l)$ as

$$\begin{aligned}\varphi(t,x-l) - \varphi(t,x+l) &= \sum_{m=1}^{l}\varphi(t,x-l+2(m-1)) - \varphi(t,x-l+2m)\\ &= \frac{1}{t}\sum_{m=1}^{l}(x-l+2m-1)\varphi(t,x-l+2m-1).\end{aligned}$$

□

Next we find the scaling behavior by proving the following asymptotics for the involved Bessel functions.

Lemma A.5. *The modified Bessel functions satisfy for every $x > 0$ the asymptotic*

$$I_{x\sqrt{t}}(t) \simeq \frac{\exp(t)}{\sqrt{2\pi t}}\exp\left(-\frac{x^2}{2}\right) \qquad \text{as } t \to \infty.$$

Consequently, it holds

$$\varphi(t,x\sqrt{t}) \simeq \frac{1}{\sqrt{4\pi t}}\exp\left(-\frac{x^2}{4}\right) \qquad \text{as } t \to \infty. \tag{A.7}$$

Proof. The proof is similar to the result in [42, Theorem 2.13]. We have

$$I_{x\sqrt{t}}(t) \simeq \frac{1}{\pi}\int_0^{\pi}\exp(t\cos(\theta))\cos(x\sqrt{t}\theta)\,\mathrm{d}\theta \qquad \text{as } t \to \infty.$$

By substituting $u = 2\sqrt{t}\sin\left(\frac{\theta}{2}\right)$ in the above integral we obtain

$$I_{x\sqrt{t}}(t) \simeq \frac{\exp(t)}{\pi\sqrt{t}}\int_0^{2\sqrt{t}}\exp\left(-\frac{u^2}{2}\right)\frac{\cos\left(x2\sqrt{t}\sin^{-1}\left(\frac{u}{2\sqrt{t}}\right)\right)}{\sqrt{1-\frac{u^2}{4t^2}}}\,\mathrm{d}u \qquad \text{as } t \to \infty.$$

Clearly the point-wise limit of the integrand is

$$\exp\Big(-\frac{u^2}{2}\Big)\frac{\cos\Big(x2\sqrt{t}\sin^{-1}\Big(\frac{u}{2\sqrt{t}}\Big)\Big)}{\sqrt{1-\frac{u^2}{4t^2}}} \to \exp\Big(-\frac{u^2}{2}\Big)\cos(xu) \qquad \text{as } t\to\infty,$$

which leads by dominated convergence to

$$I_{x\sqrt{t}}(t) \simeq \frac{\exp(t)}{\pi\sqrt{t}}\int_0^\infty \exp\Big(-\frac{u^2}{2}\Big)\cos(xu)\,\mathrm{d}u.$$

By standard Fourier analysis, we obtain in the last step

$$\int_0^\infty \exp\Big(-\frac{u^2}{2}\Big)\cos(xu)\,\mathrm{d}u = \frac{1}{2}\int_{\mathbb{R}}\exp\Big(-\frac{u^2}{2}\Big)\exp(-ixu)\,\mathrm{d}u = \sqrt{\frac{\pi}{2}}\exp\Big(-\frac{x^2}{2}\Big).$$

□

Corollary A.6. *For* $(t,x,l)\in\overline{\mathbb{R}}_+\times\overline{\mathbb{R}}_+\times\mathbb{N}$, *it holds*

$$|\psi(t,x,l)| \lesssim \frac{l}{t}\Big(\frac{x}{\sqrt{t}}+1\Big). \tag{A.8}$$

Proof. We use the identity (A.6) to estimate

$$\begin{aligned} t|\psi(t,x,l)| &\le lx\sup_{y\in[0,\infty)}|\varphi(t,y)| + \sum_{m=1}^{l}|-l+2m-1|\varphi(t,x-l+2m-1) \\ &\lesssim l\frac{x}{\sqrt{t}} + l\sum_{m\in\mathbb{Z}}\varphi(t,m) = l\Big(\frac{x}{\sqrt{t}}+1\Big). \end{aligned}$$

□

With the above bound we can pass to the limit in the representation formula for u.

Proof of Proposition A.3. By the representation formula (A.5), we have

$$t\,u\big(t,x\sqrt{t}\big) = \sum_{l=1}^{\infty} t\,\psi\big(t,x\sqrt{t},l\big)c_l.$$

Next, we note that for every fixed $k\in\mathbb{Z}$ we have

$$\sqrt{t}\,\varphi\big(t,x\sqrt{t}+k\big) \to \frac{1}{\sqrt{4\pi}}\exp\Big(-\frac{x^2}{4}\Big).$$

Indeed, the proof for $k=0$ was done in Lemma A.5 and works with minor modifications in the same way for general k. Thus the formula (A.6) for ψ implies that as $t\to\infty$

$$t\,\psi\big(t,x\sqrt{t},l\big) = \sum_{m=1}^{l} x\sqrt{t}\,\varphi\big(t,x\sqrt{t}-l+2m-1\big) + \mathcal{O}\big(t^{-\frac{1}{2}}\big) \to \frac{l\,x}{\sqrt{4\pi}}\exp\Big(-\frac{x^2}{4}\Big).$$

Furthermore, by Corollary A.6 we have that $|t\psi(t, x\sqrt{t}, l)| \lesssim l\,(x+1)$, so with dominated convergence we conclude

$$\begin{aligned}\lim_{t\to\infty} t\,u\Big(t, x\sqrt{t}\Big) &= \sum_{l=1}^{\infty}\Big(\lim_{t\to\infty} t\,\psi\Big(t, x\sqrt{t}, l\Big)\Big)c_l = \frac{x}{\sqrt{4\pi}}\exp\bigg(-\frac{x^2}{4}\bigg)\sum_{l=1}^{\infty} l\,c_l \\ &= \frac{\rho\,x}{\sqrt{4\pi}}\exp\bigg(-\frac{x^2}{4}\bigg).\end{aligned}$$

$\square$

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1964.

[2] A. Alvino, F. Brock, F. Chiacchio, A. Mercaldo, and M. Posteraro. Some isoperimetric inequalities on $\mathbb{R}^N$ with respect to weights $|x|^\alpha$. *J. Math. Anal. Appl.*, 451(1):280–318, 2017.

[3] J. M. Ball, J. Carr, and O. Penrose. The Becker-Döring cluster equations: Basic properties and asymptotic behaviour of solutions. *Commun. Math. Phys.*, 104(4):657–692, 1986.

[4] F. Barthe and C. Roberto. Sobolev inequalities for probability measures on the real line. *Studia Mathematica*, 159(3):481–497, 2003.

[5] R. Becker and W. Döring. Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. *Ann. der Physik*, 24:719–752, 1935.

[6] J. Beltrán, M. Jara, and C. Landim. A martingale problem for an absorbed diffusion: the nucleation phase of condensing zero range processes. *Probab. Theory Relat. Fields*, 169(3-4):1169–1220, 2017.

[7] E. Ben-Naim and P. L. Krapivsky. Exchange-driven growth. *Phys. Rev. E*, 68(3):031104, 2003.

[8] S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.

[9] M. Bonforte and J. L. Vazquez. Global positivity estimates and Harnack inequalities for the fast diffusion equation. *Journal of Functional Analysis*, 240(2):399–428, 2006.

[10] E. Bouin, J. Dolbeault, and C. Schmeiser. A variational proof of Nash's inequality. *Rend. Lincei-Mat. Appl.*, 31(1):211–223, 2020.

[11] J. Cao, P. Chleboun, and S. Grosskinsky. Dynamics of Condensation in the Totally Asymmetric Inclusion Process. *J. Stat. Phys.*, 155(3):523–543, 2014.

[12] E. A. Carlen and M. Loss. Sharp constant in Nash's inequality. *Int. Math. Res. Not.*, 1993(7):213–215, 1993.

[13] Y.-X. Chau, C. Connaughton, and S. Grosskinsky. Explosive condensation in symmetric mass transport models. *J. Stat. Mech. Theory Exp.*, 2015(11):P11031, 2015.

[14] M. Chen. Analytic proof of dual variational formula for the first eigenvalue in dimension one. *Sci. China Ser. A*, 42(8):805–815, 1999.

[15] M. Doi and S. F. Edwards. *The Theory of Polymer Dynamics*. Comparative Pathobiology - Studies in the Postmodern Theory of Education. Oxford University Press, 1988.

[16] C. Eichenberg. Special solutions to a nonlinear coarsening model with local interactions. *Journal of NonLinear Science*, 29(4):1343–1378, Aug. 2019.

[17] C. Eichenberg and A. Schlichting. Self-similar behavior of the exchange-driven growth model with product kernel. *Communications in Partial Differential Equations*, 46(3):498–546, 2021.

[18] S. Esedoglu and J. B. Greer. Upper bounds on the coarsening rate of discrete, ill-posed nonlinear diffusion equations. *Communications on Pure and Applied Mathematics*, 62(1):57–81, 2009.

[19] S. Esedoglu and D. Slepčev. Refined upper bounds on the coarsening rate of discrete, ill-posed diffusion equations. *Nonlinearity*, 21(12):2759, 2008.

[20] E. Esentürk. Mathematical theory of exchange-driven growth. *Nonlinearity*, 31(7):3460, 2018.

[21] E. Esentürk and J. J. L. Velázquez. Large time behavior of exchange-driven growth. *Discrete & Continuous Dynamical Systems*, 41(2):747–775, 2021.

[22] E. Esenturk. Mathematical theory of exchange-driven growth. *Nonlinearity*, 31(7):3460–3483, 2018.

[23] M. R. Evans and B. Waclaw. Condensation in models with factorized and pair-factorized stationary states. *J. Stat. Mech: Theory Exp.*, 2015(9):P09005, 2015.

[24] S. K. Friedlander. *Smoke, Dust, and Haze: Fundamentals of Aerosol Dynamics*. Topics in chemical engineering. Oxford University Press, 2000.

[25] G. Giacomin, S. Olla, and H. Spohn. Equilibrium fluctuations for ∇_φ interface model. *Ann. Probab.*, 29(3):1138–1172, 07 2001.

[26] K. Glasner and T. Witelski. Collision versus collapse of droplets in coarsening of dewetting thin films. *Physica D: Nonlinear Phenomena*, 209(1):80–104, 2005. Non-linear Dynamics of Thin Films and Fluid Interfaces.

[27] K. B. Glasner and T. P. Witelski. Coarsening dynamics of dewetting films. *Phys. Rev. E*, 67:016302, Jan 2003.

[28] C. Godrèche and J.-M. Drouffe. Coarsening dynamics of zero-range processes. *J. Phys. A Math. Theor.*, 50(1):015005, 2017.

[29] S. Grosskinsky and W. Jatuviriyapornchai. Derivation of mean-field equations for stochastic particle systems. *Stochastic Process. Appl.*, 129(4):1455–1475, 2019.

[30] S. Grosskinsky, F. Redig, and K. Vafayi. Dynamics of condensation in the symmetric inclusion process. *Electronic Journal of Probability*, 18(0), 2013.

[31] S. Grosskinsky, G. M. Schütz, and H. Spohn. Condensation in the zero range process: stationary and dynamical properties. *J. Stat. Phys.*, 113:389, 2003.

[32] E. K. O. Hellén and J. Krug. Coarsening of sand ripples in mass transfer models. *Phys. Rev. E*, 66:011304, Jul 2002.

[33] M. Helmers, B. Niethammer, and J. J. L. Velázquez. Mathematical analysis of a coarsening model with local interactions. *Journal of NonLinear Science*, 26(5):1227–1291, Oct. 2016.

[34] R. Henseler, B. Niethammer, and F. Otto. A reduced model for simulating grain growth. In P. Colli, C. Verdi, and A. Visintin, editors, *Free Boundary Problems*, pages 177–187, Basel, 2004. Birkhäuser Basel.

[35] G. M. Hidy and J. R. Brock. Topics in current aerosol research. *Top. Curr. Aerosol Res.*, 3, 1972.

[36] S. Ispolatov, P. Krapivsky, and S. Redner. Wealth distributions in asset exchange models. *Eur. Phys. J. B*, 2(2):267–276, 1998.

[37] W. Jatuviriyapornchai, P. Chleboun, and S. Grosskinsky. Structure of the condensed phase in the inclusion process. *J. Stat. Phys.*, 2019.

[38] W. Jatuviriyapornchai and S. Grosskinsky. Coarsening dynamics in condensing zero-range processes and size-biased birth death chains. *J. Phys. A Math. Theor.*, 49(18):185005, 2016.

[39] J. Ke and Z. Lin. Kinetics of migration-driven aggregation processes. *Phys. Rev. E*, 66(5):050102, 2002.

[40] R. V. Kohn and F. Otto. Upper bounds on coarsening rates. *Communications in Mathematical Physics*, 229(3):375–395, Sep 2002.

[41] P. L. Krapivsky, S. Redner, and E. Ben-Naim. *A Kinetic View of Statistical Physics*. Cambridge University Press, Cambridge, 2010.

[42] M. Kreh. The Bessel functions. http://www.math.psu.edu/papikian/Kreh.pdf, 2012. Project for the Penn State - Göttingen Summer School on Number Theory.

[43] P. Laurençot. Weak solutions to the Lifshitz-Slyozov-Wagner equation. *Indiana University Mathematics Journal*, 50(3):1319–1346, 2001.

[44] P. Laurençot. The Lifshitz–Slyozov–Wagner equation with conserved total volume. *SIAM Journal on Mathematical Analysis*, 34(2):257–272, 2002.

[45] G. F. Lawler. Notes on the Bessel process. http://www.math.uchicago.edu/~lawler/inprogress, 2018.

[46] F. Leyvraz and S. Redner. Scaling theory for migration-driven aggregate growth. *Phys. Rev. Lett.*, 88(6):068301, 2002.

[47] I. Lifshitz and V. Slyozov. The kinetics of precipitation from supersaturated solid solutions. *Journal of Physics and Chemistry of Solids*, 19(1):35–50, 1961.

[48] J. Nash. Continuity of solutions of parabolic and elliptic equations. *American Journal of Mathematics*, 80(4):931–954, 1958.

[49] B. Niethammer. On the evolution of large clusters in the Becker-Döring model. *J. Nonlinear Sci.*, 13(1):115–155, 2003.

[50] B. Niethammer. Effective theories for Ostwald ripening. https://www.wias-berlin.de:8443/people/koenig/www/FG/niethammer.pdf, 2008.

[51] B. Niethammer and R. Pego. On the initial-value problem in the Lifshitz–Slyozov–Wagner theory of ostwald ripening. *SIAM Journal on Mathematical Analysis*, 31(3):467–485, 2000.

[52] B. Niethammer and R. Pego. Well-posedness for measure transport in a family of nonlocal domain coarsening models. *Indiana University Mathematics Journal*, 54:499–530, 2005.

[53] B. Niethammer and R. L. Pego. Non-self-similar behavior in the LSW theory of ostwald ripening. *Journal of Statistical Physics*, 95(5-6):867–902, June 1999.

[54] A. Schlichting. The exchange-driven growth model: Basic properties and longtime behavior. *J. Nonlinear Sci.*, 2019.

[55] A. Schlichting. Macroscopic limit of the Becker-Döring equation via gradient flows. *ESAIM: COCV*, 25:22, 2019.

[56] M. V. Smoluchowski. Drei Vortrage uber Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen. *Phys. Zeitschrift*, 17:585–599, 1916.

[57] R. van der Meer, J. van der Weele, and D. Lohse. Bifurcation diagram for compartmentalized granular gases. *Physical review E: covering statistical, nonlinear, biological, and soft matter physics*, 63(061304):061304–1–061304–9, 2001.

[58] B. Waclaw and M. R. Evans. Explosive condensation in a mass transport model. *Phys. Rev. Lett.*, 108(7):070601, 2012.

[59] C. Wagner. Theorie der Alterung von Niederschlägen durch Umlösen (Ostwald-Reifung). *Zeitschrift für Elektrochemie, Berichte der Bunsengesellschaft für physikalische Chemie*, 65(7-8):581–591, 1961.

[60] H. T. Yau. Relative entropy and hydrodynamics of Ginzburg-Landau models. *Lett. Math. Phys.*, 22(1):63–80, 1991.