Finite Size Effects and Metastability in Zero-Range Condensation

Paul Chleboun · Stefan Grosskinsky

Abstract We study zero-range processes which are known to exhibit a condensation transition, where above a critical density a non-zero fraction of all particles accumulates on a single lattice site. This phenomenon has been a subject of recent research interest and is well understood in the thermodynamic limit. The system shows large finite size effects, and we observe a switching between metastable fluid and condensed phases close to the critical point, in contrast to the continuous limiting behaviour of relevant observables. We describe the leading order finite size effects and establish a discontinuity near criticality in a rigorous scaling limit. We also characterise the metastable phases using a current matching argument and an extension of the fluid phase to supercritical densities. This constitutes an interesting example where the thermodynamic limit fails to capture essential parts of the dynamics, which are particularly relevant in applications with moderate system sizes such as traffic flow or granular clustering.

Keywords Zero range process · Condensation · Metastability · Finite size effects · Large deviations

1 Introduction

Zero-range processes are stochastic particle systems with no restriction on the number of particles per site and with jump rates that depend only on the occupation of the departure site. This simple zero-range interaction leads to a product structure of the stationary distributions [1, 35]. These processes have been a focus of recent research interest since they can exhibit a condensation transition. This is the case for space homogeneous jump rates $g(n)$ that decay asymptotically with the number of particles n. A prototypical model with

P. Chleboun (✉) · S. Grosskinsky
Mathematics Institute and Centre for Complexity Science, University of Warwick, Coventry, CV4 7AL, UK
e-mail: P.I.Chleboun@warwick.ac.uk

S. Grosskinsky
e-mail: S.W.Grosskinsky@warwick.ac.uk

$g(n) = 1 + \frac{b}{n^\gamma}$ for $n = 1, 2, \ldots$ has been introduced in [12], where condensation occurs for parameter values $\gamma \in (0, 1)$, $b > 0$ or $\gamma = 1$, $b > 2$. If the particle density ρ exceeds a critical value ρ_c, the system phase separates into a homogeneous background with density ρ_c and a condensate, where the excess particles accumulate on a single randomly located lattice site. This transition has been established on a rigorous level in a series of papers, in the thermodynamic limit [3, 4, 21, 26], as well as on a finite system as the total number of particles diverges [16]. Dynamic aspects of the transition such as equilibration and coarsening [17, 21] and the stationary dynamics of the condensate [18] are well understood heuristically, for the latter first rigorous results have been achieved recently [6, 7]. Findings for the zero-range process could be applied to understand condensation phenomena in a variety of nonequilibrium systems (see [13] and references therein), as well as providing a generic model of domain wall dynamics and a criterion for phase separation using a mapping to one-dimensional exclusion systems [27]. The process continues to be of interest, recent work on variations of the model includes mechanisms leading to more than one condensate [29, 34, 36], or the effects of memory in the dynamics [23].

While most of the results so far consider the thermodynamic limit, finite size effects in the model with jump rates $g(n)$ and $\gamma = 1$ have been investigated in [15, 32] using saddle point methods, and in [2] for a variant of the model with a single defect site. In the condensed phase region, finite systems are found to exhibit a large overshoot of the stationary current above its value in the thermodynamic limit. In this paper we examine this phenomenon in detail for all possible values of the parameter $\gamma \in (0, 1]$. We find the leading order finite size effects that describe the current overshoot by continuing the homogeneous (fluid) phase above the critical density and characterizing the condensed phase by a current matching argument. For $\gamma < 1$, the main focus of this article, finite systems exhibit a metastable switching behaviour between the two phases, which is prevalent in Monte Carlo simulations for a wide range of parameters. To capture this phenomenon we examine the system in a scaling limit and derive a rate function, which exhibits a double well structure, describing the distribution over the bulk density (the bulk density serves as an order parameter to distinguish the fluid and condensed phases). This way we rigorously establish the discontinuous behaviour on the critical scale, even though the bulk density is a continuous function in the thermodynamic limit. This is shown to be in agreement with recent results on the condensation transition at the critical density [3]. Based on the exact scaling limit, we can also predict the lifetime of the metastable phases for large finite systems by a heuristic random walk argument.

In general, finite system size can lead to effective long-range interactions and non-convexity of thermodynamic potentials, as has been observed for various models (see for example [5, 25] and references therein). For the zero-range process it has been shown that for simple size-dependent jump rates metastability effects can be manifested even in the thermodynamic limit [20], and large crossover effects in these systems have already been observed in [12]. The finite-size behaviour of the zero-range process considered here also exhibits the above non-convexity, with the additional feature of a sharp crossover between a putative fluid and condensed phase in the non-convex part. This leads to a metastable switching behaviour which disappears in the thermodynamic limit, intriguingly contradicting the usual expectation of finite systems to behave in a smoother fashion than the limiting prediction. The onset of phase coexistence at criticality is a classical question of general interest in phase separating systems, see for example [9, 10] for the formation of equilibrium droplets in the Ising model which also form suddenly on a critical scale. Rigorous results on metastability regarding the dynamics of the condensate in zero-range processes have also been a subject of recent research interest [6–8], and our work provides a contribution in that

direction and new insight in the mechanisms of condensate dynamics on finite systems. This is explained in more detail in the discussion.

A proper understanding of the metastability phenomenon exhibited by zero-range processes on finite systems is also of particular importance for recent applications with moderate system sizes. Clustering phenomena in granular media can be described by zero-range processes (see [37, 39] and references therein), and metastable switching between homogeneous and condensed states has been observed experimentally [38, 40]. In the spirit of the mapping introduced in [27] the zero-range process with jump rates $g(n)$ has also been applied as a simplified traffic model [28, 30], where condensation corresponds to the occurance of a traffic jam. A key feature of traffic models is the existence of a broad range of densities over which metastability between free flowing and jammed states is observed (see for example [41] and references therein). The relevance of this study in application serves as a motivation, but the aim of the paper is a general understanding of finite size effects and their implications for a generic class of zero-range processes, rather than a detailed analysis of particular cases.

The paper is organized as follows: In Sect. 2 we introduce the model and summarise previous results on stationary distributions and the thermodynamic limit. In Sect. 3 we present relevant observations on finite systems which we later study in detail. In Sect. 4 we give preliminary results and a heuristic description of the finite size effects. This motivates our main results presented in Sect. 5, where we rigorously establish metastability in a suitable scaling limit. In Sect. 6 we connect these results to the lifetimes of the metastable phases, and end with a short discussion in Sect. 7.

2 Stationary Measures and the Thermodynamic Limit

2.1 The Zero Range Process

We consider a one dimensional lattice of L sites $\Lambda_L = \{1, 2, \ldots, L\}$ with periodic boundary conditions. Let $\eta_x \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$ be the number of particles on site $x \in \Lambda_L$. The state of the system is described by $\eta = (\eta_x)_{x \in \Lambda_L}$ belonging to the state space of all particle configurations $\Omega_L = \mathbb{N}_0^{\Lambda_L}$. Particles jump on the lattice at a rate that depends only on the occupation number of the departure site. A particle jumps off site x after an exponential waiting time with rate $g(\eta_x)$ and moves to a target site y according to the probability distribution $p(y - x)$. We assume that p is of finite range, i.e. $p(z) = 0$ if $|z| > R$ for some $R > 0$, normalised and irreducible on Λ_L. The main results of the paper focus on the jump rates $g : \mathbb{N}_0 \to \mathbb{R}_+$ of the form

$$g(n) = \begin{cases} 1 + \frac{b}{n^\gamma} & \text{if } n > 0 \\ 0 & \text{if } n = 0 \end{cases} \tag{1}$$

with $\gamma \in (0, 1)$ and $b > 0$. A discussion on extending these results to the case $\gamma = 1$ and $b \geq 3$ and other lattice geometries can be found in Sect. 7. These jump rates were first introduced by Evans [12] and represent a fairly general class of functions of interest for the condensation transition.

The infinitesimal generator of the process acting on suitable test functions f is given by

$$(\mathcal{L}f)(\eta) = \sum_{x, y \in \Lambda_L} g(\eta_x) p(y - x) \big(f(\eta^{x,y}) - f(\eta) \big), \tag{2}$$

where $\eta_z^{x,y} = \eta_z - \delta(z,x) + \delta(z,y)$ and δ is the Kronecker delta [1, 31]. The process conserves the total number of particles in the system, so Ω_L can be partitioned into invariant subsets $\Omega_{L,N} = \{\eta \in \Omega_L | \sum_{x \in \Lambda_L} \eta_x = N\}$ on which the zero-range process is a finite state irreducible Markov process. The process can also be defined on an infinite lattice under certain constraints, for details see [1, 24].

2.2 Stationary Measures

The following summarises well known results on stationary measures of the zero-range process, for details see [1, 12, 35]. The zero-range process with generator (2) has a family of stationary homogeneous product measures on Ω_L which we refer to as the *grand canonical ensemble*. These measures are parameterized by a fugacity ϕ and are of the form,

$$\nu_\phi^L[\eta] = \prod_{x \in \Lambda_L} \nu_\phi[\eta_x] \quad \text{where } \nu_\phi[n] = \frac{1}{z(\phi)} w(n)\phi^n. \tag{3}$$

These exist for all $\phi \in [0, \phi_c)$ where ϕ_c is the radius of convergence of the (single site) partition function

$$z(\phi) = \sum_{k=0}^{\infty} w(k)\phi^k. \tag{4}$$

The stationary weights w are given by $w(0) = 1$ and

$$w(n) = \prod_{k=1}^{n} g(k)^{-1}, \quad n > 0. \tag{5}$$

In the grand canonical ensemble the expected particle density is a function of ϕ and is given by,

$$R(\phi) := \mathbb{E}_{\nu_\phi}[\eta_1] = \sum_{k=0}^{\infty} k\nu_\phi(k) = \phi\, \partial_\phi \log z(\phi), \tag{6}$$

which is strictly increasing and $R(0) = 0$. The critical density is defined by $\rho_c = \lim_{\phi \nearrow \phi_c} R(\phi) \in (0, \infty]$, and condensation occurs if $\rho_c < \infty$, as explained below.

The expected jump rate off a site is proportional to the average stationary current or, in case the first moment $\sum_z z\, p(z)$ vanishes, to the diffusivity. Therefore for simplicity, in the rest of this paper, *current* will refer to the average jump rate off a site, which is clearly site independent under a stationary distribution. In the grand canonical ensemble the current is simply given by the fugacity ϕ,

$$j_\phi := \mathbb{E}_{\nu_\phi}[g(\eta_x)] = \frac{1}{z(\phi)} \sum_{n=0}^{\infty} g(n)w(n)\phi^n = \phi, \tag{7}$$

which follows directly from the form of the stationary weights $w(n)$, (5).

For fixed L and N the process restricted to $\Omega_{L,N}$ is ergodic, the corresponding unique stationary measures belong to the *canonical ensemble* and are given by

$$\pi_{L,N}[\eta] := \nu_\phi[\eta | S_L(\eta) = N] \quad \text{where } S_L(\eta) = \sum_{x \in \Lambda_L} \eta_x. \tag{8}$$

These measures are independent of ϕ and are given explicitly by

$$\pi_{L,N}[\boldsymbol{\eta}] = \frac{1}{Z(L,N)} \prod_{x \in \Lambda_L} w(\eta_x) \delta\left(\sum_x \eta_x, N\right),$$

where the canonical partition function is

$$Z(L,N) = \sum_{\boldsymbol{\eta} \in \Omega_{L,N}} \prod_{x \in \Lambda_L} w(\eta_x). \tag{9}$$

In the canonical ensemble the form of the stationary weights, (5), imply that the average current is given by a ratio of partition functions,

$$j_{L,N} := \mathbb{E}_{\pi_{L,N}}[g(\eta_x)] = \frac{Z(L,N-1)}{Z(L,N)}. \tag{10}$$

For the jump rates (1) the single site weights asymptotically decay as a stretched exponential

$$w(n) \sim e^{-\frac{b}{1-\gamma} n^{1-\gamma}} \quad \text{for } \gamma \in (0,1). \tag{11}$$

Throughout the paper we use '$\sim$' to mean asymptotically proportional and '$\simeq$' to mean asymptotically equal. It follows that $\phi_c = 1$ and the critical density and the variance are finite,

$$\rho_c := R(1) < \infty, \qquad \sigma_c^2 := \mathbb{E}_{\nu_{\phi_c}}[\eta_x^2] - \rho_c^2 < \infty.$$

R is strictly increasing, so invertible on $[0, \phi_c] = [0, 1]$ and we denote its inverse by

$$\tilde{\Phi}(\rho) = R^{-1}(\rho). \tag{12}$$

In this way we can parameterise the grand canonical measures by densities $\rho \in [0, \rho_c]$.

2.3 Thermodynamic Limit

In this section we summarise known results on the condensation transition in the zero-range process. It has been established in [12] and rigorously in [19, 21] as a continuous phase transition in the thermodynamic limit, as particle number N and lattice size L tend to infinity such that $N/L \to \rho$. It was shown that all finite dimensional marginals of the canonical measure $\pi_{N,L}$ converge to the grand canonical measure with density ρ if $\rho \leq \rho_c$. If $\rho > \rho_c$ then there is no grand canonical measure with density ρ and all finite dimensional marginals of $\pi_{L,N}$ converge to the grand canonical measure with density ρ_c. This result on the equivalence of ensembles holds in terms of weak convergence (for details see [19]) and can be summarised as

$$\pi_{L,N} \to \nu_{\Phi(\rho)} \quad \text{as } L \to \infty \text{ and } N/L \to \rho, \tag{13}$$

where

$$\Phi(\rho) = \begin{cases} \tilde{\Phi}(\rho) & \text{if } \rho < \rho_c, \\ \phi_c & \text{if } \rho \geq \rho_c. \end{cases} \tag{14}$$

This implies that for $\rho > \rho_c$ the excess particles in the system condense on a set of vanishing volume fraction. The result has been strengthened in [3, 4] (with partial results already in [26]), showing that the condensate typically resides on a single lattice site. Denoting the size of the maximum component $M_L(\boldsymbol{\eta}) = \max_{x \in \Lambda_L} \eta_x$ this can be written as

$$\frac{1}{L} M_L \xrightarrow{\pi_{L,N}} \begin{cases} 0 & \text{if } \rho \leq \rho_c, \\ \rho - \rho_c & \text{if } \rho > \rho_c. \end{cases} \tag{15}$$

Here the notation denotes converges in probability with respect to $\pi_{L,N}$,

$$\pi_{L,N}\left[\left|\frac{1}{L} M_L - (\rho - \rho_c)\right| > \varepsilon\right] \to 0 \quad \text{for all } \varepsilon > 0.$$

It will often be useful to express this result also in terms of the behaviour of the bulk of the system. We denote the total number of particles outside of the maximally occupied site by $S_L^{\text{bg}}(\boldsymbol{\eta}) = N - M_L(\boldsymbol{\eta})$. The background density converges as

$$\frac{1}{L-1} S_L^{\text{bg}}(\boldsymbol{\eta}) \xrightarrow{\pi_{L,N}} \begin{cases} \rho & \text{if } \rho \leq \rho_c, \\ \rho_c & \text{if } \rho > \rho_c. \end{cases} \tag{16}$$

A similar result holds for the behaviour of the canonical current in the thermodynamic limit. With (7) the convergence in (13), (14) implies that

$$j_{L,N} \to \Phi(\rho) \quad \text{as } L \to \infty \text{ and } N/L \to \rho. \tag{17}$$

So the current and the background density are both continuous with respect to the total system density in the thermodynamic limit. Also both are strictly increasing up to ρ_c and constant for $\rho > \rho_c$. If $N/L \to \rho \leq \rho_c$ the system is said to be in the *fluid* phase region and if $N/L \to \rho > \rho_c$ the system is in the *condensed* phase region.

The thermodynamic entropy is defined by the Legendre transform

$$s(\rho) = \sup_{\phi \in [0, \phi_c)} (\log z(\phi) - \rho \log \phi)$$
$$= \log z(\Phi(\rho)) - \rho \log \Phi(\rho), \tag{18}$$

where $\Phi(\rho)$ is given by (14). It has been shown in [21] that the canonical partition function converges to the thermodynamic entropy, i.e.

$$\frac{1}{L} \log Z(L, N) \to s(\rho) \quad \text{as } L \to \infty \text{ and } N/L \to \rho. \tag{19}$$

3 Observations on a Finite Systems

In this section we present results obtained from exact numerics and Monte Carlo simulations in the canonical ensemble. We can calculate the canonical current given by (10) by making use of the following recursion relation for the canonical partition functions,

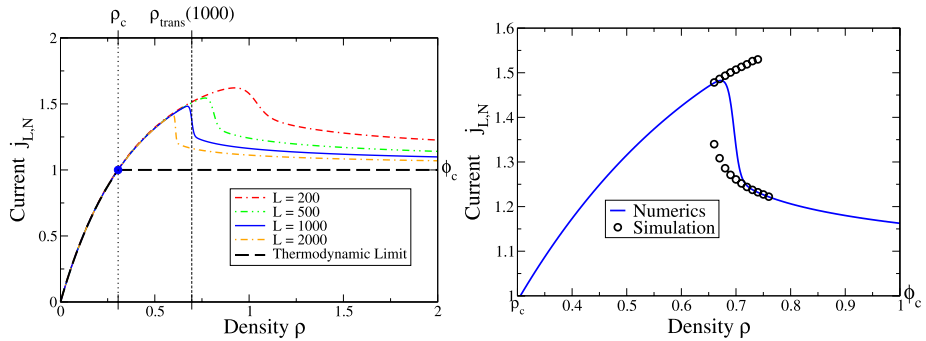$$Z(L, N) = \sum_{k=0}^{N} w(k) Z(L - 1, N - k). \tag{20}$$

Fig. 1 (Color online) Finite size effects and current overshoot for $\gamma = 0.5$ and $b = 4$. *Left*: The canonical current for various system sizes as a function of the density $\rho = N/L$ are plotted. The *dashed black line* shows the thermodynamic current as a function of the system density ρ. *Right*: The overshoot region for $L = 1000$ showing the two distinct currents measured from Monte Carlo simulations for various densities near the maximum current

Similarly we can calculate the canonical distribution of the maximum site occupation M_L. To this end we define the cut-off canonical partition function which counts configurations for which $M_L(\eta) \leq m$,

$$Q(L, N, m) = \sum_{k=0}^{\min\{m,N\}} w(k) Q(L-1, N-k, m). \tag{21}$$

This allows us to calculate

$$\pi_{L,N}[M_L = m] = \frac{Q(L, N, m) - Q(L, N, m-1)}{Z(L, N)}, \tag{22}$$

for all $m \in \mathbb{N}$ (the case $m = 0$ is trivial). We often consider the equivalent formulation using background densities

$$\pi_{L,N}[M_L = m] = \pi_{L,N}\left[\frac{S_L^{\text{bg}}}{L-1} = \rho_{\text{bg}}\right] \quad \text{where } \rho_{\text{bg}} = \frac{N-m}{L-1}.$$

This is more useful for illustrations and is more intuitive, since the background density characterises all but a single lattice site, while the formulation involving M_L is more convenient for computations since it avoids issues with non-integer numbers. Therefore keeping both formulations in parallel is the best option for a concise presentation of our results.

On large finite systems we observe significant finite size effects above the critical density ρ_c, Fig. 1 shows the typical behaviour of the canonical current. Below ρ_c the current is very close to the thermodynamic limit result even on relatively small systems ($L \sim 100$) and the leading order finite size effects can be understood immediately from the proof of the thermodynamic limit result [21]. However the canonical current and background density significantly overshoot their critical values see Fig. 1 and Fig. 2. An overshoot has been observed before for systems containing a single site defect [2] for $\gamma = 1$. For $\gamma < 1$ the current increases monotonically with density $\rho = N/L$, in a way that appears to vary only very slightly with system size, up to some size dependent maximum current at $\rho_{\text{trans}}(L)$. For $\rho < \rho_{\text{trans}}(L)$ the background density in the system is typically very close to ρ (Fig. 2).
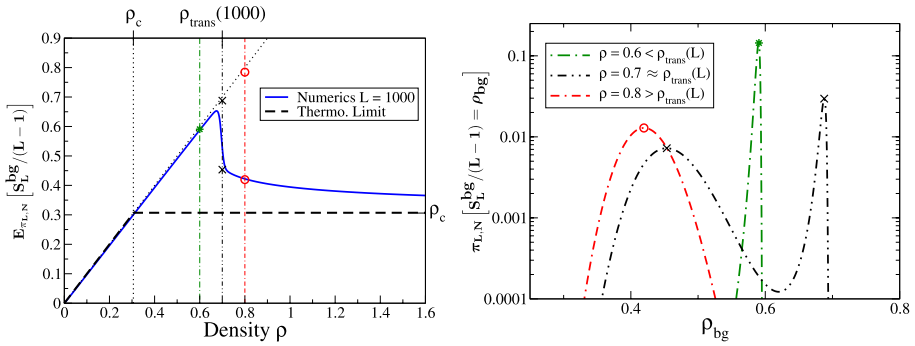
Fig. 2 (Color online) Finite size effects for the background density for $\gamma = 0.5$, $b = 4$ and $L = 1000$. *Left*: The expected value of the background density. The *dashed black line* shows the thermodynamic limit result. *Right*: Distribution of the background density at three system densities shown by corresponding *dash-dotted lines on the left*, calculated exactly using (22). The position of local maxima of the distributions are marked on both plots (∗ for $\rho = 0.6$, × for $\rho = 0.7$, ∘ for $\rho = 0.8$). The high background density maximum at $\rho = 0.8$ occurs with extremely low probability and is off the scale
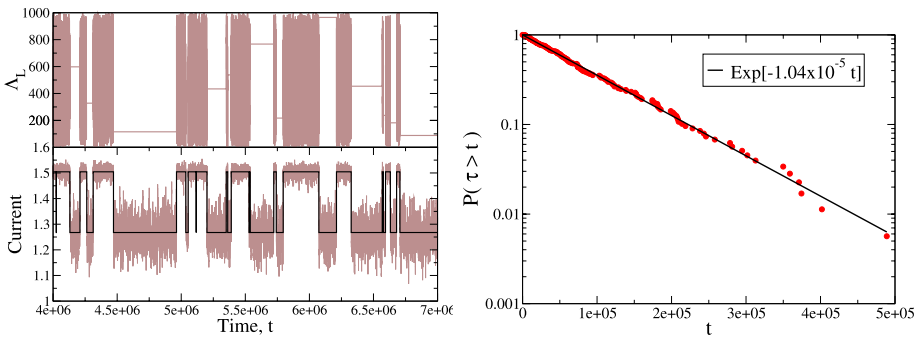

Fig. 3 (Color online) Switching dynamics for $\gamma = 0.5$, $b = 4$ and $L = 1000$ taken at $N/L = \rho_{\text{trans}}(1000) = 0.695$. *Left*: (*Bottom*) Current against time from Monte Carlo simulations, calculated by taking the average number of jumps in the system over small time windows. The *black line* indicates the transition between the two putative phases. (*Top*) The location of the maximum does not change until the system is fluid. *Right*: Cumulative tail of the distribution of waiting times in the putative fluid phase on a log linear scale. The *solid line* shows an exponential fit

For a fixed system size L we associate the region $\rho < \rho_{\text{trans}}(L)$ with a *putative fluid phase*. Close to $\rho = \rho_{\text{trans}}(L)$ there is an abrupt decrease in the current and background density. For $\rho > \rho_{\text{trans}}(L)$ the current and background density are both decreasing in ρ and tend to their respective critical value as $\rho \to \infty$. We associate the region $\rho > \rho_{\text{trans}}(L)$ with a *putative condensed phase* since the increasing density is entirely taken up by a large number of particles condensing on a single lattice site. The size of the effective fluid overshoot increases with increasing b and with decreasing γ, and is already very pronounced at $\gamma = 0.5$, $b = 4$.

For $\rho = N/L$ close to $\rho_{\text{trans}}(L)$ we observe that the canonical distribution over background densities has two maxima of similar magnitude (Fig. 2) and Monte Carlo simulations show that the system switches between the two putative phases. Figure 3 shows the typical behaviour of the current close to $\rho_{\text{trans}}(L)$ as a function of time, the putative phases can be clearly distinguished by the current. For $\rho \approx \rho_{\text{trans}}(L)$ in the condensed phase the location

of the condensate does not change, while its position fluctuates heavily in the fluid phase, supporting the fact that the particles are distributed homogeneously.

The empirical distribution of waiting times in the two putative phases is very close to an exponential (Fig. 3 right). This suggests that the switching process is approximately Markovian over the range of parameters and jump distributions observed, and constitutes a genuine metastability phenomenon. The rate of the switching depends on the parameters b and γ as well as the jump distribution $p(x)$, which will be discussed in Sect. 6 in more detail.

4 Preliminary Results

4.1 Heuristics and Current Matching

As suggested by the form of the current overshoot in Fig. 1, our approach is to approximate the putative fluid phase by extending the grand canonical current above ϕ_c, which we achieve by means of a cut-off grand canonical measure. Since the total number of particles is fixed canonically to N the system can not explore states where any single site contains more than N particles. We therefore expect the distribution of particles in the background under the canonical measure will always be closer (in any reasonable sense) to a grand canonical measure, with some suitably chosen cut-off and fugacitiy chosen to fix the correct density of particles, than it is to the unconditioned distribution. For cut-off $m \in \mathbb{N}$ the *cut-off grand canonical measures* are defined by single site marginals with support on $\{0, 1, \ldots, m\}$,

$$\nu_{\phi,m}[\eta_x = n] := \nu_\phi[\eta_x = n | \eta_x \leq m]$$

$$= \frac{1}{z_m(\phi)} w(n)\phi^n \quad \text{for } n \leq m \tag{23}$$

where the normalisation is given by the finite sum

$$z_m(\phi) = \sum_{k=0}^{m} w(k)\phi^k. \tag{24}$$

These measures are well defined for all $\phi \in (0, \infty)$ and the current is given by

$$\mathbb{E}_{\nu_{\phi,m}}[g(\eta_x)] = \phi\left(1 - \frac{w(m)\phi^m}{z_m(\phi)}\right). \tag{25}$$

Also the average density in the cut-off ensemble

$$R_m(\phi) := \mathbb{E}_{\nu_{\phi,m}}[\eta_x] = \frac{1}{z_m(\phi)} \sum_{k=0}^{m} k w(k) \tag{26}$$

is a strictly increasing function from $[0, \infty)$ onto $[0, \infty)$ and so we denote its inverse

$$\Phi_m(\rho_{\text{bg}}) := R_m^{-1}(\rho_{\text{bg}}). \tag{27}$$

To a first approximation we estimate the current in the putative fluid phase using (25),

$$j_{L,N} \approx \Phi_N(\rho) \tag{28}$$

Fig. 4 (Color online) Current estimates above the critical point $\rho > \rho_c$ and $j_{L,N} > \phi_c$, for $L = 1000$, $\gamma = 0.5$ and $b = 4$. The fluid approximation (28) and the current matching (29) agree well with canonical numerics and the metastable branches from Monte Carlo simulations (cf. Figs. 1 and 3)
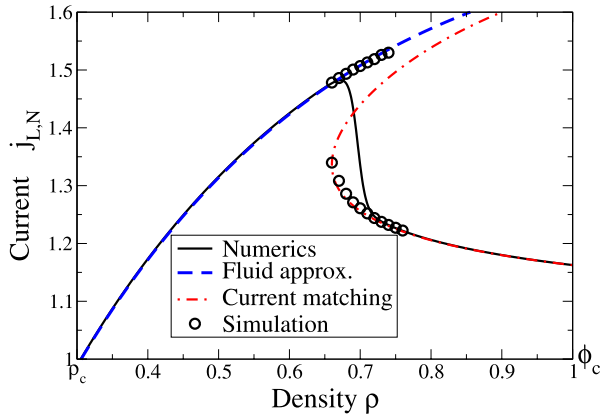
under the assumption that $(1 - w(N)\phi^N/z_N(\phi)) \approx 1$ for values of N under consideration. This approximation is shown in Fig. 4 by the blue dashed line for $\gamma = 0.5$, $b = 4$ and a system size of $L = 1000$. The approximation is extremely close to the canonical current for $\rho < \rho_{\text{trans}}(L)$ and for $\rho \approx \rho_{\text{trans}}(L)$ is in good agreement with the empirically measured fluid currents from Monte Carlo simulations. We demonstrate in the Sect. 5 that this estimate can be improved by choosing the cut-off more carefully.

We approximate the current in the putative condensed phase by a current matching argument, a similar heuristic argument has been given before in [28]. The existence of a stable condensate implies that the average rate of particles exiting the condensate must be equal to the average rate of particles entering it. Conditioned on the occupation of the condensate $M_L = m$, the exit rate is simply $g(m)$ while the entry rate is well approximated by the stationary current in the background. This is assumed to be in a putative fluid phase described by $\nu_{\phi,m}$, which leads to the current matching condition.

$$\Phi_m(\rho_{\text{bg}}) = g(m) \quad \text{where } \rho_{\text{bg}} = \frac{N - m}{L - 1}. \tag{29}$$

The lower branch of solutions to this equation define the condensed current approximation (Fig. 4 red dashed line). These are extremely close to the canonical current for $\rho > \rho_{\text{trans}}(L)$ and for $\rho \approx \rho_{\text{trans}}(L)$ are in good agreement with the empirically measured condensed currents from Monte Carlo simulations.

4.2 Thermodynamic Limit Rate Function

In the thermodynamic limit we expect the existence of a rate function I_ρ that describes the asymptotic probability of observing a background density $\rho_{\text{bg}} \in (0, \rho)$. Precisely, we will consider the limit

$$N, L, m \to \infty \quad \text{such that} \quad N/L \to \rho \quad \text{and} \quad \frac{m}{L} \to \rho - \rho_{\text{bg}}, \tag{30}$$

and expect that

$$\pi_{L,N}[M_L = m] \sim e^{-LI_\rho(\rho_{\text{bg}})}. \tag{31}$$

To derive this we define the following function for finite systems

$$I_{L,N}(m) = -\frac{1}{L}\log \pi_{L,N}[M_L = m],\tag{32}$$

which can be written as

$$\begin{aligned}I_{L,N}(m) = &-\frac{(L-1)}{L}\left(\log z_m(\phi) - \frac{N-m}{L-1}\log\phi\right)\\&-\frac{1}{L}\log w(m) + \frac{1}{L}\log Z(L,N) - \frac{1}{L}\log L\\&-\frac{1}{L}\log \nu_{\phi,m}^{L-1}[S_{L-1} = N - m] + O(e^{-\frac{b}{1-\gamma}m^{1-\gamma}}).\end{aligned}\tag{33}$$

This is derived in Appendix A.1 and is valid for any $\phi \in (0, \infty)$. The first line in (33) resembles the thermodynamic entropy of the background (cf. (18)), the first term on the second line is the contribution due to the maximum occupied site. The second term on the second line is the canonical normalisation and is independent of m. All the other terms will vanish in the limit (30) and we get the following result.

Theorem 1 *In the limit* (30) $I_{L,N}(\rho_{\mathrm{bg}})$ *converges for all* $\rho_{\mathrm{bg}} \in [0, \rho)$ *as*

$$I_\rho(\rho_{\mathrm{bg}}) := \lim_{L \to \infty} I_{L,N}(m) = s(\rho) - s(\rho_{\mathrm{bg}}),\tag{34}$$

where s *is the thermodynamic entropy given by* (18). *So for* $\rho > \rho_c$ *(supercritical case),* $I_\rho(\rho_{\mathrm{bg}}) = 0$ *for each* $\rho_{\mathrm{bg}} \geq \rho_c$ *(see Fig.* 5*).*

Proof The proof follows directly from previous results on the equivalence of ensembles, for details see [21]. It has been shown that

$$\frac{1}{L}\log Z(L,N) \to s(\rho).\tag{35}$$

We use (33) and choose $\phi = \Phi(\rho_{\mathrm{bg}})$ as defined by (14). The cutoff $m \simeq N - \rho_{\mathrm{bg}}(L-1)$ diverges linearly in L for all $\rho_{\mathrm{bg}} \in [0, \rho)$. It follows that for $\phi \leq \phi_c$ we have $z_m(\phi) \to z(\phi)$, and therefore with (18)

$$\log z_m(\Phi(\rho_{\mathrm{bg}})) - \frac{N-m}{L-1}\log \Phi(\rho_{\mathrm{bg}}) \to s(\rho_{\mathrm{bg}}).\tag{36}$$

It follows from the asymptotic behaviour of the single site weights (11) that $\frac{1}{L}\log w(m) \to 0$, and the terms $\frac{1}{L}\log L$ and $O(e^{-\frac{b}{1-\gamma}m^{1-\gamma}})$ also vanish in the limit (30). It remains to show that $\frac{1}{L}\log \nu_{\Phi(\rho_{\mathrm{bg}}),m}^{L-1}[S_{L-1} = N - m] \to 0$. In the case $\rho_{\mathrm{bg}} \leq \rho_c$ we have $\Phi(\rho_{\mathrm{bg}}) \leq \phi_c$ and the first and second moments of $\nu_{\Phi(\rho_{\mathrm{bg}}),m}$ converge to those of $\nu_{\Phi(\rho_{\mathrm{bg}})}$. The local limit theorem for triangular arrays (Theorem 1.2 in [11]) covers the sum, S_L, of independent random variables whose distribution depends on the number of terms L, which is the case here via the cut-off m. Convergence of the first two moments then implies

$$\nu_{\Phi(\rho_{\mathrm{bg}}),m}^{L-1}[S_{L-1} = N - m] \sim \frac{1}{\sqrt{L}},\tag{37}$$

Fig. 5 (Color online)
Thermodynamic limit rate
function for a fixed system
density $\rho = 0.75$ with $\gamma = 0.5$
and $b = 4$. $I_\rho(\rho_{bg})$ is shown by a
solid black line. The *dashed lines*
show $I_{L,N}(m)$ against
$\rho_{bg} = (N - m)/(L - 1)$ with
$N/L = 0.75$, for several values of
L calculated by exact numerics
using (22). Note convergence to
$I_\rho(\rho_{bg})$ (*solid black*) is slow and
initially non-monotonic above ρ_c
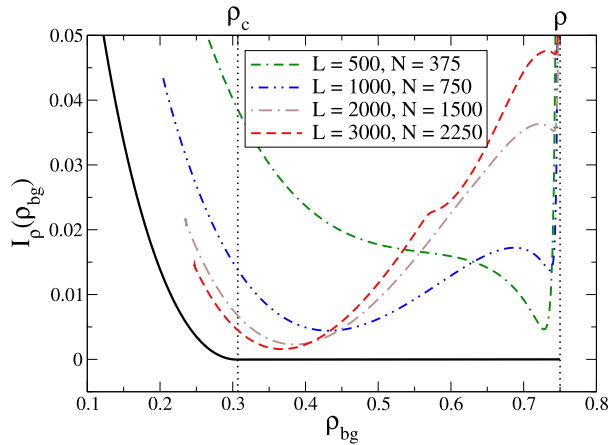
and the result follows immediately. The case $\rho_{bg} > \rho_c$ can be reduced to the previous case following [21], by arranging the excess mass $N - \rho_c(L - 1)$ in the background among a finite number of sites so that each site contains at most m particles. This provides a sub-exponential lower bound for $\nu_{\phi_c,m}^{L-1}[S_{L-1} = N - m]$ which completes the proof. $\square$

Previous results in the thermodynamic limit imply that for $N/L \to \rho > \rho_c$ the background density converges to ρ_c (see Sect. 2.3, (16)). Naturally, the thermodynamic limit result gives no indication of the sharp transition from fluid to condensed putative phases or the metastable switching observed on large finite systems. For $\rho > \rho_c$ the system will appear to be condensed for sufficiently large L, since $\rho_{trans}(L) \searrow \rho_c$ as $L \to \infty$ (see Fig. 1). Therefore the leading order finite size effects of Theorem 1 do not describe the sharp transition or the metastability, as can be seen in Fig. 5. The apparent double well structure of $I_{L,N}$ is less pronounced for higher L, and the location of the global minimum shifts towards ρ_c. For $\rho_{bg} = \rho$, i.e. $m/L \to 0$, the limit of $I_{L,N}$ actually depends on the precise scaling of m, which we do not discuss here (see [4] for more details). To capture the sharp transition and the metastability we will replace the limit (30) by an appropriate scaling limit in Sect. 5.

4.3 Heuristics on Metastability

We can understand the metastability on a heuristic level in terms of a simple current matching argument. $I_{L,N}(m)$ can be calculated efficiently for system sizes $L < 4000$ using exact numerics (see (22)). In Fig. 5 we see a clearly defined double well structure close to $\rho_{trans}(L)$ which accounts for the observed switching behaviour. We may choose $\phi = \Phi_m(\rho_{bg})$ instead of $\phi = \Phi(\rho_{bg})$ in (33) so that the mean under the cut-off grand-canonical distribution is ρ_{bg}. The gradient of $I_{L,N}(m)$ is given by a one step finite difference. If we assume for large L that $\nu_{\Phi_m(\rho_{bg}),m}^{L-1}[S_{L-1}/(L - 1) \asymp \rho_{bg}] \sim \frac{1}{\sqrt{2\pi L \sigma_c^2}}$ (which is valid in the scaling limit under certain conditions according to a local limit for triangular arrays (Theorem 1.2 in [11]), see Theorem 3 for details) and replace $m - 1$ by m, then a straightforward calculation shows

$$I_{L,N}(m - 1) - I_{L,N}(m) \approx \frac{1}{L}\Big(\log \Phi_m(\rho_{bg}) - \log g(m)\Big), \tag{38}$$
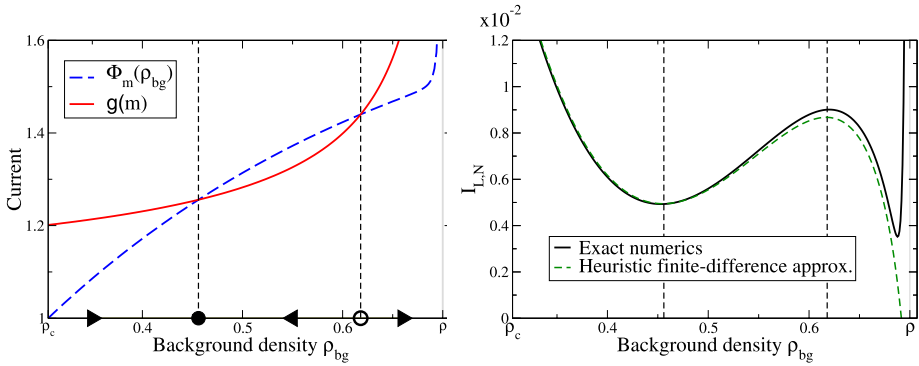
Fig. 6 (Color online) Current matching on a finite system with $L = 1000$ and $\rho = 0.7$ close to $\rho_{\text{trans}}(1000)$ for $\gamma = 0.5$, $b = 4$. *Left*: The current in the background $\Phi_m(\rho_{\text{bg}})$ and current out of the maximum $g(m)$ plotted against $\rho_{\text{bg}} = (N - m)/(L - 1)$. *Right*: Exact numerics of $I_{L,N}$ using (22) shown in *solid black* and approximation (38) in *dashed green*. *Vertical lines* show the correspondence between current matching and critical points of $I_{L,N}$. The first local minimum corresponds to the lower stable branch of the current matching curve in Fig. 4 and the local maximum corresponds to the upper unstable branch

where $\rho_{\text{bg}} = (N - m)/(L - 1)$. The first term on the right hand side follows in direct analogy with the thermodynamic limit for which $\partial_\rho s(\rho) = -\log \Phi(\rho)$ (cf. [19]), and the second term is the finite difference of the maximum site contribution using (5).

This result holds rigorously in scaling limit and even for relatively small systems ($L \approx 1000$) $I_{L,N}$ is well approximated by (38) as is shown in Fig. 6. The approximation breaks down for background densities close to ρ since the cut off m becomes small and so the approximation $m - 1 \approx m$ is no longer valid. By the above argument we expect that solutions to the current matching equation (29) correspond to local maxima or minima of the rate function. Whilst the current out of the maximum occupied site is greater than the current in the background we expect the background density to increase and vice versa. Therefore the first point that the two currents cross is locally stable and so a local minimum of the rate function. The next point they cross is a local maximum by the same argument, and there is another local minimum at a point close to ρ associated with fluid configurations, which the current matching argument predicts to be at the boundary. This is shown in Fig. 6.

5 Rigorous Scaling Limit

In this section we explore the finite size effects by examining the leading order behaviour at the critical scale on which metastability persists. As $L \to \infty$ this scale is given by

$$
\begin{aligned}
N &= \rho_c L + \delta\rho L^{1-\alpha} + o(L^{1-\alpha}) \\
m &= (\delta\rho - \delta\rho_{\text{bg}})L^{1-\alpha} + o(L^{1-\alpha}) \quad \text{for } \rho_{\text{bg}} \in [0, \delta\rho]
\end{aligned}
\tag{39}
$$

for suitable $\alpha \in (0, 1)$. We may equivalently express it as

$$
\begin{aligned}
N/L &= \rho_c + \delta\rho L^{-\alpha} + o(L^{-\alpha}) \\
\frac{N - m}{L - 1} &= \rho_c + \delta\rho_{\text{bg}} L^{-\alpha} + o(L^{-\alpha}).
\end{aligned}
$$

The correct scaling exponent α to capture the canonical overshoot and metastability can be determined heuristically by our previous current matching argument (29). If we assume that close to ρ_c the fluid current equation (28) can be approximated by the first term in the Taylor expansion around ρ_c, then $(\Phi_m(\rho_{\text{bg}}) - 1) \simeq \frac{\delta\rho_{\text{bg}}}{\sigma_c^2} L^{-\alpha}$. The current matching equation (29) then implies

$$\frac{\delta\rho_{\text{bg}}}{\sigma_c^2} L^{-\alpha} \simeq g(m) - 1 \simeq b/m^\gamma = b(\delta\rho - \delta\rho_{\text{bg}})^{-\gamma} L^{-\gamma(1-\alpha)},$$

which leads to $\alpha = \frac{\gamma}{1+\gamma}$. In order to make this argument rigorous, also to find the transition point and describe the metastability, we study the canonical distribution of the background density (or equivalently the maximally occupied site) in the scaling limit.

5.1 The Rate Function

We examine the asymptotic behaviour of $I_{L,N}(m)$ introduced in (32). We will see in Theorem 3 that there exists a unique $1 > \beta > 0$ such that $\lim_{L\to\infty} L^{1-\beta} I_{L,N}(m)$ is finite and non-zero at the critical scale, we define

$$I_{\delta\rho}^{(2)}(\delta\rho_{\text{bg}}) := \lim_{L\to\infty} L^{1-\beta} I_{L,N}(m), \tag{40}$$

where N and m are given by (39). Since $I_\rho(\rho_{\text{bg}}) = 0$ for $\rho_{\text{bg}} \geq \rho_c$ (see Sect. 4.2) this definition implies

$$\pi_{L,N}[M_L = m] \sim e^{-L^\beta I_{\delta\rho}^{(2)}(\delta\rho_{\text{bg}})} \tag{41}$$

as $L \to \infty$, we recall that if $M_L = m$ the background density is $\frac{N-m}{L-1}$ and $\delta\rho$ and $\delta\rho_{\text{bg}}$ are given by (39). We begin by describing the scaling limit for the current in the fluid bulk and out of the maximally occupied site.

Theorem 2 (Current scaling limits) *For the scaling limit* (39), *with* $\alpha = \frac{\gamma}{1+\gamma}$, *the current out of the most occupied site and the average current in the fluid background are asymptotically given by*

$$g(m) = 1 + \frac{b}{(\delta\rho - \delta\rho_{\text{bg}})^\gamma} L^{-\alpha} + o(L^{-\alpha}),$$

$$\Phi_m\left(\frac{N-m}{L-1}\right) = 1 + \frac{1}{\sigma_c^2}\delta\rho_{\text{bg}} L^{-\alpha} + o(L^{-\alpha})$$

provided $\frac{1}{\sigma_c^2}\delta\rho_{\text{bg}} < \frac{b}{(1-\gamma)(\delta\rho-\delta\rho_{\text{bg}})^\gamma}$.

Proof The result for the current out of the condensate $g(m)$ is immediate since $m = (\delta\rho - \delta\rho_{\text{bg}})L^{1-\alpha} + o(L^{1-\alpha})$, $g(m) = 1 + bm^{-\gamma}$ and $(1-\alpha)\gamma = \alpha$.

The proof of the fluid result follows from a Taylor expansion of the truncated density function $R_m(\phi)$ introduced in (26), details on the expansion are contained in Lemmas 1 and 2 in the Appendix. For convenience we write $\phi = e^\mu$, then

$$R_m(e^{\mu L^{-\alpha}}) = \rho_c + \sigma_c^2 \mu L^{-\alpha} + o(L^{-\alpha})$$

for all $\mu \in [0, \frac{b}{(1-\gamma)(\delta\rho-\delta\rho_{\text{bg}})^\gamma})$. The result follows directly since

$$e^{\mu L^{-\alpha}} = \Phi_m \left(R_m(e^{\mu L^{-\alpha}}) \right)$$

$$\Rightarrow 1 + \mu L^{-\alpha} + o(L^{-\alpha}) = \Phi_m \left(\rho_c + \sigma_c^2 \mu L^{-\alpha} + o(L^{-\alpha}) \right)$$

and choosing μ so that $\sigma_c^2 \mu = \delta\rho_{\text{bg}}$.   □

Although the proofs in this section are restricted to $\frac{1}{\sigma_c^2}\delta\rho_{\text{bg}} < \frac{b}{(1-\gamma)(\delta\rho-\delta\rho_{\text{bg}})}$ this includes the point of intersection of the two currents (cf. Fig. 6) and we discuss a possible extension to the whole region $\delta\rho_{\text{bg}} \in (0, \delta\rho)$ at the end of this Subsection. Theorem 2 states that in the scaling limit the average current in the fluid background and the jump rate out of the condensate converge to the critical current $\phi_c = 1$ on the same scale as the density N/L converges to ρ_c. So for N and m at the critical scale, given by (39), $(1 - \Phi_m(\rho_{\text{bg}}))L^\alpha$ and $(1 - g(m))L^\alpha$ converge to unique functions of the scaled variables $\delta\rho$ and $\delta\rho_{\text{bg}}$. We now show that in the scaling limit the rate function is exactly the integral of the difference in the two limiting currents (up to normalisation).

Theorem 3 (Scaling limit rate function) *In the scaling limit (39), with $\alpha = \frac{\gamma}{1+\gamma}$, the rescaled rate function (40) with $\beta = \frac{1-\gamma}{1+\gamma}$ converges as*

$$I_{\delta\rho}^{(2)}(\delta\rho_{\text{bg}}) = \lim_{L\to\infty} L^{1-\beta} I_{L,N}(m)$$

$$= \frac{\delta\rho_{\text{bg}}^2}{2\sigma_c^2} + \frac{b}{1-\gamma}(\delta\rho - \delta\rho_{\text{bg}})^{1-\gamma} - \inf_{r\in(0,\delta\rho)} \left\{ \frac{r^2}{2\sigma_c^2} + \frac{b}{1-\gamma}(\delta\rho - r)^{1-\gamma} \right\}$$

provided $\frac{1}{\sigma_c^2}\delta\rho_{\text{bg}} < \frac{b}{(1-\gamma)(\delta\rho-\delta\rho_{\text{bg}})^\gamma}$.

Proof Firstly consider the unnormalised measure

$$\mu_{L,N} = \frac{Z(L,N)}{z^L(1)} \pi_{L,N}, \tag{42}$$

for which it follows from definitions in Sect. 2.2 and (32)

$$\frac{1}{L^\beta} \log \mu_{L,N}[M_L = m] = -L^{1-\beta} I_{L,N}(m) + \frac{1}{L^\beta} \log Z(L,N) - L^{1-\beta} \log z(1). \tag{43}$$

Re-writing $I_{L,N}$ using (33) we find

$$-\frac{1}{L^\beta} \log \mu_{L,N}[M_L = m] = -\frac{(L-1)}{L^\beta} \left(\log z_m(\Phi_m(\rho_{\text{bg}}^L)) - \rho_{\text{bg}}^L \log \Phi_m(\rho_{\text{bg}}^L) \right)$$

$$+ L^{1-\beta} \log z(1) - \frac{1}{L^\beta} \log w(m) - \frac{1}{L^\beta} \log L$$

$$- \frac{1}{L^\beta} \log \nu_{\Phi_m(\rho_{\text{bg}}^L),m}^{L-1} [S_{L-1} = N - m] - o(1), \tag{44}$$

where we used the shorthand $\rho_{\text{bg}}^L = (N - m)/(L - 1)$. By applying Lemma 2, given in the appendix, the variance under the cut-off distribution $\nu_{\Phi_m(\rho_{\text{bg}}^L),m}$ converges to the variance of

the critical measure for $\frac{1}{\sigma_c^2}\delta\rho_{\text{bg}} < \frac{b}{(1-\gamma)(\delta\rho-\delta\rho_{\text{bg}})^\gamma}$. By the local limit theorem for triangular arrays (Theorem 1.2 in [11]) we have

$$\nu^{L-1}_{\Phi_m(\rho^L_{\text{bg}}),m}\left[S_{L-1} = N - m\right] \sim \frac{1}{\sqrt{2\pi L\sigma_c^2}} \tag{45}$$

since $\rho^L_{\text{bg}} \to \rho_c$ and $\mathbb{E}_{\nu_{\Phi_m(\rho^L_{\text{bg}}),m}}[\eta_x] \to \rho_c$ as $L \to \infty$ with the scaling (39). So the contribution of this term to (44) vanishes. Also since $\beta > 0$ we have $L^{-\beta}\log L \to 0$ as $L \to \infty$.

The condensate contribution to (44) is given by

$$\frac{1}{L^\beta}\log w(m) \simeq -\frac{1}{L^\beta}\frac{b}{1-\gamma}(\delta\rho - \delta\rho_{\text{bg}})^{1-\gamma}L^{(1-\alpha)(1-\gamma)}$$

$$= -\frac{b}{1-\gamma}(\delta\rho - \delta\rho_{\text{bg}})^{1-\gamma} \tag{46}$$

where we have used the asymptotic behaviour of the single site weights (see (11)) and $(1-\alpha)(1-\gamma) = \beta$.

By applying Theorem 2 to $\Phi_m(\rho^L_{\text{bg}})$ and taking Taylor expansion of z_m according to Lemma 2 (Appendix),

$$L^{1-\beta}\left(\rho_c + \delta\rho_{\text{bg}}L^{-\alpha}\right)\log\Phi_m(\rho_c + \delta\rho_{\text{bg}}L^{-\alpha}) = \frac{\delta\rho_{\text{bg}}}{\sigma_c^2}L^\alpha\rho_c + \frac{\delta\rho_{\text{bg}}^2}{\sigma_c^2} + o(1) \tag{47}$$

and

$$L^{1-\beta}\log z_m\left(\Phi_m(\rho^L_{\text{bg}})\right)$$

$$= L^{1-\beta}\log\left(z(1) + \frac{\delta\rho_{\text{bg}}}{\sigma_c^2}L^{-\alpha}z'(1) + \frac{\delta\rho_{\text{bg}}^2}{2\sigma_c^4}L^{-2\alpha}z''(1) + o(L^{-2\alpha})\right) \tag{48}$$

$$= L^{1-\beta}\log z(1) + \frac{\delta\rho_{\text{bg}}}{\sigma_c^2}L^\alpha\rho_c + \frac{\delta\rho_{\text{bg}}^2}{2\sigma_c^2} + o(1). \tag{49}$$

since $(1-\beta) = 2\alpha$. Combining (45) to (48) in (44) implies,

$$-\frac{1}{L^\beta}\log\mu_{L,N}\left[M_L = m\right] = \frac{\delta\rho_{\text{bg}}^2}{2\sigma_c^2} + \frac{b}{1-\gamma}(\delta\rho - \delta\rho_{\text{bg}})^{1-\gamma} + o(1) \tag{50}$$

in the scaling limit (39) as $L \to \infty$, provided $\frac{1}{\sigma_c^2}\delta\rho_{\text{bg}} < \frac{b}{(1-\gamma)(\delta\rho-\delta\rho_{\text{bg}})^\gamma}$.

By definition

$$Z(L, N) = z^L(1)\sum_{m=1}^{N}\mu_{L,N}\left[M_L = m\right].$$

So we may bound $Z(L, N)$ as follows,

$$z^L(1)\max_m\{\mu_{L,N}\left[M_L = m\right]\} \leq Z(L, N) \leq Nz^L(1)\max_m\{\mu_{L,N}\left[M_L = m\right]\}.$$

Applying (50) the asymptotic behaviour of $L^{-\beta} \log Z(L, N)$ in the scaling limit is given by

$$L^{-\beta} \log Z(L, N) = L^{1-\beta} \log z(1) - \inf_{r \in (0, \delta \rho)} \left\{ \frac{r^2}{2\sigma_c^2} + \frac{b}{1 - \gamma} (\delta \rho - r)^{1-\gamma} \right\} + o(1), \quad (51)$$

provided that the global minimum of $-\frac{1}{L^\beta} \log \mu_{L,N}[M_L = m]$ is attained for $\frac{1}{\sigma_c^2} \delta \rho_{\text{bg}} < \frac{b}{(1-\gamma)(\delta \rho - \delta \rho_{\text{bg}})^\gamma}$ in the scaling limit. This follows from results in [3]. Combining (50) and (51) the result follows from definitions (33) and (42)

$$L^{1-\beta} I_{L,N}(\rho_{\text{bg}}) = \frac{\delta \rho_{\text{bg}}^2}{2\sigma_c^2} + \frac{b}{1 - \gamma} (\delta \rho - \delta \rho_{\text{bg}})^{1-\gamma}$$

$$- \inf_{r \in (0, \delta \rho)} \left\{ \frac{r^2}{2\sigma_c^2} + \frac{b}{1 - \gamma} (\delta \rho - r)^{1-\gamma} \right\} + o(1). \qquad \square$$

The result may break down for $\frac{1}{\sigma_c^2} \delta \rho_{\text{bg}} > \frac{b}{(1-\gamma)(\delta \rho - \delta \rho_{\text{bg}})^\gamma}$ due to the large probability of a second sub-condensate forming on the same scale as the condensate. The conditions under which this occurs could be found by conditioning on a maximum site occupation and using the same methods as for Theorem 3 to find the distribution over the second highest occupied site. By using this method and iterating when necessary, the results here could be extended to the entire region $\delta \rho_{\text{bg}} \in (0, \delta \rho)$. However our results already cover the relevant critical points of the rate function, the fluid minimum, condensed minimum and local maximum that constitutes a potential barrier.

5.2 Current Matching and Overshoot

Theorem 3 together with Theorem 2 imply that in the scaling limit given by (39) the derivative of the scaling rate function $I_{\delta \rho}^{(2)}$ is given by the difference in the current out of the condensate and the average current in the background. This is made precise in the following result, which is a rigorous version of the current matching argument of Sect. 4.

Corollary 1 (Current matching)

$$\partial_{\delta \rho_{\text{bg}}} I_{\delta \rho}^{(2)} (\delta \rho_{\text{bg}}) = \frac{\delta \rho_{\text{bg}}}{\sigma_c^2} - \frac{b}{(\delta \rho - \delta \rho_{\text{bg}})^\gamma} \tag{52}$$

and the two terms on the right-hand side are exactly the limiting current curves from Theorem 2 above the critical point.

This implies that

$$\partial_{\delta \rho_{\text{bg}}} I_{\delta \rho}^{(2)} (\delta \rho_{\text{bg}}) \to -\infty \quad \text{as } \delta \rho_{\text{bg}} \to \delta \rho, \tag{53}$$

so that $I_{\delta \rho}^{(2)}$ always exhibits a boundary minimum at $\delta \rho_{\text{bg}} = \delta \rho$.

Depending on the value of $\delta \rho$ the rescaled rate function has one of three qualitative forms characterised by the number of its extreme points, or equivalently the number of roots of (52), and the position of the global minimum. With the threshold

$$c_0(\gamma, b) = \frac{1 + \gamma}{\gamma} \left(\sigma_c^2 \gamma b \right)^{1/(1+\gamma)}, \tag{54}$$
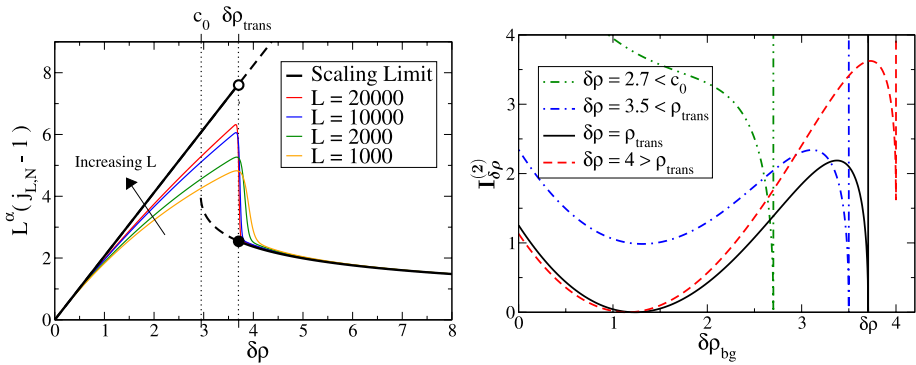
Fig. 7 (Color online) Scaling limit for $\gamma = 0.5$ and $b = 4$. *Left*: The *thick black line* shows the scaling limit result including metastable branches (*dashed*). Thin *coloured lines* show the rescaled canonical currents from exact numerics, slowly approaching the scaling limit as L increases. The transition point is given by $\delta\rho_{\text{trans}}$. *Right*: $I_{\delta\rho}^{(2)}$ for various values of $\delta\rho$. *Solid black* corresponds to $\delta\rho = \delta\rho_{\text{trans}}$ where the depth of the two local minima are equal. The *green dashed-dot-dot curve* shows $I_{\delta\rho}^{(2)}$ for $\delta\rho < c_0$, *blue dot-dashed curve* corresponds to $c_0 < \delta\rho < \delta\rho_{\text{trans}}$ and *red dashed* $\delta\rho > \delta\rho_{\text{trans}}$

Equation (52) has no real roots for $\delta\rho < c_0$, exactly one for $\delta\rho = c_0$, and two for $\delta\rho > c_0$. The latter correspond to a local minimum at $\delta\rho_{\text{bg}} = r_0 \in (0, \delta\rho)$ and a local maximum at $\delta\rho_{\text{bg}} = r^*$, with $r_0 < r^*$ (cf. Fig. 7 on the right). As usual, minima of $I_{\delta\rho}^{(2)}$ correspond to metastable phases, and the depth of the local minimum at r_0 as compared to the one at the boundary determines which of the phases is stable (i.e. corresponds to the global minimum). So c_0 marks the threshold above which (for $\delta\rho > c_0$) there exists a metastable condensed phase. Stability of the phases changes when both minima have the same depth, the density where this is the case is given by $\rho_{\text{trans}}(L)$ introduced in Sect. 3. In the scaling limit (39) this behaves as

$$\rho_{\text{trans}}(L) = \rho_c + \delta\rho_{\text{trans}} L^{-\alpha} + o(L^{-\alpha}),$$

and from Theorem 3 and Corollary 1 we get the explicit expression

$$\delta\rho_{\text{trans}} = \sigma_c^{\frac{2}{1+\gamma}} (1 + \gamma)(2\gamma)^{-\frac{\gamma}{1+\gamma}} \left(\frac{b}{1 - \gamma} \right)^{\frac{1}{1+\gamma}}. \tag{55}$$

Note that both c_0 and $\delta\rho_{\text{trans}}$ are increasing with γ and b. Their ratio simplifies to

$$\delta\rho_{\text{trans}}/c_0 = \frac{1}{2}\left(\frac{2}{1 - \gamma} \right)^{1/(1+\gamma)}, \tag{56}$$

which also increases monotonically with γ. In particular, $\delta\rho_{\text{trans}} > c_0$ for all $\gamma > 0$ which implies existence of an extended metastability region.

We can now summarise the critical behaviour of the system at scale (39) in three cases, which are also illustrated in Fig. 7 (right).

Case 1 ($\delta\rho < c_0$) There exists a unique minimum of the scaling rate function at $\delta\rho_{\text{bg}} = \delta\rho$,

$$I_{\delta\rho}^{(2)}(\delta\rho) = 0,$$

corresponding to a stable fluid phase. The most likely background density under the canonical measure is $S_L^{\text{bg}}/(L-1) = \rho_c + \delta\rho L^{-\alpha} + o(L^{-\alpha})$.

Case 1 ($c_0 < \delta\rho < \delta\rho_{\text{trans}}$) There are two local minima of the scaling rate function

$$I_{\delta\rho}^{(2)}(r_0) > I_{\delta\rho}^{(2)}(\delta\rho) = 0,$$

corresponding to a stable fluid phase with $\delta\rho_{\text{bg}} = \delta\rho$ and a metastable condensed phase with $\delta\rho_{\text{bg}} = r_0$. The most likely background density under the canonical measure is still $S_L^{\text{bg}}/(L-1) = \rho_c + \delta\rho L^{-\alpha} + o(L^{-\alpha})$.

Case 1 ($\delta\rho > \delta\rho_{\text{trans}}$) There are two local minima of the scaling rate function

$$I_{\delta\rho}^{(2)}(\delta\rho) > I_{\delta\rho}^{(2)}(r_0) = 0,$$

corresponding to a metastable fluid phase with $\delta\rho_{\text{bg}} = \delta\rho$ and a stable condensed phase with $\delta\rho_{\text{bg}} = r_0$. The most likely background density under the canonical measure is now $S_L^{\text{bg}}/(L-1) = \rho_c + r_0 L^{-\alpha} + o(L^{-\alpha})$ and a finite fraction of the excess mass condenses on a single lattice site with $m = (\delta\rho - r_0)L^{1-\alpha} + o(L^{1-\alpha})$.

The value of $\delta\rho_{\text{trans}}$, and for $\delta\rho > \delta\rho_{\text{trans}}$ the position of the global minimum, agree with recent results on the distribution of the maximum at the critical scale [3]. The expected value of the maximum site occupation (equivalently background density) under the canonical measure can be interpreted in terms of the stable solution of the current matching argument as described above.

Corollary 2 (Canonical current overshoot) *Under the conditions of Theorem* 3, *for* $\delta\rho < \delta\rho_{\text{trans}}$

$$j_{L,N} - 1 = L^{-\alpha}\frac{\delta\rho}{\sigma_c^2} + o(1) \tag{57}$$

and for $\delta\rho > \delta\rho_{\text{trans}}$

$$j_{L,N} - 1 = \frac{bL^{-\alpha}}{(\delta\rho - r_0)^\gamma} + o(1). \tag{58}$$

Proof Follows directly by expressing $j_{L,N}$ in terms of the ratio of canonical partition functions (10) and applying (51). □

This result is illustrated in Fig. 7 (left). At the transition point $\delta\rho = \delta\rho_{\text{trans}}$ we must examine higher order terms to demonstrate that the system will asymptotically be in a condensed phase, see [3] and Sect. 6.

6 Estimating the Lifetime of Metastable Phases

Having established the existence of metastable phases we can use our previous results to get an estimate of their lifetime. Previous studies include results on the dynamics of fluctuations in a finite system [22] and the dynamics of the condensate [17, 18, 21]. Our approach here follows mostly the one in [20] where metastability in a different zero-range process has been studied, and [18] where a random walk argument was used to find the characteristic time of the motion of the condensate.

In this section we consider the dynamics in the scaling limit defined by (39). In Sect. 5.2 we have seen that for $\delta\rho > c_0$ the system exhibits two metastable phases, we assume in the following that we are in this region. So for sufficiently large L we know $I_{L,N}(m)$ exhibits a local maximum between the fluid and condensed minima. Let m^* be the local maximum of $I_{L,N}$ which is well approximated by the closest integer to the largest root of (29) (see Sect. 4). In the scaling limit $\frac{N-m^*}{L-1} = \rho_c + r^* L^{-\alpha}$ (see Sect. 5). We define the metastable fluid (condensed) phase as all states for which the maximum site occupation is below (above) m^*, i.e.

$$
\begin{aligned}
\pi_{L,N}^{\text{fluid}} &:= \pi_{L,N}[\cdot | M_L(\eta) \leq m^*] \\
\pi_{L,N}^{\text{cond}} &:= \pi_{L,N}[\cdot | M_L(\eta) > m^*].
\end{aligned}
\tag{59}
$$

Since the process spends very little time close to the maximum of $I_{L,N}$, the precise choice of m^* is not relevant for results in the scaling limit (cf. also [7] for a slightly different approach). The lifetimes of the two metastable phases can be expressed in terms of the following hitting time,

$$
\tau_L := \inf\left\{ t \geq 0 : M_L(\eta(t)) = m^* \right\}.
$$

We take

$$
T_{\text{fluid}} := 2\mathbb{E}_{\pi_{L,N}^{\text{fluid}} e^{\mathcal{L}t}}[\tau_L] \quad \text{and} \quad T_{\text{cond}} := 2\mathbb{E}_{\pi_{L,N}^{\text{cond}} e^{\mathcal{L}t}}[\tau_L],
\tag{60}
$$

where the expectation is with respect to the dynamics equation (2) with initial distribution given by $\pi_{N,L}^{\text{fluid}}$ or $\pi_{N,L}^{\text{cond}}$. The factor of 2 comes from the fact that once m^* is reached the process can still return to the same metastable phase with probability $1/2$.

By ergodicity of the process $(\eta(t) : t \geq 0)$ on $\Omega_{L,N}$ the ratio of lifetimes is directly related to the stationary distribution,

$$
\frac{T_{\text{fluid}}}{T_{\text{cond}}} = \frac{\pi_{L,N}[M_L(\eta) \leq m^*]}{\pi_{L,N}[M_L(\eta) > m^*]} = \frac{\sum_{k=0}^{m^*} e^{-L I_{L,N}(k)}}{\sum_{k=m^*+1}^{N} e^{-L I_{L,N}(k)}}.
\tag{61}
$$

Applying the scaling limit result of Theorem 3, we get for large L

$$
\begin{aligned}
\frac{T_{\text{fluid}}}{T_{\text{cond}}} &\simeq \int_{r^*}^{\delta\rho} e^{-L^\beta I_{\delta\rho}^{(2)}(x)} \, dx \left(\int_0^{r^*} e^{-L^\beta I_{\delta\rho}^{(2)}(x)} \, dx \right)^{-1} \\
&\simeq \frac{\Gamma(1 + \frac{1}{1-\gamma})(\frac{1-\gamma}{b})^{\frac{1}{1-\gamma}} \sqrt{\partial_x^2 I_{\delta\rho}^{(2)}(r_0)}}{\sqrt{2\pi L}} e^{-L^\beta (I_{\delta\rho}^{(2)}(\delta\rho) - I_{\delta\rho}^{(2)}(r_0))},
\end{aligned}
\tag{62}
$$

where the second line follows by a saddle point approximation of the denominator at the local minimum r_0 that corresponds to condensed configurations, and by expanding $I_{\delta\rho}^{(2)}(x)$ in the numerator at the boundary minimum (corresponding to fluid configurations), keeping the leading order singular term. To leading order in the exponent this result implies that the ratio of the lifetimes is given by the relative depth of the minima, as expected. The lower order term $\sqrt{L}$ in the denominator implies that at the transition density $\delta\rho_{\text{trans}}$ in the scaling limit the condensed phase is stable and the fluid phase is metastable, in accordance with results in [3]. However the scaling with L is difficult to verify in Monte Carlo simulations, since system sizes have to be relatively small (at most of the order of $L = 1000$) to get good

statistics on switching times. In this regime higher order finite size effects still play a role and can affect the location of $\rho_{\text{trans}}(L)$ and therefore the relative depth of the minima of $I_{L,N}$.

We estimate the two lifetimes by approximating the number of particles in the condensate by a continuous time random walk. Since it is difficult to validate the scaling with L, as discussed above, we will demonstrate the validity of our approach by estimating the lifetimes directly for finite systems. The actual lifetimes are not only related to the potential barrier given by the maximum of $I_{L,N}$, but also depend on the underlying dynamics. The occupation of the maximum $\big(M_L(\boldsymbol{\eta}(t)) : t \geq 0\big)$ is a non-Markovian, stationary, ergodic process with state space $\Omega^* = \{\lceil N/L \rceil, \lceil N/L \rceil + 1, \ldots, m^*, m^* + 1, \ldots, N\}$ and stationary distribution $\pi^*(m) = e^{-L\,I_{L,N}(m)}$. Since the process exhibits only single steps, it can be approximated by a continuous time random walk on Ω^* (the validity of this Markovian assumption is discussed later) where a particle leaves the maximum ($m \to m - 1$) with rate $g(m)$. The rates corresponding to $m \to m + 1$ are then fixed by the stationary distribution π^*. In the scaling limit this implies that particles enter the condensate with rate given by the current in the background, see Corollary 1. The lifetime of the metastable fluid phase is approximated by the mean first passage time of the random walk starting in state $\lceil N/L \rceil$ to reach m^*. For the metastable condensed phase it is given by the mean first passage time starting at N, and both have a factor of 2 in front due to the possibility of reaching m^* and not actually switching phase. These can be calculated using standard techniques (see e.g. [33] for the discrete time analog),

$$
\begin{aligned}
T_{\text{fluid}} &\approx 2 \sum_{i=\lceil N/L \rceil+1}^{m^*} \frac{\pi^*(i-1)}{g(i)\pi^*(i)} + 2 \sum_{i=\lceil N/L \rceil+2}^{m^*} \frac{1}{g(i)\pi^*(i)} \sum_{j=\lceil N/L \rceil}^{i-2} \pi^*(j) \\
T_{\text{cond}} &\approx 2 \sum_{i=m^*+1}^{N} \frac{1}{g(i)} + 2 \sum_{i=m^*+1}^{N} \frac{1}{g(i)\pi^*(i)} \sum_{j=i-1}^{N} \pi^*(j).
\end{aligned}
\tag{63}
$$

These two approximations can be calculated numerically for reasonably large systems (up to $L \approx 4000$) since we can calculate $\pi^*(m) = e^{-L\,I_{L,N}(m)}$ exactly using (22). Applying the scaling limit result and using saddle point approximations we recover an Arrhenius estimate with a prefactor for the lifetime of the two metastable phases,

$$
\begin{aligned}
T_{\text{fluid}} &\approx 2 \int_{r^*}^{\delta\rho} e^{L^\beta I_{\delta\rho}^{(2)}(x)} \int_x^{\delta\rho} e^{-L^\beta I_{\delta\rho}^{(2)}(y)}\, dy\, dx \\
&\approx \Gamma\left(1 + \frac{1}{1-\gamma}\right)\left(\frac{b}{1-\gamma}\right)^{\frac{-1}{1-\gamma}} \sqrt{\frac{2\pi L}{\partial_r^2 I_{\delta\rho}^{(2)}(r^*)}}\, e^{L^\beta (I_{\delta\rho}^{(2)}(r^*) - I_{\delta\rho}^{(2)}(\delta\rho))} \\
T_{\text{cond}} &\approx 2 \int_0^{r^*} e^{L^\beta I_{\delta\rho}^{(2)}(x)} \int_0^x e^{-L^\beta I_{\delta\rho}^{(2)}(y)}\, dy\, dx \\
&\approx \frac{2\pi L}{\sqrt{\partial_r^2 I_{\delta\rho}^{(2)}(r^*)\partial_r^2 I_{\delta\rho}^{(2)}(r_0)}}\, e^{L^\beta (I_{\delta\rho}^{(2)}(r^*) - I_{\delta\rho}^{(2)}(r_0))}.
\end{aligned}
\tag{64}
$$

Here r^* is the location of the local maximum of the rate function, r_0 is the local minimum corresponding to condensed configurations and $\delta\rho$ is the boundary minimum corresponding to fluid configurations. To leading order the lifetime of the metastable fluid (condensed) phase is given by the height of the local maximum of the rate function (potential barrier) above the fluid (condensed) minimum. This result agrees with the ratio of the lifetimes

Table 1 Lifetime of metastable phases for $L = 1000$ and $N = 695$, with $\gamma = 0.5$ and $b = 4$. Using approximation (63) compared to measurements from Monte Carlo simulations on a fully connected lattice (MF), a one dimensional lattice with $p = 1$ and with $p = 3/4$

	Approx. (63)	MF	1D $p = 1$	1D $p = 3/4$
T_{fluid}	1.61×10^5	1.65×10^5	1.21×10^5	2.05×10^5
T_{cond}	1.51×10^5	1.52×10^5	1.07×10^5	1.94×10^5

already discussed. However as with the previous results the asymptotic form is difficult to validate due to higher order finite size effects, the exact expression (63) is therefore more appropriate for relevant system sizes.

We expect the above random walk approximation to be accurate if particles exiting the maximum equilibrate in the bulk before returning to the condensate. This condition is best fulfilled for a fully connected lattice (mean-field geometry) and we expect it to be a reasonably good approximation for one dimensional totally asymmetric systems since particles have to pass through the whole fluid bulk before returning. For partial asymmetry and in higher dimensions a return without penetrating the fluid bulk is possible. To a first approximation the lifetimes have to be multiplied by the inverse probability of the event that particles escape into the bulk, since only such particles have a chance to equilibrate and contribute on the right scale. This pre-factor can be estimated using another random walk argument. If a particle jumps to the right with probability p and to the left with probability q where $p + q = 1$ and $p \neq 1/2$ then the probability of a particle reaching some macroscopic distance before returning to the condensate is asymptotically $|p - q|$. We therefore expect the lifetime to increase by a factor of $1/|p - q|$ for partial asymmetry. As a special case, for symmetric systems in one dimension this leads to an increase of lifetimes by a factor of L by the same argument. For a detailed investigation and validation of this Markovian ansatz see [18]. We find that this argument gives a good approximation for fully connected lattices, and totally asymmetric jumps in one dimension, see Table 1. We observe that the one dimensional totally asymmetric case is a factor of approximately 1.4 faster than the mean-field case, which is due to internal structure in the fluid background, and has also been observed and discussed in [18]. For partial asymmetry in one dimension with $p = 3/4$ the process is slower than the totally asymmetric case by a factor of approximately 1.8 supporting the arguments above which predict an increase by a factor of 2.

The characteristic time for the motion of the condensate in the thermodynamic limit, $N/L \to \rho > \rho_c$, can be approximated by considering (39) with a sequence of $\delta\rho$ increasing like L^α. The corresponding rescaled location of the condensed minimum $r_0 \to 0$, since the condensed phase has limiting background density ρ_c. Our rigorous results do not technically hold in this limit, however our estimates of the lifetimes of the metastable phases agree with the characteristic times found by Godreche and Luck [17, 18]. Our results demonstrate that close to the transition point in finite systems (i.e. $\delta\rho \approx \delta\rho_{\text{trans}}$) the condensate typically moves via the system entering the fluid phase (cf. Fig. 3), and therefore its new position is expected to be chosen uniformly at random on Λ_L. For $\delta\rho$ large, however, the lifetime of the fluid phase becomes small and the condensate can re-locate whilst the system remains in the condensed phase, via the mechanism described in [18]. A sub-condensate starts to grow in the fluid bulk, and the potential barrier the process has to cross for condensate motion is associated with the probability of the excess mass shared equally between two sites. This is the relevant mechanism in the thermodynamic limit. For finite systems one can estimate which of these re-location processes dominates, by considering the distribution of the second

largest site using the same techniques as in the previous section. It has recently been shown that the second mechanism via a sub-condensate can lead to a non-uniform relocation of the condensate [6].

7 Discussion

We have shown that a prototypical class of zero-range processes, that are known to undergo a condensation transition, exhibit large finite size effects including a metastable switching phenomenon for a large range of system parameters. We have characterized this behaviour using the background density as an order parameter, by rigorously deriving a large deviation rate function for its distribution in an appropriate scaling limit, which shows a double well structure. These results agree with recent work on the zero-range process at the critical scale [3], which we extend by establishing metastable fluid and condensed phases. Our methods give rise to a simple interpretation of these results in terms of a stationary current balance between the fluid background and the condensate. All results presented here, except those in Sect. 6, concern properties of the canonical stationary distribution and are therefore independent of the geometry or dimension of the lattice, so long as it permits homogeneous stationary distributions.

Our results allow us to estimate the lifetime of the two metastable phases using a heuristic random walk argument. The estimates agree with previous studies of the dynamics of the condensate in the thermodynamic limit [18]. We show that in the critical scaling limit the lifetime of the two phases is growing with system size (as a stretched exponential) and derive the appropriate Arrhenius law including prefactor. We have also demonstrated that on finite systems the re-location dynamics of the condensate is dominated by switching to the metastable fluid phase for a large parameter range, rather than the mechanism via a second sub-condensate assumed in previous results, which is relevant in the thermodynamic limit.

The zero-range process provides an interesting example where the thermodynamic limit fails to capture essential parts of the dynamics, and contradicts the usual expectation that finite systems should behave in a smoother fashion than the limiting prediction. Although the aim of this paper is not to discuss particular applications in detail, the finite size effects described here are particularly important for a proper understanding of various real condensation phenomena. To name two relevant examples, the condensing zero-range process has recently been applied as a simplified traffic model as well as to describe clustering phenomena in granular media [28, 30, 37, 39]. Both of these applications exhibit metastability phenomena, and typical system sizes are of order 10^2–10^3, the region where metastability effects in our analysis are most relevant. Since our results hold for a large generic class of zero-range models, they suggest that the apparent metastability in many applications is not necessarily the result of a particular choice of the jump rates, but rather a generic finite size phenomenon. While the results in the scaling limit hold for all zero-range processes with the same tail behaviour as the rates (1), the strength of the effect for relevant system sizes will of course depend on the details of the particular application. Traffic modelling is an example where metastability is particularly pronounced [41], and it is an interesting question to investigate in detail the applicability of the zero-range process and our results in that area.

Using the methods discussed in this paper, a direct application of the findings in [3] allows us to extend our results on the current overshoot to zero-range processes with rates (1) where $\gamma = 1$ and $b > 3$, see Fig. 8. The critical scale in this case turns out to be of order $\sqrt{L/\log L}$, and the system also shows a discontinuous current overshoot in the scaling limit.

Fig. 8 (Color online) The current overshoot in the scaling limit for $\gamma = 1$ and $b = 4$. The *thick black line* shows the scaling limit result and the rescaled canonical currents are shown for various system sizes by *thin coloured lines*, they slowly approaching the scaling limit as L increases. The transition point is given by $c_0(\gamma, b)$ which is given in [3]
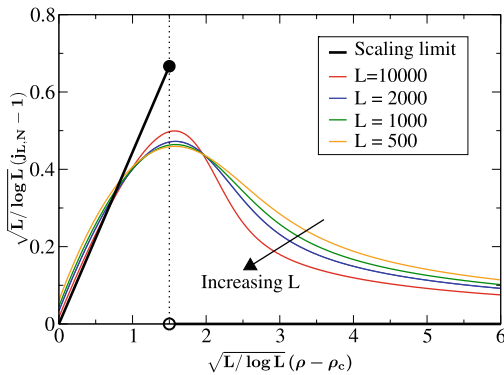
However, as can be seen in Fig. 8, convergence is very slow and there are no metastable phases on the critical scale. Although metastability might occur on higher order scales, it will hardly be relevant in any finite size simulation or application.

The scaling limit rate function in Theorem 3 could also be derived using a slightly modified version of the saddle point methods applied in [14, 15, 32]. However, the direct approximation of the fluid and condensed metastable phases used in this paper seems more intuitive for our purpose, and can be made rigorous with less effort than a saddle point computation.

Appendix

A.1 A Formula for $I_{L,N}$

In (32) we introduce $I_{L,N}(m) = -\frac{1}{L} \log \pi_{L,N}[M_L = m]$ for the canonical probability that the maximum site occupation is m. The following upper bound over counts configurations in which more than one site contains m particles

$$\pi_{L,N}[M_L = m] \leq \frac{Lw(m) \sum_{\boldsymbol{\eta} \in \hat{X}} \prod_{x=1}^{L-1} w(\eta_x)}{Z(L,N)}, \tag{65}$$

where $\hat{X} = \{\boldsymbol{\eta} : \eta_1, \ldots, \eta_{L-1} \leq m, \sum_{x=1}^{L-1} \eta_x = N - m\}$. The following lower bound does not count any configurations in which more than one site contains m particles,

$$\pi_{L,N}[M_L = m] \geq \frac{Lw(m) \sum_{\boldsymbol{\eta} \in \check{X}} \prod_{x=1}^{L-1} w(\eta_x)}{Z(L,N)}, \tag{66}$$

where $\check{X} = \{\boldsymbol{\eta} : \eta_1, \ldots, \eta_{L-1} \leq m - 1, \sum_{x=1}^{L-1} \eta_x = N - m\}$. It follows that

$$\pi_{L,N}[M_L = m] = \frac{Lw(m) \sum_{\boldsymbol{\eta} \in \hat{X}} \prod_{x=1}^{L-1} w(\eta_x)}{Z(L,N)} \left(1 - O(Le^{-\frac{-b}{1-\gamma} m^{1-\gamma}})\right) \tag{67}$$

as $L, m \to \infty$. We may write the last term in the numerator in terms of the cut-off grand canonical measure and get

$$\pi_{L,N}[M_L = m]$$

$$= \frac{Lw(m)z_m^{L-1}(\phi)\phi^{-(N-m)}v_{\phi,m}^{L-1}[S_{L-1} = N - m]}{Z(L,N)}\left(1 - O(Le^{-\frac{-b}{1-\gamma}m^{1-\gamma}})\right), \quad (68)$$

which holds for all $\phi \in (0, \infty)$ (cf. Sect. 4.1). Finally, taking logarithm,

$$-\frac{1}{L}\log \pi_{L,N}[M_L = m] = -\frac{(L-1)}{L}\left(\log z_m(\phi) - \frac{N-m}{L-1}\log \phi\right)$$

$$-\frac{1}{L}\log w(m) + \frac{1}{L}\log Z(L,N) - \frac{1}{L}\log L$$

$$-\frac{1}{L}\log v_{\phi,m}^{L-1}[S_{L-1} = N - m] + O(e^{-\frac{-b}{1-\gamma}m^{1-\gamma}}). \quad (69)$$

A.2 Convergence of Moments under the Cut-off Ensemble

Lemma 1 *Consider two sequences $m_n \in \mathbb{N}$, $\mu_n \in (0, \infty)$, such that $m_n \to \infty$, $\mu_n \to 0$ and $m_n^\gamma \mu_n \to C$ as $n \to \infty$ with $C \in [0, \frac{b}{1-\gamma})$. Then for each $i \in \mathbb{N}_0$*

$$\limsup_{n\to\infty} \sum_{k\le m_n} k^i w(k)e^{\mu_n k} < \infty.$$

Proof From the asymptotic behaviour of $w(k)$ (11) we know there exists a $C_0 \in (0, \infty)$ such that $w(k) \le C_0 e^{\frac{-b}{1-\gamma}k^{1-\gamma}}$ for all k. So,

$$\sum_{k\le m_n} k^i w(k)e^{\mu_n k} \le C_0 \sum_{k\le m_n} k^i e^{\frac{-b}{1-\gamma}k^{1-\gamma}+\mu_n k}. \quad (70)$$

Fix $0 \le \varepsilon < (\frac{b}{1-\gamma} - C)$. Then there exists $\bar{n} \in \mathbb{N}$ such that $\mu_n \le (C + \varepsilon)m_n^{-\gamma}$ for $n \ge \bar{n}$. It follows that $\mu_n \le (C + \varepsilon)k^{-\gamma}$ for all $k \le m_n$ and $n \ge \bar{n}$. By applying this upper bound on μ_n we can bound the sum above uniformly in $n > \bar{n}$ as follows,

$$\sum_{k\le m_n} k^i w(k)e^{\mu_n k} \le C_0 \sum_{k\le m_n} k^i e^{\frac{-b}{1-\gamma}k^{1-\gamma}+(C+\varepsilon)k^{1-\gamma}}$$

$$\le \int_0^\infty x^i e^{\frac{-b}{1-\gamma}x^{1-\gamma}+(C+\varepsilon)x^{1-\gamma}} < \infty. \quad (71)$$

$\square$

Lemma 2 *Under the assumptions of Lemma 1 we may bound the remainder in the Taylor expansion of each moment of the cut-off measure v_{ϕ_n, m_n} introduced in (23). Writing $\phi_n = e^{\mu_n}$ we have*

$$z_{m_n}(e^{\mu_n}) = z(1) + z'(1)\mu_n + o(\mu_n),$$

$$R_{m_n}(e^{\mu_n}) = \rho_c + \sigma_c^2 \mu_n + o(\mu_n),$$

$$\text{and} \quad \mathbb{E}_{v_{\mu_n, m_n}}[\eta_1^i] = \mathbb{E}_{v_{\mu_c}}[\eta_1^i] + \mu_n \partial_\mu \mathbb{E}_{v_{\mu_c}}[\eta_1^i] + o(\mu_n).$$

Proof For each $n \in \mathbb{N}$ the ith moment of the cut-off ensemble ν_{ϕ_n, m_n} is given by a finite sum. We let,

$$f_{m_n}(\mu) = \sum_{k \leq m_n} k^i w(k) e^{\mu k}. \tag{72}$$

then $f_{m_n} \in C^\infty(\mathbb{R}_+, \mathbb{R}_+)$. By Taylors theorem expanding around $\mu = 0$,

$$f_{m_n}(\mu_n) = f_{m_n}(0) + f'_{m_n}(0)\mu_n + h_n(\mu_n),$$

where the remainder term may be expressed as

$$h_n(\mu_n) = \frac{f_{m_n}^{(2)}(s)}{2}\mu_n^2 \quad \text{for some } s \in [0, \mu_n].$$

$f_{m_n}^{(2)}$ is non-negative and increasing and so the remainder is bounded by

$$0 \leq h_n(\mu_n) \leq \frac{f_{m_n}^{(2)}(\mu_n)}{2}\mu_n^2.$$

By Lemma 1, we know that there exists $U > 0$ such that $f_{m_n}^{(2)}(e^{\mu_n}) \leq U$ for all $n \in \mathbb{N}$. It follows that,

$$f_{m_n}(e^{\mu_n}) = f_{m_n}(1) + f'_{m_n}(1)\mu_n + o(\mu_n).$$

The result then follows by considering $|f_\infty(1) - f_{m_n}(1)| \sim \sum_{k > m_n} k^i e^{\frac{-b}{1-\gamma}k^{1-\gamma} + \mu_n k}$ which is $o(\mu_n)$ since $\mu_n m_n^\gamma \to C$, so the tail of the sum converges as a stretched exponential, this allows us to replace the expected values under the cut-off with $\mu = 0$ with the expected values under the true critical measure. $\qquad\square$

References

1. Andjel, E.D.: Invariant measures for the zero range process. Ann. Probab. **10**(3), 525–547 (1982)
2. Angel, A.G., Evans, M.R., Mukamel, D.: Condensation transitions in a one-dimensional zero-range process with a single defect site. J. Stat. Mech. **4**, 04001 (2004)
3. Armendáriz, I., Grosskinsky, S., Loulakis, M.: Zero range condensation at criticality (2009). arXiv:0912.1793v1 [math.PR]
4. Armendáriz, I., Loulakis, M.: Thermodynamic limit for the invariant measures in supercritical zero range processes. Probab. Theory Relat. Fields **145**(1), 175–188 (2009)
5. Behringer, H., Pleimling, M.: Continuous phase transitions with a convex dip in the microcanonical entropy. Phys. Rev. E **74**(1), 011108 (2006)
6. Beltrán, J., Landim, C.: Meta-stability and condensed zero-range processes on finite sets (2008). arXiv:0802.2171 [math.PR]
7. Beltrán, J., Landim, C.: Metastability of reversible condensed zero range processes on a finite set (2009). arXiv:0910.4089v1 [math.PR]
8. Beltrán, J., Landim, C.: Tunneling and metastability of continuous time Markov chains (2009). arXiv:0910.4088v1 [math.PR]
9. Biskup, M., Chayes, L., Kotecký, R.: On the formation/dissolution of equilibrium droplets. Europhys. Lett. **60**(1), 21 (2002)
10. Biskup, M., Chayes, L., Kotecký, R.: Critical region for droplet formation in the two-dimensional Ising model. Commun. Math. Phys. **242**(1), 137–183 (2003)
11. Davis, B., McDonald, D.: An elementary proof of the local central limit theorem. J. Theor. Probab. **8**(3), 693–701 (1995)

12. Evans, M.R.: Phase transitions in one-dimensional nonequilibrium systems. Braz. J. Phys. **30**(1), 42–57 (2000)
13. Evans, M.R., Hanney, T.: Nonequilibrium statistical mechanics of the zero-range process and related models. J. Phys. A, Math. Gen. **38**(19), R195–R240 (2005)
14. Evans, M.R., Majumdar, S.N.: Condensation and extreme value statistics. J. Stat. Mech., Theory Exp. **2008**(05), P05004 (2008)
15. Evans, M.R., Majumdar, S.N., Zia, R.K.P.: Canonical analysis of condensation in factorised steady states. J. Stat. Phys. **123**(2), 357–390 (2006)
16. Ferrari, P.A., Landim, C., Sisko, V.V.: Condensation for a fixed number of independent random variables. J. Stat. Phys. **128**(5), 1153–1158 (2007)
17. Godreche, C.: Dynamics of condensation in zero-range processes. J. Phys. A, Math. Gen. **36**(23), 6313–6328 (2003)
18. Godreche, C., Luck, J.M.: Dynamics of the condensate in zero-range processes. J. Phys. A, Math. Gen. **38**(33), 7215–7237 (2005)
19. Grosskinsky, S.: Equivalence of ensembles for two-species zero-range invariant measures. Stoch. Process. Appl. **118**(8), 1322–1350 (2008)
20. Grosskinsky, S., Schütz, G.M.: Discontinuous condensation transition and nonequivalence of ensembles in a zero-range process. J. Stat. Phys. **132**(1), 77–108 (2008)
21. Grosskinsky, S., Schütz, G.M., Spohn, H.: Condensation in the zero range process: Stationary and dynamical properties. J. Stat. Phys. **113**(3), 389–410 (2003)
22. Gupta, S., Barma, M., Majumdar, S.N.: Finite-size effects on the dynamics of the zero-range process. Phys. Rev. E **76**, 4 (2007)
23. Hirschberg, O., Mukamel, D., Schütz, G.M.: Condensation in temporally correlated zero-range dynamics. Phys. Rev. Lett. **103**(9), 090602 (2009)
24. Holley, R.: A class of interactions in an infinite particle system. Adv. Math. **5**(2), 291–309 (1970)
25. Hüller, A.: Finite size scaling at first order phase transition? Z. Phys. B, Condens. Matter **95**, 63–66 (1994)
26. Jeon, I., March, P., Pittel, B.: Size of the largest cluster under zero-range invariant measures. Ann. Probab. **28**(3), 1162–1194 (2000)
27. Kafri, Y., Levine, E., Mukamel, D., Schütz, G.M., Török, J.: Criterion for phase separation in one-dimensional driven systems. Phys. Rev. Lett. **89**(3), 035702 (2002)
28. Kaupuzs, J., Mahnke, R., Harris, R.J.: Zero-range model of traffic flow. Phys. Rev. E **72**(5, Part 2), 5 (2005)
29. Kim, S.W., Lee, J., Noh, J.D.: Particle condensation in pair exclusion process (2010). arXiv:1001.3563v1 [cond-mat.stat-mech]
30. Levine, E., Ziv, G., Gray, L., Mukamel, D.: Traffic jams and ordering far from thermal equilibrium. Physica A **340**, 636–646 (2004)
31. Liggett, T., Spitzer, F.: Ergodic theorems for coupled random walks and other systems with locally interacting components. Probab. Theory Relat. Fields **56**(4), 443–468 (1981)
32. Majumdar, S.N., Evans, M.R., Zia, R.K.P.: Nature of the condensate in mass transport models. Phys. Rev. Lett. **94**(18), 180601 (2005)
33. Murthy, K.P.N., Kehr, K.W.: Mean first-passage time of random walks on a random lattice. Phys. Rev. A **40**(4), 2082–2087 (1989)
34. Schwarzkopf, Y., Evans, M.R., Mukamel, D.: Zero-range processes with multiple condensates: statics and dynamics. J. Phys. A, Math. Theor. **41**(20), 205001 (2008) (21pp)
35. Spitzer, F.: Interaction of Markov processes. Adv. Math. **5**(246–290), 2 (1970)
36. Thompson, A.G., Tailleur, J., Cates, M.E., Blythe, R.A.: Zero-range processes with saturated condensation: the steady state and dynamics. J. Stat. Mech., Theory Exp. **2010**(02), P02013 (2010)
37. Török, J.: Analytic study of clustering in shaken granular material using zero-range processes. Physica A **355**(2–4), 374–382 (2005)
38. van der Meer, D., van der Weele, K., Lohse, D.: Sudden collapse of a granular cluster. Phys. Rev. Lett. **88**(17), 174302 (2002)
39. van der Meer, D., van der Weele, K., Reimann, P., Lohse, D.: Compartmentalized granular gases: flux model results. J. Stat. Mech. **07**, P07021 (2007)
40. van der Weele, K., van der Meer, D., Versluis, M., Lohse, D.: Hysteretic clustering in granular gas. Europhys. Lett. **53**(3), 328–334 (2001)
41. Wilson, R.E.: Mechanisms for spatio-temporal pattern formation in highway traffic models. Philos. Trans. R. Soc. A **366**(1872), 2017–2032 (2008)