



**HAL**  
open science

## Cutoff for the mean-field zero-range process with bounded monotone rates

Jonathan Hermon, Justin Salez

► **To cite this version:**

Jonathan Hermon, Justin Salez. Cutoff for the mean-field zero-range process with bounded monotone rates. *Annals of Probability*, 2020, 48 (2), pp.742-759. 10.1214/19-AOP1373 . hal-03035272

**HAL Id: hal-03035272**

**<https://hal.science/hal-03035272v1>**

Submitted on 6 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Cutoff for the mean-field zero-range process with bounded monotone rates

Jonathan Hermon, Justin Salez

November 9, 2018

## Abstract

We consider the zero-range process with arbitrary bounded monotone rates on the complete graph, in the regime where the number of sites diverges while the density of particles per site converges. We determine the asymptotics of the mixing time from any initial configuration, and establish the cutoff phenomenon. The intuitive picture is that the system separates into a slowly evolving solid phase and a quickly relaxing liquid phase: as time passes, the solid phase dissolves into the liquid phase, and the mixing time is essentially the time at which the system becomes completely liquid. Our proof uses the path coupling technique of Bubley and Dyer, and the analysis of a suitable hydrodynamic limit. To the best of our knowledge, even the order of magnitude of the mixing time was unknown, except in the special case of constant rates.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Model . . . . .	2
1.2	Main result . . . . .	3
1.3	Proof outline . . . . .	4
<b>2</b>	<b>Fast mixing in the absence of a solid phase</b>	<b>6</b>
2.1	Preliminaries . . . . .	6
2.2	Uniform downward drift . . . . .	7
2.3	Path coupling via tagged particles . . . . .	9
<b>3</b>	<b>Dissolution of the solid phase</b>	<b>12</b>
3.1	Identification of the hydrodynamic limit . . . . .	12
3.2	Proxy for the empirical distribution . . . . .	14
3.3	Tightness and convergence . . . . .	17

# 1 Introduction

## 1.1 Model

Introduced by Spitzer in 1970 [24], the *zero-range process* is a widely studied model of interacting random walks, see, e.g. [18, 19, 9] and the references therein. It describes the evolution of  $m \geq 1$  indistinguishable particles randomly hopping across  $n \geq 1$  sites. The interaction is specified by a function  $r: \{1, 2, \dots\} \rightarrow (0, \infty)$ , where  $r(k)$  indicates the rate at which particles are expelled from a site when  $k$  particles are present on it. We will here focus on the *mean-field* version of the model, where all jump destinations are uniformly distributed. More formally, we consider a continuous-time Markov chain  $\mathbf{X} := (X(t): t \geq 0)$  taking values in the state space

$$\Omega := \left\{ x = (x_1, \dots, x_n) \in \mathbb{Z}_+^n : \sum_{i=1}^n x_i = m \right\}, \quad (1)$$

and whose Markov generator  $\mathcal{L}$  acts on observables  $\varphi: \Omega \rightarrow \mathbb{R}$  as follows:

$$(\mathcal{L}\varphi)(x) = \frac{1}{n} \sum_{1 \leq i, j \leq n} r(x_i) (\varphi(x + \delta_j - \delta_i) - \varphi(x)). \quad (2)$$

Here  $(\delta_i)_{1 \leq i \leq n}$  denotes the canonical basis of  $\mathbb{Z}_+^n$ , and we adopt the convention that  $r(0) = 0$  (no jumps from empty sites). The generator  $\mathcal{L}$  is clearly irreducible, with reversible law

$$\pi(x) \propto \prod_{i=1}^n \prod_{k=1}^{x_i} \frac{1}{r(k)}. \quad (3)$$

The present paper is concerned with the problem of estimating the speed at which the convergence to equilibrium occurs, as quantified by the so-called *mixing times*:

$$t_{\text{MIX}}(x; \varepsilon) := \min \{t \geq 0: \|\mathbb{P}_x(X(t) \in \cdot) - \pi\|_{\text{TV}} \leq \varepsilon\}. \quad (4)$$

In this definition,  $\|\mu - \nu\|_{\text{TV}} = \max_{A \subseteq \Omega} |\mu(A) - \nu(A)|$  denotes the total-variation distance, and the parameters  $x \in \Omega$  and  $\varepsilon \in (0, 1)$  specify the initial state and the desired precision, respectively. Of particular interest is the *worst-case* mixing time, obtained by maximizing over all initial states:

$$t_{\text{MIX}}(\varepsilon) := \max \{t_{\text{MIX}}(x; \varepsilon): x \in \Omega\}. \quad (5)$$

Understanding this fundamental parameter – and in particular, its dependency in the precision  $\varepsilon \in (0, 1)$  – is in general a challenging task, see the books [21, 17] for a comprehensive account. Our current knowledge on the total-variation mixing time of the zero-range process is embarrassingly limited in comparison with the numerous functional-analytic estimates that have been established over the past decades [16, 13, 4, 6, 2, 22, 5, 11, 10, 12]. In fact, to the best of our knowledge, the exact order of magnitude of the mixing time of the zero-range process has only been determined in the very special case where the rate function  $r$  is constant [15, 14, 20, 12].

## 1.2 Main result

The rate function  $r$  will remain fixed throughout the paper, and will only be assumed to be non-decreasing and bounded. Upon re-scaling time by a constant factor if necessary, we take

$$\lim_{k \rightarrow \infty} \uparrow r(k) = 1. \quad (6)$$

Our results will be expressed in terms of a certain generating series associated with  $r$ , namely

$$\Psi(z) := \frac{zR'(z)}{R(z)}, \quad \text{where} \quad R(z) := \sum_{k=0}^{\infty} \frac{z^k}{r(1) \cdots r(k)}. \quad (7)$$

All asymptotic statements will refer to the regime where the density of particles per site stabilizes:

$$n \rightarrow \infty, \quad \frac{m}{n} \rightarrow \rho \in [0, \infty). \quad (8)$$

To lighten the notation, we will keep the dependency upon  $n$  implicit as often as possible. Our main result is the following explicit asymptotics for the worst-case mixing time:

**Theorem 1** (Worst-case mixing time). *For any fixed  $\varepsilon \in (0, 1)$ ,*

$$\frac{t_{\text{MIX}}(\varepsilon)}{n} \xrightarrow{n \rightarrow \infty} \gamma := \int_0^\rho \frac{ds}{1 - \Psi^{-1}(s)}, \quad (9)$$

where  $\Psi^{-1}: [0, \infty) \rightarrow [0, 1)$  denotes the inverse of the (increasing) bijection  $\Psi: [0, 1) \rightarrow [0, \infty)$ .

Although it seems intuitively clear that the worst-case mixing time should be achieved by initially placing all particles on the same site, there does not appear to be any direct justification of this fact. We will thus determine the asymptotics of the mixing time from *every* possible configuration  $x \in \Omega$ , see Corollary 1 below for the detailed result. The notable disappearance of  $\varepsilon$  on the right-hand side of (9) reveals a sharp transition in the convergence to equilibrium of the process, known as a *cutoff* [1, 7]. To the best of our knowledge, the occurrence of this phenomenon for the mean-field zero-range process was only known in the special *rate-one* case, where the function  $r$  is simply constant equal to 1 [20]. This choice trivially fits our setting (6), with

$$R(z) = \frac{1}{1-z}, \quad \Psi(z) = \frac{z}{1-z}, \quad \Psi^{-1}(s) = \frac{s}{1+s}, \quad \gamma = \rho + \frac{\rho^2}{2}. \quad (10)$$

Beyond the obvious complications raised by the non-explicit nature of the rates, Theorem 1 requires new ideas for at least two reasons. First, the crucial spectral gap estimate of Morris [22], on which the whole argument of [20] ultimately relies, is only available in the rate-one case. Second, the stationary distribution (3) is no longer uniform, making the entropy computations from [11, 20] unapplicable. As a result, even the order of magnitude  $t_{\text{MIX}}(\varepsilon) = \Theta(n)$  appears to be new. We circumvent these obstacles by resorting to the powerful *path coupling* method of Bubley and Dyer [3]. This alternative route turns out to be so efficient that the proof of our generalization ends up being significantly shorter than that of the original result [20], without using anything from it.

### 1.3 Proof outline

Intuitively, the system may be viewed as consisting of two regions evolving on different time-scales:

- a slow *solid phase*, consisting of those sites which are occupied by  $\Theta(n)$  particles
- a quick *liquid phase*, formed by those sites which are occupied by  $o(n)$  particles.

The presence of a solid phase is a clear indication that the system is out of equilibrium, since under the stationary law  $\pi$ , the maximum occupancy is easily seen to be  $\Theta(\log n)$  (see, e.g., (30) below). What is less obvious, but true, is that conversely, any completely liquid system reaches equilibrium in negligible time. To make this rigorous, we use the path coupling method of Bubley and Dyer [3]. Note that in the regime (8), there is  $\bar{\rho} < \infty$ , independent of  $n$ , such that

$$\frac{m}{n} \leq \bar{\rho}. \quad (11)$$

By a *dimension-free constant*, we will always mean a number which depends only on  $\bar{\rho}$  and  $r$ .

**Theorem 2** (Fast mixing). *There is a dimension-free constant  $\kappa < \infty$  such that*

$$t_{\text{MIX}}(x; \varepsilon) \leq \kappa \|x\|_{\infty} + (\ln n)^{\kappa}, \quad (12)$$

for every  $x \in \Omega$  and every  $\varepsilon \in (0, 1)$ , provided  $n \geq \kappa/\varepsilon$ .

When combined with the worst-case bound  $\|x\|_{\infty} \leq m$ , this already yields the correct order of magnitude  $t_{\text{MIX}}(\varepsilon) = \mathcal{O}(n)$ , for any fixed  $\varepsilon \in (0, 1)$ . However, the real interest of Theorem 2 lies in the linear dependency in  $\|x\|_{\infty}$ , which implies that the equilibrium is attained in negligible time when the initial configuration  $x = x(n)$  is completely liquid: for any fixed  $\varepsilon \in (0, 1)$ , we have

$$\|x\|_{\infty} = o(n) \implies t_{\text{MIX}}(x; \varepsilon) = o(n). \quad (13)$$

By the Markov property, this reduces our task to that of understanding the time it takes for an arbitrary initial condition  $x \in \Omega$  to become completely liquid. By symmetry, we may assume without loss of generality that the coordinates of  $x$  are arranged in decreasing order:

$$x_1 \geq \dots \geq x_n. \quad (14)$$

In the regime (8), we may further assume (upon passing to a subsequence) that for each  $k \geq 1$ ,

$$\frac{x_k}{n} \xrightarrow[n \rightarrow \infty]{} u_k. \quad (15)$$

Note that the limiting profile  $(u_k)_{k \geq 1}$  must then satisfy  $u_1 \geq u_2 \geq \dots \geq 0$  and  $\sum_{k=1}^{\infty} u_k \leq \rho$ . In this setting, our second step will consist in establishing a deterministic approximation of the form

$$\frac{X_k(nt)}{n} \approx [u_k - f(t)]_+, \quad (16)$$

for fixed  $t \geq 0$  and  $k \geq 1$ , where  $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a smooth increasing function describing the dissolution of the solid phase, and where we have used the notation  $[a]_+ = \max(a, 0)$ .

**Theorem 3** (Hydrodynamic limit). *If  $x = x(n)$  satisfies (14)-(15), then for any fixed  $T \geq 0$ ,*

$$\mathbb{E}_x \left[ \sup_{k \in [n], t \in [0, T]} \left| \frac{X_k(nt)}{n} - [u_k - f(t)]_+ \right| \right] \xrightarrow{n \rightarrow \infty} 0, \quad (17)$$

where the function  $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is characterized by the differential equation

$$f'(t) = 1 - \Psi^{-1} \left( \rho - \sum_{k=1}^{\infty} [u_k - f(t)]_+ \right), \quad f(0) = 0. \quad (18)$$

Moreover, for each  $i \geq 1$ , we have the explicit expression

$$f^{-1}(u_i) = \sum_{k=i}^{\infty} \frac{1}{k} \int_{\rho_k}^{\rho_{k-1}} \frac{ds}{1 - \Psi^{-1}(s)}, \quad (19)$$

where the numbers  $\rho = \rho_0 \geq \rho_1 \geq \dots \geq 0$  are given by  $\rho_k := \rho + ku_{k+1} - \sum_{i=1}^k u_i$ .

The proof relies on a *separation of timescales* argument: the liquid phase relaxes so quickly that, on the relevant time-scale, the solid phase can be considered as inert. Consequently, the liquid phase is permanently maintained in a metastable state resembling the true equilibrium, except that the density is lower because a macroscopic number of particles are “stuck” in the solid phase. This imposes a simple asymptotic relation between the number of particles in the solid phase and the dissolution rate, from which the autonomous equation (18) arises. Note that this limiting description gives access to the dissolution time of the system: for any  $t \geq 0$ , we have

$$\|X(nt)\|_{\infty} = o(n) \iff t \geq f^{-1}(u_1). \quad (20)$$

Combining this with (13), we readily obtain the following complete description of the mixing time.

**Corollary 1** (Mixing time from any state). *In the regime (14)-(15), we have for fixed  $\varepsilon \in (0, 1)$ ,*

$$\frac{t_{\text{MIX}}(x; \varepsilon)}{n} \xrightarrow{n \rightarrow \infty} f^{-1}(u_1) = \sum_{k=1}^{\infty} \frac{1}{k} \int_{\rho_k}^{\rho_{k-1}} \frac{ds}{1 - \Psi^{-1}(s)}. \quad (21)$$

Note that in the degenerate case where  $u_1 = 0$ , the right-hand side of (21) vanishes, in agreement with (13). From this detailed description, the worst-case mixing time can finally be extracted by maximizing the right-hand side of (21) over all possible profiles  $(u_k)_{k \geq 1}$ : we trivially always have

$$\sum_{k=1}^{\infty} \frac{1}{k} \int_{\rho_k}^{\rho_{k-1}} \frac{ds}{1 - \Psi^{-1}(s)} \leq \int_0^{\rho} \frac{ds}{1 - \Psi^{-1}(s)}, \quad (22)$$

and the equality is moreover attained when  $u_1 = \rho$  and  $u_2 = u_3 = \dots = 0$ . The maximizer corresponds to placing all particles on the same site, as anticipated. This clearly establishes Theorem 1, and the remainder of the paper is devoted to the proofs of Theorems 2 and 3, in Sections 2 and 3.

## 2 Fast mixing in the absence of a solid phase

### 2.1 Preliminaries

We start by enumerating a few standard facts that will be used repeatedly in the sequel.

**Graphical construction.** Let  $\Xi$  be a Poisson point process of intensity  $\frac{1}{n} dt \otimes du \otimes \text{Card} \otimes \text{Card}$  on  $[0, \infty) \times [0, 1] \times [n] \times [n]$  (where  $\text{Card}$  denotes the counting measure), and consider the piece-wise constant process  $\mathbf{X} = (X(t): t \geq 0)$  defined by the initial condition  $X(0) = x$  and the following jumps: for each point  $(t, u, i, j) \in \Xi$ ,

$$X(t) := \begin{cases} X(t-) + \delta_j - \delta_i & \text{if } r(X_i(t-)) \geq u \\ X(t-) & \text{otherwise.} \end{cases} \quad (23)$$

Then  $\mathbf{X}$  is a Markov process with generator  $\mathcal{L}$  and initial state  $x$ . We always use this construction.

**Monotony.** Since  $r$  is non-decreasing, the above construction provides a *monotone coupling* of trajectories: if we start from two configurations  $x, y \in \mathbb{Z}_+^n$  satisfying  $x \leq y$  (coordinate-wise), then this property is preserved by the jumps (23), so the resulting processes  $\mathbf{X}, \mathbf{Y}$  satisfy

$$\forall t \geq 0, \quad X(t) \leq Y(t). \quad (24)$$

This classical fact will play an important role in our proof.

**Stochastic regularity** For any  $i \in [n]$  and  $0 \leq s \leq t$ , we have by construction

$$-\Xi([s, t] \times [0, 1] \times \{i\} \times [n]) \leq X_i(t) - X_i(s) \leq \Xi([s, t] \times [0, 1] \times [n] \times \{i\}), \quad (25)$$

and the Poisson random variables  $\Xi(\cdot)$  appearing on both sides have mean  $t - s$ .

**Mean-field jump rate.** For any observable  $\varphi: \Omega \rightarrow \mathbb{R}$ , the process  $\mathbf{M} = (M(t): t \geq 0)$  given by

$$M(t) := \varphi(X(t)) - \varphi(x) - \int_0^t (\mathcal{L}\varphi)(X(s)) ds \quad (26)$$

is a zero-mean martingale, see e.g. [8]. In particular, when  $\varphi(y) = y_i$  ( $i \in [n]$ ), we obtain

$$M(t) = X_i(t) - x + \int_0^t r(X_i(s)) ds - \int_0^t \zeta(s) ds, \quad (27)$$

where the *mean-field jump rate*  $\zeta(t)$  is defined by

$$\zeta(t) := \frac{1}{n} \sum_{j=1}^n r(X_j(t)). \quad (28)$$

Understanding the evolution of  $(\zeta(t): t \geq 0)$  will constitute an important step in the proof.

## 2.2 Uniform downward drift

Our first task consists in showing that the number of particles on a site can not stay large for long.

**Proposition 1** (Uniform downward drift). *There are dimension-free constants  $\theta, \delta > 0$  such that*

$$\mathbb{E}_x \left[ e^{\theta X_i(t)} \right] \leq 2 \left( 1 + e^{\theta(x_i - \delta t)} \right), \quad (29)$$

for all  $x \in \Omega$ ,  $i \in [n]$  and  $t \in \mathbb{R}_+$ . In particular, for any  $a \geq 0$ ,

$$\mathbb{P}_x \left( X_i(t) \geq [x_i - \delta t]_+ + a \right) \leq 4e^{-\theta a}. \quad (30)$$

The intuition behind this result is as follows: if  $X_i(t)$  is large, then by (6) and (27), the drift of  $X_i(t)$  is essentially  $\zeta(t) - 1$ , which is uniformly negative thanks to the following lemma.

**Lemma 1** (Mean-field jump rate). *There is a dimension-free constant  $\varepsilon > 0$  such that*

$$\mathbb{P}_x \left( \zeta(t) \geq 1 - \varepsilon \right) \leq e^{-\varepsilon n},$$

for all  $x \in \Omega$  and all  $t \in [1, \infty)$ , provided  $n \geq 2$ .

*Proof.* Since  $r$  is  $[0, 1]$ -valued with  $r(0) = 0$ , we clearly have

$$\zeta(t) \leq 1 - \frac{1}{n} \sum_{j=1}^n \mathbf{1}_{(X_j(t)=0)}. \quad (31)$$

Now, because of (11), we can find a region  $\mathcal{I} \subseteq [n]$  of size  $|\mathcal{I}| = \lceil n/2 \rceil$  such that  $\max_{i \in \mathcal{I}} x_i \leq 2\bar{\rho}$ . Note that  $|\mathcal{I}^c| = \lfloor n/2 \rfloor \geq n/3$ , since  $n \geq 2$ . For each site  $i \in \mathcal{I}$ , consider the ‘‘good’’ event

$$G_i := \{ \Xi([0, 1] \times [0, 1] \times [n] \times \{i\}) = 0 \} \cap \{ \Xi([0, 1] \times [0, r(1)] \times \{i\} \times \mathcal{I}^c) \geq 2\bar{\rho} \}.$$

The first part forbids any new arrival at  $i$  during the time-interval  $[0, 1]$ , while the second ensures that the  $x_i \leq 2\bar{\rho}$  particles will depart (recall that  $r(k) \geq r(1)$  for  $k \geq 1$ ). Thus,  $G_i \subseteq \{X_i(1) = 0\}$ . Writing  $\mathcal{P}(\lambda; k)$  for the probability that a Poisson variable with mean  $\lambda$  is at least  $k$ , we have

$$\mathbb{P}(G_i) = e^{-1} \mathcal{P} \left( \frac{r(1)\lfloor n/2 \rfloor}{n}; \lfloor 2\bar{\rho} \rfloor \right) \geq e^{-1} \mathcal{P} \left( \frac{r(1)}{3}; \lfloor 2\bar{\rho} \rfloor \right) =: q. \quad (32)$$

Since the events  $(G_i)_{i \in \mathcal{I}}$  are moreover independent, we conclude that the sum  $\sum_{i=1}^n \mathbf{1}_{(X_i(1)=0)}$  stochastically dominates a Binomial random variable with parameters  $\lceil n/2 \rceil$  and  $q$ . In particular, Hoeffding’s inequality implies

$$\mathbb{P} \left( \sum_{i=1}^n \mathbf{1}_{(X_i(1)=0)} \leq \frac{nq}{4} \right) \leq \exp \left( -\frac{nq^2}{4} \right),$$

so that  $\varepsilon = q^2/4$  satisfies the claim for  $t = 1$ . Since the result is uniform in the choice of the initial state  $x \in \Omega$ , the claim automatically propagates to any time  $t \geq 1$  by the Markov property.  $\square$



*Proof of Proposition 1.* For any observable  $\varphi: \Omega \rightarrow \mathbb{R}$ , the formula (26) implies that

$$\frac{d}{dt} \mathbb{E}_x [\varphi(X(t))] = \mathbb{E}_x [(\mathcal{L}\varphi)(X(t))]. \quad (33)$$

We take  $\varphi(y) = e^{\theta y_i}$ , and we bound  $\mathcal{L}\varphi$  in terms of  $\varphi$  to obtain a differential inequality. We have

$$\frac{(\mathcal{L}\varphi)(y)}{\varphi(y)} = (e^\theta - 1) \left\{ \frac{1}{n} \sum_{j \in [n] \setminus \{i\}} (r(y_j) - e^{-\theta} r(y_i)) \right\}.$$

The term  $\{\cdot\}$  is always less than 1, and is even less than  $\lambda := 1 - \varepsilon - e^{-\theta} r(k)$  if  $y \in A \cap B$ , where

$$A := \left\{ \frac{1}{n} \sum_{j=1}^n r(y_j) < 1 - \varepsilon \right\} \quad \text{and} \quad B := \{y_i \geq k\}. \quad (34)$$

The parameters  $\varepsilon \in (0, 1)$  and  $k \in \mathbb{N}$  are arbitrary for now and will be adjusted later. Thus,

$$\begin{aligned} \mathcal{L}\varphi &\leq (e^\theta - 1) (\lambda\varphi + (1 - \lambda)\varphi \mathbf{1}_{(A \cap B)^c}) \\ &\leq (e^\theta - 1) (\lambda\varphi + 2\varphi (\mathbf{1}_{A^c} + \mathbf{1}_{B^c})) \\ &\leq (e^\theta - 1) (\lambda\varphi + 2e^{\theta m} \mathbf{1}_{A^c} + 2e^{\theta k}), \end{aligned}$$

where we have used the simple facts  $\lambda \in (-1, 1)$ ,  $\|\varphi\|_\infty = e^{\theta m}$  and  $\varphi \mathbf{1}_{B^c} \leq e^{\theta k}$ . By (33), we obtain

$$\frac{d}{dt} \mathbb{E}_x [e^{\theta X_i(t)}] \leq (e^\theta - 1) \left\{ \lambda \mathbb{E}_x [e^{\theta X_i(t)}] + 2e^{\theta m} \mathbb{P}_x(\zeta(t) \geq 1 - \varepsilon) + 2e^{\theta k} \right\}. \quad (35)$$

We now choose the dimension-free constants  $\theta, \varepsilon, k$  as follows. We take  $\varepsilon$  as in Lemma 1. Since  $\lambda \rightarrow -\varepsilon$  as  $(k, \theta) \rightarrow (\infty, 0)$ , we may then choose  $k \in \mathbb{N}$  and  $\theta > 0$  so that  $\lambda < 0$ . Upon further reducing  $\theta$  if necessary, we may assume that  $\theta \leq \varepsilon/\bar{p}$ , so that  $\theta m \leq \varepsilon n$ . For  $t \geq 1$ , we then have

$$\frac{d}{dt} \mathbb{E}_x [e^{\theta X_i(t)}] \leq \alpha - \delta \mathbb{E}_x [e^{\theta X_i(t)}], \quad (36)$$

where  $\alpha, \delta > 0$  are dimension-free constants. It is classical that this differential inequality implies

$$\mathbb{E}_x [e^{\theta X_i(t)}] \leq \frac{\alpha}{\delta} + \left( \mathbb{E}_x [e^{\theta X_i(1)}] - \frac{\alpha}{\delta} \right) e^{-\delta(t-1)}, \quad (37)$$

for  $t \geq 1$ . On the other hand, for  $t \in [0, 1]$ , the domination (25) implies  $\mathbb{E}_x [e^{\theta X_i(t)}] \leq e^{\theta x + e^\theta - 1}$ . Combining these two facts, we conclude that there is a dimension-free  $\kappa \in (0, \infty)$  such that

$$\mathbb{E}_x [e^{\theta X_i(t)}] \leq \kappa \left( 1 + e^{\theta x_i - \delta t} \right), \quad (38)$$

for all  $t \geq 0$ . By Jensen's inequality and  $(1 + u)^p \leq 1 + u^p$ , the conclusion still holds if we replace  $(\kappa, \theta, \delta)$  with  $(\kappa^p, \theta p, \delta p)$  for any  $p \in (0, 1)$ . Choosing  $p$  sufficiently small so that  $\kappa^p \leq 2$  and  $\theta p \leq 1$  completes the proof of (29). The claim (30) is then a consequence of Chernov's bound.  $\square$

### 2.3 Path coupling via tagged particles

Our proof of Theorem 2 will rely on the introduction of *tagged particles*. For  $k \in \mathbb{Z}_+$ , define

$$\Delta(k) := r(k+1) - r(k) \geq 0. \quad (39)$$

Let  $\Theta$  be a Poisson process of intensity  $\frac{1}{n}\text{Leb} \otimes \text{Leb} \otimes \text{Card}$  on  $\mathbb{R}_+ \times [0, 1] \times [n]$ , independent of the Poisson process  $\Xi$  used in the graphical construction of  $\mathbf{X}$ , and construct an  $[n]$ -valued process  $\mathbf{I} = (I(t): t \geq 0)$  by setting  $I(0) = i$  and imposing the following jumps: for each  $(t, u, k)$  in  $\Theta$ ,

$$I(t) := \begin{cases} k & \text{if } \Delta(X_{I(t-)}(t-)) \geq u \\ I(t-) & \text{else.} \end{cases} \quad (40)$$

In other words, conditionally on the *background process*  $\mathbf{X}$ , the *tagged particle*  $\mathbf{I}$  performs a time-inhomogeneous random walk starting from  $i$  and jumping from a site  $\ell$  to a uniformly chosen site at the time-varying rate  $\Delta(X_\ell(t))$ . The elementary but crucial observation is that the process

$$(X(t) + \delta_{I(t)}: t \geq 0) \quad (41)$$

is then distributed as a zero-range process starting from  $x + \delta_i$ . Now, if  $j$  is another site, we may introduce a second tagged particle  $\mathbf{J} = (J(t): t \geq 0)$  by setting  $J(0) = j$  and for each  $(t, u, k)$  in  $\Theta$ ,

$$J(t) := \begin{cases} k & \text{if } \Delta(X_{J(t-)}(t-)) \geq u \\ J(t-) & \text{else.} \end{cases} \quad (42)$$

We emphasize that we use the same processes  $\mathbf{X}, \Theta$  to generate  $\mathbf{I}$  and  $\mathbf{J}$ . This produces two coupled zero-range processes  $(X(t) + \delta_{I(t)}: t \geq 0)$  and  $(X(t) + \delta_{J(t)}: t \geq 0)$  starting from  $x + \delta_i$  and  $x + \delta_j$ . We clearly have  $\{I(s) = J(s)\} \subseteq \{I(t) = J(t)\}$  for  $s \leq t$ , and we will estimate the *coalescence time*:

$$\tau := \inf \{t \geq 0: I(t) = J(t)\} = \sup \{t \geq 0: I(t) \neq J(t)\}. \quad (43)$$

**Proposition 2** (Coalescence time). *There is a dimension-free constant  $\kappa < \infty$  such that*

$$\mathbb{P}(\tau > \kappa (\|x\|_\infty \vee (\ln n)^\kappa)) \leq \frac{\kappa}{n^2}.$$

Let us first quickly see how this leads to Theorem 2.

*Proof of Theorem 2.* Set  $P_x^t = \mathbb{P}_x(X(t) \in \cdot)$ . By stationarity of  $\pi$  and convexity of  $d_{\text{TV}}(\cdot, \cdot)$ , we have

$$d_{\text{TV}}(P_x^t, \pi) \leq \sum_{y \in \Omega} \pi(y) d_{\text{TV}}(P_x^t, P_y^t). \quad (44)$$

Call  $x, y \in \Omega$  *adjacent* if they differ by a single jump, i.e.  $y = x + \delta_j - \delta_i$  for some  $1 \leq i \neq j \leq n$ . When this is the case, Proposition 2 (with  $m - 1$  background particles) ensures that

$$t \geq \kappa (\|x\|_\infty \vee \|y\|_\infty \vee (\ln n)^\kappa) \implies d_{\text{TV}}(P_x^t, P_y^t) \leq \frac{\kappa}{n^2}. \quad (45)$$

Now if  $x, y \in \Omega$  are arbitrary, one can always connect them by a path, i.e. a sequence  $(w_0, w_1, \dots, w_k)$  where  $w_0 = x$ ,  $w_k = y$  and  $w_{\ell-1}$  is adjacent to  $w_\ell$  for  $1 \leq \ell \leq k$ . By the triangle inequality, we have

$$d_{\text{TV}}(P_x^t, P_y^t) \leq \sum_{\ell=1}^k d_{\text{TV}}(P_{w_{\ell-1}}^t, P_{w_\ell}^t). \quad (46)$$

Choosing a shortest path further ensures that  $k \leq m$  and  $\max_{1 \leq \ell < k} \|w_\ell\|_\infty \leq \|x\|_\infty \vee \|y\|_\infty$ , so that

$$t \geq \kappa (\|x\|_\infty \vee \|y\|_\infty \vee (\ln n)^\kappa) \implies d_{\text{TV}}(P_x^t, P_y^t) \leq \frac{\kappa m}{n^2}. \quad (47)$$

In particular, if  $t \geq \kappa (\|x\|_\infty \vee (\ln n)^\kappa)$ , then the restriction of the sum in (44) to the index set  $A := \{y \in \Omega: \|y\|_\infty \leq (\ln n)^\kappa\}$  is at most  $\frac{\kappa m}{n^2}$ . On the other hand, the remaining part is at most  $\pi(A^c) \leq e^{-\theta(\ln n)^\kappa}$ , as can be seen by taking the  $t \rightarrow \infty$  limit in (30). In conclusion, for any  $x \in \Omega$ ,

$$t \geq \kappa (\|x\|_\infty \vee (\ln n)^\kappa) \implies d_{\text{TV}}(P_x^t, \pi) \leq \frac{m\kappa}{n^2} + 4e^{-\theta(\ln n)^\kappa}. \quad (48)$$

Upon replacing  $\kappa$  by a larger constant if necessary, we obtain the claim.  $\square$

The remainder of the section is devoted to the proof of Proposition 2. It is clear from (40),(42) that if the two tagged particles manage to jump at the same time, then they immediately coalesce. Note, however, that their jumps may be severely hindered by the background process: in the rate-one case for example, we have  $\Delta(k) = \mathbf{1}_{(k=0)}$ , so that the tagged particles can not jump unless they are alone! Our first step will thus consist in controlling the number of co-occupants of the tagged particles. We will then complement this estimate by showing that, when the tagged particles do not have too many co-occupants, they have a decent chance to coalesce within a short time-interval.

**Lemma 2** (Co-occupants of the tagged particles). *There exist dimension-free constants  $\kappa_1, \kappa_2 < \infty$  such that for  $t = \kappa_1 \|x\|_\infty$  and  $a = \kappa_2 \ln(1 + \|x\|_\infty)$ , we have*

$$\mathbb{P}(X_{I(t)}(t) \vee Y_{J(t)}(t) \leq a) \geq \frac{1}{2}. \quad (49)$$

*Proof.* Since  $I(t), J(t)$  can only move by jumps of the form (40),(42), we necessarily have

$$I(t), J(t) \in \{i, j\} \cup \{k \in [n]: \Theta([0, t] \times [0, 1] \times \{k\}) \geq 1\}. \quad (50)$$

Note that the random set on the right-hand side contains at most  $2 + t$  elements on average. Taking a union bound over all these possibilities, and using the independence of  $\Theta, \mathbf{X}$ , we obtain

$$\mathbb{P}(X_{I(t)}(t) \vee X_{J(t)}(t) > a) \leq (2 + t) \max_{k \in [n]} \mathbb{P}(X_k(t) > a).$$

To make this less than a half, we may choose  $t = \frac{1}{\delta} \|x\|_\infty$  with  $\delta$  as in (30), and  $a = \frac{1}{\delta} \ln(16 + 8t)$ .  $\square$

**Lemma 3** (Quick coalescence). *Setting  $h := \frac{3(x_i \vee x_j) + 1}{r(1)}$ , we have  $\mathbb{P}(\tau \leq h) \geq \frac{e^{-3h}}{8}$ , provided  $n \geq 3$ .*

*Proof.* Write  $h = t + s$  with  $t = 3\frac{x_i \vee x_j}{r(1)}$  and  $s = \frac{1}{r(1)}$ , and note that  $\{\tau \leq h\} \supseteq G_i \cap G_j \cap F$ , where

$$\begin{aligned} G_i &:= \{\Xi([0, t + s] \times [0, 1] \times [n] \times \{i\}) = 0\} \cap \{\Xi([0, t] \times [0, r(1)] \times \{i\} \times [n] \setminus \{i, j\}) \geq x_i\} \\ G_j &:= \{\Xi([0, t + s] \times [0, 1] \times [n] \times \{j\}) = 0\} \cap \{\Xi([0, t] \times [0, r(1)] \times \{j\} \times [n] \setminus \{i, j\}) \geq x_j\} \\ F &:= \{\Theta([0, t] \times [0, 1] \times [n]) = 0\} \cap \{\Theta([t, t + s] \times [0, r(1)] \times [n]) \geq 1\}. \end{aligned}$$

Indeed, the events  $G_i, G_j$  guarantee that  $X_i, X_j$  are zero over the time-interval  $[t, t + s]$ , while  $F$  ensures that the tagged particles will remain in positions  $i, j$  until time  $t$ , and then make an attempt to jump over  $[t, t + s]$ . The first such attempt will be successful for both particles, because the conditions in (40),(42) are met (note that  $\Delta(0) = r(1)$ ). Now,  $F, G_i, G_j$  are independent, with

$$\mathbb{P}(G_k) = e^{-t-s} \mathcal{P}\left(\frac{(n-2)r(1)t}{n}; x_k\right) \quad \text{and} \quad \mathbb{P}(F) = e^{-t} \mathcal{P}(r(1)s; 1),$$

where we recall that  $\mathcal{P}(\lambda; k)$  is the probability that a Poisson variable with mean  $\lambda$  is at least  $k$ . The claim now easily follows from the classical estimate  $\mathcal{P}(\lambda; k) > \frac{1}{2}$ , valid for any  $\lambda \geq k \geq 0$ .  $\square$

**Corollary 2.** *There is a dimension-free constant  $\beta < \infty$  such that  $\mathbb{P}(\tau \leq \beta \|x\|_\infty) \geq (1 + \|x\|_\infty)^{-\beta}$ .*

*Proof.* Set  $t = \kappa_1 \|x\|_\infty$ ,  $a = \kappa_2 \ln(1 + \|x\|_\infty)$ ,  $h = (3a + 1)/r(1)$  with  $\kappa_1, \kappa_2$  as in Lemma 2. Then,

$$\mathbb{P}(\tau \leq t + h) \geq \mathbb{P}(X_{I(t)}(t) \vee X_{J(t)}(t) \leq a) \mathbb{P}(\tau \leq t + h | X_{I(t)}(t) \vee X_{J(t)}(t) \leq a).$$

The first term is at least  $\frac{1}{2}$  and the second at least  $\frac{1}{8}e^{-3h}$ , by Lemma 3 and the Markov property.  $\square$

*Proof of Proposition 2.* Let  $\beta$  be as in the above corollary, and let  $t, a \geq 0$  be parameters to be adjusted later. Consider the increasing sequence of events  $(A_k)_{k \geq 0}$  defined by

$$A_k := \{\tau > t + ka\beta\} \cap \bigcap_{\ell=0}^{k-1} \{\|X(t + \ell a\beta)\|_\infty \leq a\}. \quad (51)$$

By the above corollary and the Markov property, we have  $\mathbb{P}(A_{k+1}|A_k) \leq 1 - (1 + a)^{-\beta}$ . Thus,

$$\mathbb{P}(A_k) \leq \left(1 - (1 + a)^{-\beta}\right)^k \leq e^{-k(1+a)^{-\beta}}. \quad (52)$$

On the other hand, it is clear from the definition of  $A_k$  that

$$\mathbb{P}(\tau > t + ka\beta) \leq \mathbb{P}(A_k) + k \sup_{s \geq t} \mathbb{P}(\|X(s)\|_\infty > a). \quad (53)$$

Recalling (30), we conclude that for  $t = \frac{1}{8}\|x\|_\infty$ ,

$$\mathbb{P}(\tau > t + ka\beta) \leq e^{-k(1+a)^{-\beta}} + 4kne^{-\theta a}. \quad (54)$$

Choosing  $a = \frac{4}{\theta} \ln n$  and  $k = \lfloor (\ln n)^{2+\beta} \rfloor$  ensures that the right-hand side is  $O(\frac{1}{n^2})$ , as desired.  $\square$

### 3 Dissolution of the solid phase

#### 3.1 Identification of the hydrodynamic limit

With the setting of Theorem 3 in mind, we fix a sequence of numbers  $u_1 \geq u_2 \dots \geq 0$  such that

$$\sum_{k=1}^{\infty} u_k \leq \rho. \quad (55)$$

**Proposition 3** (Resolution). *There is a unique measurable functions  $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  satisfying*

$$f(t) = \int_0^t \left\{ 1 - \Psi^{-1} \left( \rho - \sum_{k=1}^{\infty} [u_k - f(s)]_+ \right) \right\} ds, \quad (56)$$

for all  $t \geq 0$ . Moreover,  $f$  is an increasing bijection from  $\mathbb{R}_+$  to  $\mathbb{R}_+$  and for each  $i \geq 1$ ,

$$f^{-1}(u_i) = \sum_{k=i}^{\infty} \frac{1}{k} \int_{\rho_k}^{\rho_{k-1}} \frac{ds}{1 - \Psi^{-1}(s)}, \quad (57)$$

where the numbers  $\rho_0 \geq \rho_1 \geq \dots \geq 0$  are given by  $\rho_k := \rho + ku_{k+1} - \sum_{i=1}^k u_i$ .

*Proof of uniqueness.* Fix  $t > 0$ . Since  $\Psi^{-1}: \mathbb{R}_+ \rightarrow [0, 1)$  is increasing, any solution to (56) satisfies

$$t(1 - \Psi^{-1}(\rho)) \leq f(t) \leq t.$$

Setting  $\kappa(t) := \max \{k \geq 1: u_k > t(1 - \Psi^{-1}(\rho))\}$  (which is finite), we deduce that

$$\sum_{k=1}^{\infty} [u_k - f(t)]_+ = \sum_{k=1}^{\kappa(t)} [u_k - f(t)]_+. \quad (58)$$

In particular, if  $g$  is another solution to (56), we have

$$\left| \sum_{k=1}^{\infty} [u_k - f(t)]_+ - \sum_{k=1}^{\infty} [u_k - g(t)]_+ \right| \leq \kappa(t) |f(t) - g(t)|. \quad (59)$$

Now,  $\Psi^{-1}$  is continuously differentiable and hence  $\alpha$ -Lipschitz on  $[0, \rho]$  for some  $\alpha < \infty$ . Therefore,

$$\left| \Psi^{-1} \left( \rho - \sum_{k=1}^{\infty} [u_k - f(t)]_+ \right) - \Psi^{-1} \left( \rho - \sum_{k=1}^{\infty} [u_k - g(t)]_+ \right) \right| \leq \alpha \kappa(t) |f(t) - g(t)|. \quad (60)$$

Integrating this inequality and recalling (56), we obtain the differential inequality

$$|f(t) - g(t)| \leq \alpha \int_0^t \kappa(s) |f(s) - g(s)| ds. \quad (61)$$

This will force  $f = g$  by Grönwall's Lemma, provided we can show that  $\kappa \in L^1(\mathbb{R}_+)$ . But

$$\int_0^{\infty} \kappa(t) dt = \frac{1}{1 - \Psi^{-1}(\rho)} \sum_{k=1}^{\infty} u_k, \quad (62)$$

by Fubini's Theorem, and the right-hand side is clearly finite, concluding the proof.  $\square$

*Explicit resolution.* Let  $\Phi: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be the function defined by

$$\Phi(t) := \int_0^t \frac{1}{1 - \Psi^{-1}(s)} ds. \quad (63)$$

Note that  $\Phi$  increases continuously from 0 to  $+\infty$ , so that  $\Phi^{-1}: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is well-defined. Now, let  $\rho = \rho_0 \geq \rho_1 \geq \dots \geq 0$  and  $t_1 \geq t_2 \geq \dots \geq 0$  be defined by

$$\rho_k := \rho + k u_{k+1} - \sum_{i=1}^k u_i \quad (64)$$

$$t_k := \sum_{i=k}^{\infty} \frac{\Phi(\rho_{i-1}) - \Phi(\rho_i)}{i}. \quad (65)$$

Finally, define a function  $f$  separately on each  $[t_{k+1}, t_k], k \geq 0$  (with the convention  $t_0 = +\infty$ ) by

$$\forall t \in [t_{k+1}, t_k), \quad f(t) := u_{k+1} + \frac{\Phi^{-1}(\Phi(\rho_k) + k(t - t_{k+1})) - \rho_k}{k}. \quad (66)$$

Note that for any  $k \geq 1$ , we have  $f(t_k) = u_k$ . Moreover, the left-limit of  $f$  at  $t_k$  is

$$\begin{aligned} f(t_k-) &= u_{k+1} + \frac{\Phi^{-1}(\Phi(\rho_k) + k(t_k - t_{k+1})) - \rho_k}{k} \\ &= u_{k+1} + \frac{\rho_{k-1} - \rho_k}{k} \\ &= u_k. \end{aligned}$$

This shows that  $f$  is continuous at each  $t_k, k \geq 1$ . Since  $f$  is clearly continuously increasing on each  $[t_k, t_{k-1})$ , we deduce that  $f$  is continuously increasing on the whole of  $(0, \infty)$ . Moreover,

$$f(0+) = \lim_{k \rightarrow \infty} \downarrow f(t_k) = \lim_{k \rightarrow \infty} \downarrow u_k = 0,$$

so setting  $f(0) := 0$  extends  $f$  into a continuously increasing function on  $\mathbb{R}_+$ . The strict monotony together with the fact that  $f(t_k) = u_k$  shows that for all  $t \in \mathbb{R}_+$  and  $k \geq 1$ ,

$$f(t) < u_k \iff t < t_k. \quad (67)$$

Finally,  $f$  is continuously differentiable on each  $(t_{k+1}, t_k)$  and for  $t \in (t_{k+1}, t_k)$ , we easily compute

$$\begin{aligned} f'(t) &= 1 - \Psi^{-1} \left( \rho + k f(t) - \sum_{i=1}^k u_i \right) \\ &= 1 - \Psi^{-1} \left( \rho - \sum_{i=1}^{\infty} [u_i - f(t)]_+ \right), \end{aligned}$$

where the second equality follows from (67). Thus,  $f$  is a solution to (56), and (57) is clear.  $\square$

### 3.2 Proxy for the empirical distribution

The purpose of this section is to obtain a good approximation for the empirical measure  $\frac{1}{n} \sum_{i=1}^n \delta_{X_i(t)}$ , out of which we will then extract a good approximation for the mean-field jump rate  $\zeta(t)$ . For each  $z \in (0, 1)$ , we define a probability distribution  $\mathbf{q}(z) = (\mathbf{q}(z; k))_{k \geq 0}$  on  $\mathbb{Z}_+$  by the formula

$$\mathbf{q}(z; k) := \frac{1}{R(z)} \frac{z^k}{r(1) \cdots r(k)}. \quad (68)$$

We extend this definition to  $z = 0$  by setting  $\mathbf{q}(0) = \delta_0$ . Note that the mean of  $\mathbf{q}(z)$  is precisely  $\Psi(z)$ . It will be convenient to re-parameterize  $\mathbf{q}$  in terms of its mean by setting  $\bar{\mathbf{q}}(s) := \mathbf{q}(\Psi^{-1}(s))$  for  $s \in (0, \infty)$ . We start by showing that  $\bar{\mathbf{q}}(\rho)$  is the limiting empirical distribution at equilibrium.

**Lemma 4** (Empirical distribution at equilibrium). *In the regime (8), we have for any fixed  $\varepsilon > 0$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \pi \left( \left\{ x \in \Omega : d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{x_i}, \bar{\mathbf{q}}(\rho) \right) \geq \varepsilon \right\} \right) < 0. \quad (69)$$

*Proof.* In the degenerate case  $\rho = 0$ , the claim is trivial since any law on  $\mathbb{Z}_+$  is at total-variation distance at most its mean from the Dirac mass  $\delta_0$ . We henceforth assume that  $\rho > 0$ , and we set  $z = \Psi^{-1}(\rho) \in (0, 1)$ . Consider a random vector  $X := (X_1, \dots, X_n)$  whose coordinates are i.i.d. with law  $\mathbf{q}(z)$ . Then for any  $x = (x_1, \dots, x_n) \in \Omega$ , we have

$$\mathbb{P}(X = x) = \frac{z^m}{(R(z))^n} \prod_{i=1}^n \prod_{k=1}^{x_i} \frac{1}{r(k)}. \quad (70)$$

Thus,  $x \mapsto \mathbb{P}(X = x)$  is proportional to  $\pi$  on  $\Omega$ , and hence for any  $A \subseteq \Omega$ , we have the representation

$$\pi(A) = \frac{\mathbb{P}(X \in A)}{\mathbb{P}(X \in \Omega)}. \quad (71)$$

We now fix  $\varepsilon > 0$  and take

$$A := \left\{ x \in \Omega : d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{x_i}, \mathbf{q}(z) \right) \geq \varepsilon \right\}. \quad (72)$$

Since the coordinates of  $X$  are i.i.d. with law  $\mathbf{q}(z)$ , standard large deviation estimates imply

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(X \in A) < 0; \quad (73)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \in [0, m]) = 0; \quad (74)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \in [m, 2m]) = 0, \quad (75)$$

where  $S_n = X_1 + \dots + X_n$ . On the other hand, since  $\mathbf{q}(z)$  is log-concave, and since this property is preserved under convolutions, the law of  $S_n$  is log-concave and hence unimodal, so for any  $a \in \mathbb{N}$ ,

$$(a+1)\mathbb{P}(S_n = m) \geq \mathbb{P}(S_n \in [m-a, m]) \wedge \mathbb{P}(S_n \in [m, m+a]). \quad (76)$$

Taking  $a = m$  and using (74)-(75), we get  $\frac{1}{n} \log \mathbb{P}(X \in \Omega) \rightarrow 0$ ; (71)-(73) completes the proof.  $\square$

**Remark 1** (Monotony and regularity of  $\rho \mapsto \bar{q}(\rho)$ ). *The monotonicity (24) shows that the stationary law  $\pi$  is stochastically increasing in the number  $m$  of particles. In view of Lemma 4, we deduce that  $\bar{q}(\rho)$  is stochastically increasing in  $\rho$ : if  $\rho \leq \rho'$ , then there is a coupling  $(Z, Z')$  of  $\bar{q}(\rho), \bar{q}(\rho')$  such that  $Z \leq Z'$  almost-surely. Since  $Z, Z'$  are integer-valued, we may then write*

$$d_{\text{TV}}(\bar{q}(\rho), \bar{q}(\rho')) \leq \mathbb{E}[|Z' - Z|] = \mathbb{E}[Z'] - \mathbb{E}[Z] = \rho' - \rho. \quad (77)$$

*In conclusion,  $\rho \mapsto \bar{q}(\rho)$  is increasing and 1-Lipschitz.*

We will show that the approximation  $\frac{1}{n} \sum_{i=1}^n \delta_{X_i} \approx \bar{q}(\rho)$  remains valid out of equilibrium, provided  $\rho$  is replaced by an *effective density*, obtained by ignoring the particles in the solid phase. To formalize this, we assume that the solid phase is initially restricted to some fixed region  $\{1, \dots, L\}$ :

$$\max_{L < i \leq n} x_i = o(n). \quad (78)$$

Note that by (30), this property is preserved by the dynamics in the sense that for any fixed  $t \geq 0$ ,

$$\max_{L < i \leq n} X_i(nt) = o(n), \quad (79)$$

almost-surely (as long as all processes live on the same probability space).

**Proposition 4** (Proxy for the empirical measure). *If  $x = x(n)$  satisfies (78), then for fixed  $t > 0$ ,*

$$\mathbb{E}_x \left[ d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{X_i(nt)}, \bar{q} \left( \frac{1}{n} \sum_{i=L+1}^n X_i(nt) \right) \right) \right] \xrightarrow{n \rightarrow \infty} 0. \quad (80)$$

*Since the rate function  $r$  has mean  $z$  under the law  $\mathfrak{q}(z)$ , we have in particular*

$$\mathbb{E}_x \left[ \left| \zeta(nt) - \Psi^{-1} \left( \frac{m}{n} - \frac{1}{n} \sum_{i=1}^L X_i(nt) \right) \right| \right] \xrightarrow{n \rightarrow \infty} 0. \quad (81)$$

Our proof will consist in comparing the system with one where the solid phase is removed, so that Theorem 2 becomes applicable. We will rely on the following lemma.

**Lemma 5** (Truncation). *Fix  $x \in \mathbb{Z}_+^n$ , and let  $\hat{x}$  be obtained by zeroing the first  $L$  coordinates. Then, the processes  $\mathbf{X}, \hat{\mathbf{X}}$  obtained by applying the graphical construction to  $x, \hat{x}$  satisfy*

$$\mathbb{E} \left[ d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{X_i(t)}, \frac{1}{n} \sum_{i=1}^n \delta_{\hat{X}_i(t)} \right) \right] \leq \frac{L(1+t)}{n}.$$

*Proof.* By (24), we have  $\hat{X}(t) \leq X(t)$  for all  $t \geq 0$ . In particular,

$$\sum_{i=L+1}^n |X_i(t) - \hat{X}_i(t)| = \sum_{i=L+1}^n X_i(t) - \sum_{i=L+1}^n \hat{X}_i(t).$$



Now, observe that the right-hand side equals zero when  $t = 0$ , and that the only jumps of the form (23) that may increment it (by 1 unit each time) are those whose source  $i$  is in  $[L]$ . Consequently,

$$\mathbb{E} \left[ \sum_{i=L+1}^n \left| \widehat{X}_i(t) - X_i(t) \right| \right] \leq \mathbb{E} [\Xi([0, t] \times [0, 1] \times [L] \times [n])] = Lt. \quad (82)$$

On the other hand, by definition of the total-variation distance, we have

$$d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{X_i(t)}, \frac{1}{n} \sum_{i=1}^n \delta_{\widehat{X}_i(t)} \right) \leq \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(X_i(t) \neq \widehat{X}_i(t))}.$$

To conclude, we simply bound  $\mathbf{1}_{(X_i(t) \neq \widehat{X}_i(t))}$  by 1 for  $i \leq L$ , and by  $|X_i(t) - \widehat{X}_i(t)|$  for  $i > L$ .  $\square$

*Proof of proposition 4.* If  $L = 0$ , then  $\|x\| = o(n)$ , so Theorem 2 ensures that for fixed  $t > 0$ ,

$$\max_{A \subseteq \Omega} |\mathbb{P}_x(X(nt) \in A) - \pi(A)| \xrightarrow{n \rightarrow \infty} 0 \quad (83)$$

Since the event  $A$  in Lemma 4 satisfies  $\pi(A) \rightarrow 0$ , we must have  $\mathbb{P}_x(X(nt) \in A) \rightarrow 0$ . Thus,

$$\mathbb{E}_x \left[ d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{X_i(nt)}, \bar{\mathfrak{q}}(\rho) \right) \right] \xrightarrow{n \rightarrow \infty} 0.$$

On the other hand, we have  $d_{\text{TV}}(\bar{\mathfrak{q}}(\frac{m}{n}), \bar{\mathfrak{q}}(\rho)) \rightarrow 0$ , so the case  $L = 0$  is proved. Now, assume that  $x$  satisfies (78) for some  $L \geq 1$ , and let  $\widehat{x}$  be as in Lemma 5. Then  $\|\widehat{x}\|_{\infty} = o(n)$  by construction, so the case  $L = 0$  with  $\widehat{m} := m - (x_1 + \dots + x_L)$  particles instead of  $m$  implies

$$\mathbb{E} \left[ d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{\widehat{X}_i(nt)}, \bar{\mathfrak{q}} \left( \frac{\widehat{m}}{n} \right) \right) \right] \xrightarrow{n \rightarrow \infty} 0.$$

On the other hand, under the coupling of Lemma 5, we have

$$\mathbb{E} \left[ d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{\widehat{X}_i(nt)}, \frac{1}{n} \sum_{i=1}^n \delta_{X_i(nt)} \right) \right] \leq L \left( t + \frac{1}{n} \right).$$

Finally, Remark 1 implies that

$$d_{\text{TV}} \left( \bar{\mathfrak{q}} \left( \frac{\widehat{m}}{n} \right), \bar{\mathfrak{q}} \left( \sum_{i=L+1}^n \frac{X_i(nt)}{n} \right) \right) \leq \frac{1}{n} \left| \sum_{i=1}^L X_i(nt) - \sum_{i=1}^L x_i \right|,$$

and the right-hand side has mean at most  $2Lt$  by (25). By the triangle inequality, we conclude that

$$\limsup_{n \rightarrow \infty} \mathbb{E}_x \left[ d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{X_i(nt)}, \bar{\mathfrak{q}} \left( \frac{1}{n} \sum_{i=L+1}^n X_i(nt) \right) \right) \right] \leq 3Lt.$$

This may seem rather weak compared to what we want to establish. However, by (79), we may apply this result with  $x$  replaced by  $X(ns)$  and then invoke the Markov property to obtain

$$\limsup_{n \rightarrow \infty} \mathbb{E}_x \left[ d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{X_i(ns+nt)}, \bar{q} \left( \frac{1}{n} \sum_{i=L+1}^n X_i(ns+nt) \right) \right) \right] \leq 3Lt,$$

for any  $s \geq 0$  and  $t > 0$ . Replacing  $t$  with  $\varepsilon$  and  $s$  with  $t - \varepsilon$ , we see that for any  $0 < \varepsilon \leq t$ ,

$$\limsup_{n \rightarrow \infty} \mathbb{E}_x \left[ d_{\text{TV}} \left( \frac{1}{n} \sum_{i=1}^n \delta_{X_i(nt)} - \bar{q} \left( \frac{1}{n} \sum_{i=L+1}^n X_i(nt) \right) \right) \right] \leq 3L\varepsilon.$$

Since  $\varepsilon$  can be made arbitrarily small, the result follows.  $\square$

### 3.3 Tightness and convergence

We are now ready to prove Theorem 3, using the classical tightness-uniqueness strategy. Define

$$U_i^n(t) := \frac{X_i(nt)}{n} \quad \text{and} \quad V^n(t) := \int_0^t (1 - \zeta(ns)) \, ds. \quad (84)$$

The fact that  $U_i^n(t) \in [0, \bar{\rho}]$  and the domination (25) suffice to guarantee the tightness of  $(U_i^n)_{n \geq 1}$  in the Skorokhod space  $D(\mathbb{R}_+, \mathbb{R})$ , and the continuity of any weak sub-sequential limit  $U_i^*$ . The same conclusion applies to  $(V^n)_{n \geq 1}$ , because  $\zeta$  is  $[0, 1]$ -valued. Our objective is to show that necessarily,

$$U_i^* = [u_i - f]_+. \quad (85)$$

By diagonal extraction, we may find a sub-sequence along which we have the joint convergence

$$(V^n, U_1^n, U_2^n, \dots) \rightarrow (V^*, U_1^*, U_2^*, \dots), \quad (86)$$

with respect to the product topology. By Skorokhod's Theorem, we may even assume for convenience that the convergence (86) is almost-sure. Our plan is to pass to the limit in the martingale

$$M_i^n(t) := U_i^n(t) - U_i^n(0) + V^n(t) - \int_0^t (1 - r(nU_i^n(s))) \, ds, \quad (87)$$

which is just a rescaled version of (27). Since  $U_i^n$  has jumps of size at most  $\frac{1}{n}$  occurring at rate at most  $2n$ , a classical exponential concentration estimate for martingales (see, e.g., [23]) ensures that

$$M_i^n(t) \xrightarrow[n \rightarrow \infty]{} 0, \quad (88)$$

almost-surely. On the other hand, (30) easily imply that for fixed  $t, h \geq 0$ ,

$$\limsup_{n \rightarrow \infty} \max_{i \in [n]} \{U_i^n(t+h) - [U_i^n(t) - \delta h]_+\} \leq 0, \quad (89)$$

which shows in particular that  $U_i^*(t+h) \leq [U_i^*(t) - \delta h]_+$ . Thus,  $U_i^*$  is non-increasing. Consequently, on the event  $\{U_i^*(t) > 0\}$ , we have  $U_i^*(s) > 0$  for all  $s \leq t$ , and hence  $r(nU_i^n(s)) \rightarrow 1$ , by our assumption (6). Passing to the limit in (87), we conclude that the equality

$$U_i^*(t) = u_i - V^*(t), \quad (90)$$

holds as long as  $U_i^*(t) > 0$ . But both sides of (90) are continuous and non-increasing, so they must reach zero at the same time. Since the left-hand side is non-negative, we conclude that for all  $t \geq 0$ ,

$$U_i^*(t) = [u_i - V^*(t)]_+. \quad (91)$$

Comparing this with (85), we now only have to show that  $V^*$  solves (56), i.e.

$$V^*(t) = \int_0^t \left\{ 1 - \Psi^{-1} \left( \rho - \sum_{i=1}^{\infty} U_i^*(s) \right) \right\} ds. \quad (92)$$

Fix a non-negative integer  $L$ . Taking  $t = 0$  in (89), we deduce that

$$\limsup_{n \rightarrow \infty} \max_{L < i \leq n} U_i^n(h) \leq [u_{L+1} - \delta h]_+. \quad (93)$$

Choosing  $h = \frac{u_{L+1}}{\delta}$  ensures that the right-hand side is zero. Consequently, we may apply (81) with the initial state being  $X(nh)$  and use Markov's property to obtain that for any fixed  $t > h$ ,

$$\mathbb{E}_x \left[ \left| \zeta(nt) - \Psi^{-1} \left( \frac{m}{n} - \sum_{i=1}^L U_i^n(t) \right) \right| \right] \xrightarrow{n \rightarrow \infty} 0. \quad (94)$$

On the other hand, by continuity of  $\Psi^{-1}$ , we have almost-surely,

$$\Psi^{-1} \left( \frac{m}{n} - \sum_{i=1}^L U_i^n(t) \right) \xrightarrow{n \rightarrow \infty} \Psi^{-1} \left( \rho - \sum_{i=1}^L U_i^*(t) \right). \quad (95)$$

Moreover, we can safely replace  $L$  by  $\infty$  on the right-hand side, because (93) ensures that  $U_i^*(t) = 0$  for all  $i > L$ . Combining this with (94), we arrive at

$$\mathbb{E}_x \left[ \left| \zeta(nt) - \Psi^{-1} \left( \rho - \sum_{i=1}^{\infty} U_i^*(t) \right) \right| \right] \xrightarrow{n \rightarrow \infty} 0. \quad (96)$$

This is true for any  $t > h$ , but  $h = \frac{u_{L+1}}{\delta}$  can be made arbitrarily small by choosing  $L$  large, so (96) holds for any  $t > 0$ . Integrating over  $t$ , we easily deduce (92). Finally, note that the convergence

$$U_i^n \xrightarrow{n \rightarrow \infty} [u_i - f]_+ \quad (97)$$

is automatically uniform on compact subsets of  $\mathbb{R}_+$ , because the limit is continuous. It is also uniform in  $i$ , because (93) ensures that  $\max_{L < i \leq n} U_i^n(h)$  can be made arbitrarily small by choosing  $L$  large enough, uniformly in  $n$ . This concludes the proof of Theorem 3.

## References

- [1] David Aldous. Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on probability, XVII*, volume 986 of *Lecture Notes in Math.*, pages 243–297. Springer, Berlin, 1983.
- [2] Anne-Severine Boudou, Pietro Caputo, Paolo Dai Pra, and Gustavo Posta. Spectral gap estimates for interacting particle systems via a Bochner-type identity. *J. Funct. Anal.*, 232(1):222–258, 2006. MR2200172.
- [3] Russ Bubley and Martin Dyer. Path coupling: A technique for proving rapid mixing in markov chains. In *Proceedings of the 38th Annual Symposium on Foundations of Computer Science, FOCS '97*, pages 223–, Washington, DC, USA, 1997. IEEE Computer Society.
- [4] Pietro Caputo. Spectral gap inequalities in product spaces with conservation laws. In *Stochastic analysis on large scale interacting systems*, volume 39 of *Adv. Stud. Pure Math.*, pages 53–88. Math. Soc. Japan, Tokyo, 2004. MR2073330.
- [5] Pietro Caputo and Gustavo Posta. Entropy dissipation estimates in a zero-range dynamics. *Probab. Theory Related Fields*, 139(1-2):65–87, 2007. MR2322692.
- [6] Paolo Dai Pra and Gustavo Posta. Logarithmic Sobolev inequality for zero-range dynamics. *Ann. Probab.*, 33(6):2355–2401, 2005. MR2184099.
- [7] Persi Diaconis. The cutoff phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci. U.S.A.*, 93(4):1659–1664, 1996.
- [8] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [9] M. R. Evans and T. Hanney. Nonequilibrium statistical mechanics of the zero-range process and related models. *J. Phys. A*, 38(19):R195–R240, 2005. MR2145800.
- [10] Max Fathi and Jan Maas. Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.*, 26(3):1774–1806, 2016.
- [11] Benjamin T. Graham. Rate of relaxation for a mean-field zero-range process. *Ann. Appl. Probab.*, 19(2):497–520, 2009. MR2521877.
- [12] Jonathan Hermon and Justin Salez. A version of Aldous’ spectral-gap conjecture for the zero range process. *ArXiv e-prints*, August 2018.

- [13] E. Janvresse, C. Landim, J. Quastel, and H. T. Yau. Relaxation to equilibrium of conservative dynamics. I. Zero-range processes. *Ann. Probab.*, 27(1):325–360, 1999. MR1681098.
- [14] Hubert Lacoin. The cutoff profile for the simple exclusion process on the circle. *Ann. Probab.*, 44(5):3399–3430, 2016. MR3551201.
- [15] Hubert Lacoin. Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion. *Ann. Probab.*, 44(2):1426–1487, 2016. MR3474475.
- [16] C. Landim, S. Sethuraman, and S. Varadhan. Spectral gap for zero-range dynamics. *Ann. Probab.*, 24(4):1871–1902, 1996. MR1415232.
- [17] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017. Second edition of [MR2466937], With a chapter on “Coupling from the past” by James G. Propp and David B. Wilson.
- [18] Thomas M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. MR1717346.
- [19] Thomas M. Liggett. *Interacting particle systems*. Classics in Mathematics. Springer-Verlag, Berlin, 2005. Reprint of the 1985 original.
- [20] Mathieu Merle and Justin Salez. Cutoff for the mean-field zero-range process. *arXiv preprint arXiv:1804.04608*, 2018.
- [21] Ravi Montenegro and Prasad Tetali. Mathematical aspects of mixing times in Markov chains. *Found. Trends Theor. Comput. Sci.*, 1(3):x+121, 2006. MR2341319.
- [22] Ben Morris. Spectral gap for the zero range process with constant rate. *Ann. Probab.*, 34(5):1645–1664, 2006. MR2271475.
- [23] Galen R. Shorack and Jon A. Wellner. *Empirical processes with applications to statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.
- [24] Frank Spitzer. Interaction of Markov processes. *Advances in Math.*, 5:246–290 (1970), 1970. MR0268959.