

Mass condensation in one dimension with pair-factorized steady states

B Waclaw¹, J Sopik², W Janke¹ and H Meyer-Ortmanns²

¹ Institut für Theoretische Physik, Universität Leipzig, Postfach 100 920, 04009 Leipzig, Germany

² School of Engineering and Science, Jacobs University, PO Box 750561, 28725 Bremen, Germany

E-mail: Bartlomiej.Waclaw@itp.uni-leipzig.de, julien.sopik@googlemail.com, Wolfhard.Janke@itp.uni-leipzig.de and h.ortmanns@jacobs-university.de

Received 24 July 2009

Accepted 22 September 2009

Published 28 October 2009

Online at stacks.iop.org/JSTAT/2009/P10021

[doi:10.1088/1742-5468/2009/10/P10021](https://doi.org/10.1088/1742-5468/2009/10/P10021)

Abstract. We consider stochastic rules of mass transport which lead to steady states that factorize over the links of a one-dimensional ring. Based on the knowledge of the steady states, we derive the onset of a phase transition from a liquid to a condensed phase that is characterized by the existence of a condensate. For various types of weight functions which enter the hopping rates, we determine the shape of the condensate, its scaling with the system size, and the single-site mass distribution as characteristic static properties. As it turns out, the condensate's shape and its scaling are not universal, but depend on the competition between local and ultralocal interactions. So we can tune the shape from a delta-like envelope to a parabolic-like or a rectangular one. While we treat the liquid phase in the grand-canonical formalism, we develop a different analytical approach for the condensed phase. Its predictions are well confirmed by numerical simulations. Possible extensions to higher dimensions are indicated.

Keywords: driven diffusive systems (theory), transport processes/heat transfer (theory), stationary states, large deviations in non-equilibrium systems

ArXiv ePrint: [0907.4148](https://arxiv.org/abs/0907.4148)

Contents

1. Introduction	2
2. The model and its stationary states	4
3. Short-range interactions	6
3.1. Critical density for condensation	7
3.2. Extension of the condensate	9
3.3. Shape of the condensate	11
3.4. Fluctuations	17
3.5. Distribution of particles	20
4. The case when $p(m) \neq 1$ for all m	21
4.1. Fixed-envelope approximation	22
5. The case when $K(x)$ has heavy tails	24
5.1. Double power law: $K(x) \sim x^{-\nu}$ and $p(m) \sim m^{-b}$	24
5.2. Mixed decay law: $K(x) \sim x^{-\nu}$ and $p(m) = e^{U\delta_{m,0}}$	25
6. Conclusions and outlook	26
Acknowledgments	26
Appendix A	27
Appendix B	27
References	29

1. Introduction

Stochastic mass transport describes the movement of a generic mass from one space point to another as a classical stochastic process. ‘Mass’ stands for a generic quantity that is conserved throughout the process. It can be as different as a vehicle in macroscopic traffic or a macromolecule moving along the cytoskeleton [1]. Above a certain density, mass condensation may be observed, so that even in the limit of infinitely many masses a finite fraction of all masses piles up at a certain site, although the dynamic rules are fully symmetric. This is an expression of spontaneous symmetry breaking. In these systems it can happen even in one dimension, as they are in general out-of-equilibrium. When the mass transport describes vehicles, mass condensation corresponds to traffic jams. Other applications of condensation are found in granular flow [2], clustering [3], they are manifest as gelation in networks where a single node takes a finite fraction of all links [4], as Bose–Einstein condensation, or in phase transitions of quantum gravity [5].

Many of these systems can be modeled as a set of particles occupying discrete levels, or ‘boxes’ on a one-dimensional grid. The balls-in-boxes model (B-in-B) [6], its non-equilibrium version, the zero-range process (ZRP) [7], or a more general model [8] with continuous masses are well-known examples. Although these models are more abstract than realistic ones, they serve as paradigms since the stationary state can be derived

analytically as fully factorizing over the sites of the grid, or, more generally, over the nodes of an arbitrary graph [9]. The factorization is due to ultralocal (‘zero-range’) rules assumed to govern the dynamics of particles. This makes the phase structure of the systems accessible and enables comparison with experiments on condensation phenomena, at least on a qualitative level.

As a natural consequence of the lack of interactions between particles at neighboring sites, the condensates in the B-in-B or ZRP models—when they occur—always occupy a single site. The question then arises as to how the shape of the condensate changes in the presence of interactions between sites that tend to flatten out the condensate’s profile but still preserve translational symmetry and conserve the current. A first answer has been given in [10] for a model that is related to a solid-on-solid (SOS) model [11, 12], supplied with dynamical rules that drive the system out-of-equilibrium. There, the steady state factorizes over *pairs* of sites of a one-dimensional ring topology which allows for nearest-neighbor interactions while making the system analytically solvable. The condensate is extended over a range that scales as \sqrt{N} if N is the system size. The hopping rates were determined by weights $g(m, n)$ that factorize into an ultralocal part, depending on the occupation numbers of a single site, and a local part, depending on the occupation number of neighboring sites. For this model, the scaling of the condensate’s extension W with \sqrt{N} was derived as the first-passage time of a random walker after W time steps [10]. Since the first-passage-time density is asymptotically (for long times) independent of the distribution from which steps are drawn [13], one may naively expect that this analogy carries over also to other choices of weight functions, so that an extension $\sim\sqrt{N}$ would be a universal feature of condensates with PFSS. This argument, however, does not take into account that the number of particles in the condensate (which corresponds to the area under a random walker’s trajectory) is to a good approximation fixed, so that the random events of drawing the walker’s step size get correlated.

In this paper we shall show that, for the reason mentioned above, the scaling is non-universal. Actually, we can tune the shape of the condensate and its extension via different appropriately chosen weight functions. In terms of a random walk, this corresponds to taking only those paths into account which have a certain area beneath, which crucially affects the correlations of the walker’s steps. As it turns out, the relevant characteristics for both local and ultralocal weight factors is the range of interactions in occupation number space: we call short-range interactions local weight factors $K(x) = K(|m - n|)$ and ultralocal factors $p(m)$ which decay faster than any power in their argument, and long-range interactions those which decay as a power of their argument (x or m , respectively). For local interactions with exponential decay (short-range) and an ultralocal part that approaches a constant above some occupation number m_{\max} , we are able to analytically derive the static characteristics of the condensate: the critical density for condensation, the extension of the condensate for a given size, the shape of the condensate, that is its average occupation number as a function of spatial position, the fluctuations of the condensate and therefore also the mass distribution, that is the probability for finding m unit masses at a given site of the ring. The derivation is not restricted to the fluid phase for which the description in terms of a grand-canonical partition function holds up to the critical mass density. In our approach the condensed phase becomes analytically accessible due to the assumption that the partition function factorizes into a part Z_c that describes the contribution of the condensate, and a critical background Z_b , for which the uniform

particle distribution has just the critical density, so that all excess mass is absorbed by the condensate. The factor Z_c can be explicitly calculated when it is independent of the ultralocal part of the weight, i.e., independent of $p(m)$. This happens in particular for the choice of weight functions as it was used in [10]. Otherwise we use what we call the fixed-envelope approximation. Here we approximate Z_c via the probability $P(W)$ of having a condensate extended over W sites with unknown but fixed-envelope $h(t)$ for which we use a form that is inspired by numerical simulations of the condensation. The condensate's extension follows in these cases from the maximum of the probability $P(W)$.

As it turns out, within the same fixed-envelope approximation scheme, when local and ultralocal weights are 'short-ranged', we can tune the extension to scale with the system size as N^α with $0 \leq \alpha \leq 1/2$. When both parts of the weights are long-ranged, the condensate gets localized to a single site as in the ZRP. For a long-range local part and short-range ultralocal part, the condensate takes a rectangular form, its height scales proportionally to the system size, while its extension remains constant: features that are reminiscent of finite-size scaling at a first-order phase transition. Therefore hopping rates, leading to PFSS, do not necessarily lead to extended condensates, and if they do so, the shape of the condensate is non-universal.

The paper is organized as follows: in section 2 we introduce the model and derive its stationary state from the corresponding master equation. In section 3 we discuss short-range interactions with an ultralocal part that is constant above some threshold m_{\max} . Under mild assumptions on the factorization of the partition function we derive the critical density for condensation, the scaling of the extension of the condensate, the condensate's shape and its fluctuations, and the single-site distribution of particles, in particular also in the condensed phase. In sections 4 and 5 we use a more heuristic argumentation to predict the shape of the condensate for all other choices of weight functions, in particular for the power-like decay of local weights. Section 6 contains the summary and conclusions. Some supplementary material is provided in the appendices.

2. The model and its stationary states

We assume that we have M particles placed initially at random on N sites forming a closed chain. Each site i carries a certain number m_i of particles, which may range from zero to M . The periodicity implies that $m_{N+1} \equiv m_1$. With rate $u(m_i|m_{i-1}, m_{i+1})$ the particle jumps out of site i it occupies, and moves to one of the two neighboring sites. The hopping may be asymmetric: with probability r the particle departs to the right, and with probability $1 - r$ to the left. In a computer simulation, this corresponds to choosing at each time step the departure site at random and, if it is not empty, moving a single particle with probability proportional to $u(m_i|m_{i-1}, m_{i+1})$ (after a proper normalization) to one of the neighboring sites. Furthermore, we assume that

$$u(m_i|m_{i-1}, m_{i+1}) \equiv f(m_i, m_{i-1})f(m_i, m_{i+1}), \quad (1)$$

with some non-negative function $f(m, n)$. The hopping rate is thus a product of two factors, one for each nearest neighbor. As we shall see, this is the crucial assumption that leads to the pair-factorized steady state, but certainly it is not the only possibility. For instance, the model is very similar to the one from [10], with the only difference that there the hopping rate was a more general product of two factors $f_1(m_i, m_{i-1})$ and

$f_2(m_i, m_{i+1})$. Some aspects of the model (1) were already analyzed in [14]. In what follows we shall considerably expand our analysis in various directions. In particular, we shall (i) show how the factorized steady state emerges from a master equation for probability currents, (ii) present a more detailed derivation of the condensate's shape and extension, (iii) discuss Monte Carlo simulations of the system, (iv) calculate fluctuations around the average envelope, and (v) show how to use this information to find the distribution of particles in the condensed state and to explain finite-size effects. We shall also discuss some special choices of hopping rates more thoroughly, calculating the critical density (or at least showing conditions when it is finite) and the condensate's extension.

At each time t , the state of the system is specified by the set of occupation numbers $\vec{m} = \{m_1, \dots, m_N\}$. Let us denote by $P(\vec{m}, t)$ the probability of having the system in a particular state \vec{m} at time t . The evolution of $P(\vec{m}, t)$ is governed by the master equation:

$$\frac{dP(\vec{m}, t)}{dt} = \sum_{\vec{m}'} [W(\vec{m}' \rightarrow \vec{m})P(\vec{m}', t) - W(\vec{m} \rightarrow \vec{m}')P(\vec{m}, t)], \quad (2)$$

where $W(\vec{m}' \rightarrow \vec{m})$ is the transition rate from state \vec{m}' to state \vec{m} . In this paper we are interested in the steady-state solution of (2), that is in time-independent probabilities $P(\vec{m}) \equiv P(\vec{m}, t \rightarrow \infty)$. As follows from (2),

$$\sum_{\vec{m}'} P(\vec{m})W(\vec{m} \rightarrow \vec{m}') = \sum_{\vec{m}'} P(\vec{m}')W(\vec{m}' \rightarrow \vec{m}). \quad (3)$$

If both sides were equal separately for each \vec{m}' , this equation would simply correspond to the detailed balance condition and hence to the equilibrium case. We know, however, that for $r \neq 1/2$ the system cannot be at equilibrium, because there is a net current of particles to the right ($r > 1/2$) or to the left ($r < 1/2$). It turns out, however, that the solution of (3) does not depend on the hopping asymmetry r , and takes the same pair-factorized form either out of or in equilibrium:

$$P(\vec{m}) = \frac{1}{Z} \prod_{i=1}^N g(m_i, m_{i+1}) \delta_{m_1 + \dots + m_N, M}, \quad (4)$$

where Z stands for a normalization constant, provided that $g(m, n)$ is some symmetric function $g(m, n) = g(n, m)$ and

$$f(m, n) = \frac{g(m-1, n)}{g(m, n)}. \quad (5)$$

In what follows, we shall skip the Kronecker delta function $\delta_{m_1 + \dots + m_N, M}$ for brevity. We shall also skip the normalization constant Z but still call $P(\vec{m})$ a probability although, strictly speaking, it should be called statistical weight. To prove (4), let us observe that the transition probability $W(\vec{m} \rightarrow \vec{m}')$ is non-zero and reads $u(m_i | m_{i-1}, m_{i+1})$ only if $\vec{m}' = \{\dots, m_{i-1} - 1, m_i + 1, \dots\}$ or $\vec{m}' = \{\dots, m_i + 1, m_{i+1} - 1, \dots\}$, i.e., there is one more particle in the configuration \vec{m}' on an arbitrary site i and one less on one of its neighbors. From (3) we obtain:

$$\begin{aligned} \sum_i u(m_i | m_{i-1}, m_{i+1}) P(\vec{m}) &= \sum_i [ru(m_i + 1 | m_{i+1} - 1, m_{i-1}) P(\dots, m_i + 1, m_{i+1} - 1, \dots) \\ &+ (1 - r)u(m_{i+1} + 1 | m_i - 1, m_{i+2}) P(\dots, m_i - 1, m_{i+1} + 1, \dots)]. \end{aligned} \quad (6)$$

We proceed by inserting the guessed pair-factorized steady state (4), and

$$u(m_i|m_{i-1}, m_{i+1}) = \frac{g(m_i - 1, m_{i-1}) g(m_i - 1, m_{i+1})}{g(m_i, m_{i-1}) g(m_i, m_{i+1})}, \quad (7)$$

as follows from (1) and (5), into equation (6). After canceling some terms we obtain

$$\begin{aligned} & \sum_i g(m_{i-1}, m_i - 1) g(m_i - 1, m_{i+1}) g(m_{i+1}, m_{i+2}) R_i \\ &= r \sum_i g(m_{i-1}, m_i) g(m_i, m_{i+1} - 1) g(m_{i+1} - 1, m_{i+2}) R_i \\ &+ (1 - r) \sum_i g(m_{i-1}, m_i - 1) g(m_i - 1, m_{i+1}) g(m_{i+1}, m_{i+2}) R_i, \end{aligned} \quad (8)$$

where we used the symmetry of $g(m, n) = g(n, m)$ and $R_i = \prod_{j \neq \{i-1, i, i+1\}} g(m_j, m_{j+1})$. Both terms without the factor r cancel immediately. The remaining terms proportional to r cancel after shifting the index $i \rightarrow i + 1$ in the last sum. We have therefore proved that when the hopping rate takes the particular form (7), the system possesses a steady state factorized over pairs of neighboring sites. The proof can be extended to an arbitrary graph [15].

From the above discussion one sees that the hopping rate may be defined by the two-point weight function $g(m, n)$. It is this very weight function that we shall vary throughout the following sections. In general we shall factorize it into a local interaction factor $K(|m - n|)$, and an ultralocal factor $p(m)$ (playing the role of an on-site potential) according to

$$g(m, n) = K(|m - n|) \sqrt{p(m)p(n)}, \quad (9)$$

where both $K(x)$ and $p(m)$ are some positive functions of x and m , respectively. When $K(x) = 1$, $g(m, n)$ factorizes and we recover the ZRP with the weight $p(m)$ since every $\sqrt{p(m)}$ appears twice in the product over sites in (4). As indicated in section 1, we call it short-range interactions in occupation number space when $K(|m - n|)$ or $p(m)$ decay exponentially, and long-range interactions when they decay with some power of the argument.

The most intriguing phenomenon in the ZRP is certainly condensation when a finite fraction of all particles occupies a single site. Condensation is triggered by the on-site potential $p(m)$, which acts only on particles being at the same site. When we consider in the following PFSS with weights $K(x)$ depending on the difference of occupation numbers $x \equiv |m_i - m_{i+1}|$ at neighboring sites i and $i+1$, we can explore the effect of nearest-neighbor interactions on the very condensation. Therefore, in what follows we shall determine the critical density for condensation and the static properties of the condensate for various choices of $K(x)$ and $p(m)$.

3. Short-range interactions

In the first part of this section we will assume the following weights:

$$K(x) = e^{-Jx}, \quad p(m) = e^{U\delta_{m,0}}, \quad (10)$$

with parameters J and U , proposed in [10]. The steady state reads:

$$P(m_1, \dots, m_N) = \exp \left(-J \sum_i |m_i - m_{i+1}| + U \sum_i \delta_{m_i, 0} \right). \quad (11)$$

The delta function, reflecting the conservation of particles, was dropped for brevity. This will be our prime example, although we discuss later a more general case for which $K(x)$ is an arbitrary function that falls off faster than any power law, and $p(m)$ is constant for m greater than some m_{\max} . From (11) we see how to interpret J and U . Firstly, because $|m_i - m_{i+1}|$ measures the rate of change of m_i with i , fast-changing profiles are more suppressed, the larger J is. Thus, J gives a kind of surface stiffness, which tends to flatten out the profile. Secondly, the term $\sum_i \delta_{m_i, 0}$ measures the total number of free sites. The larger U is, the more probable are configurations with many empty sites. An interesting observation follows from comparison to the solid-on-solid (SOS) model [11]. In the SOS model without pinning potential, the probability of a microstate is essentially given by (11) with $U = 0$. Such a model does not have a phase transition in 1D. We will see below that it is the insertion of the U -term that leads to a liquid-condensed phase transition, even in the 1D equilibrium system.

3.1. Critical density for condensation

In order to study the properties of the steady state we can ignore the fact that the system is out-of-equilibrium, and treat it with conventional methods of statistical mechanics. Thinking of $P(\vec{m})$ as a probability of a microstate, we define the canonical partition function:

$$Z(N, M) = \sum_{\{m_i\}} \prod_i g(m_i, m_{i+1}) \delta_{\sum_i m_i, M}, \quad (12)$$

as well as the grand-canonical one

$$Z_N(z) = \sum_M Z(N, M) z^M = \sum_{\{m_i\}} z^{\sum_i m_i} \prod_i g(m_i, m_{i+1}), \quad (13)$$

which is just the discrete analog of a Laplace transform of $Z(N, M)$, and the fugacity z is determined from

$$\rho = \frac{1}{N} \left\langle \sum_i m_i \right\rangle = \frac{z}{N} \frac{\partial \ln Z_N(z)}{\partial z}. \quad (14)$$

The partition function $Z_N(z)$ grows monotonously with z , and so does its derivative. Thus the left-hand side of (14) grows also with z . If the radius of convergence of $Z_N(z)$ is infinite, then for any finite ρ there exists some $z > 0$ which obeys (14). Both ensembles, the canonical and the grand-canonical one, are then in our case equivalent in the thermodynamic limit. We will use this fact to calculate the distribution of particles $\pi(m)$:

$$\pi(m) = \frac{1}{N} \left\langle \sum_i \delta_{m, m_i} \right\rangle, \quad (15)$$

that is the probability of having m particles at a randomly chosen site. Taking the average in the grand-canonical ensemble we obtain

$$\pi(m) = \frac{1}{Z_N(z)} \sum_{m_2, \dots, m_N} T_{mm_2} T_{m_2 m_3} \cdots T_{m_N m}, \quad (16)$$

$$Z_N(z) = \sum_{m_1, \dots, m_N} T_{m_1 m_2} T_{m_2 m_3} \cdots T_{m_N m_1} = \text{Tr } T(z)^N, \quad (17)$$

where

$$T_{mn} = z^{(m+n)/2} g(m, n). \quad (18)$$

If we define now ϕ_m to be a normalized eigenvector of T_{mn} to the largest eigenvalue λ_{\max} ,

$$\sum_n T_{mn} \phi_n = \lambda_{\max} \phi_m, \quad (19)$$

we obtain for large N that $Z_N(z) \cong \lambda_{\max}^N$ and $\pi(m) = \phi_m^2$. The eigenvector ϕ_m has to decay with m . Otherwise, ρ calculated from (14) would be infinite in the thermodynamic limit. By analogy to the ZRP we can thus say that the system is in the liquid state—there is no condensation.

On the other hand, if $Z_N(z)$ has some finite radius of convergence z_c , the derivative in (14) can either grow to infinity for $z \rightarrow z_c$, or tend to some constant. In the first case we have again no condensation, because for any ρ there is some real $z < z_c$ which obeys (14). In the second case, there exists a critical density

$$\rho_c = \sum_m m \phi_m^2, \quad (20)$$

with ϕ_m being now the eigenvector for $z = z_c$, above which the grand-canonical ensemble does not exist. This in turn indicates a phase transition from the liquid to the condensed state.

In order to calculate the critical density, it is convenient to assume that $z_c = 1$. If $z_c \neq 1$, we can always shift it to one by rescaling the weight $g(m, n)$. If ϕ_m and λ_{\max} are now the eigenvector and the maximal eigenvalue of $T_{mn}(z = 1) = g(m, n)$, respectively, the recipe for obtaining the critical density is (i) to find the eigenvector of $g(m, n)$ to the largest eigenvalue, (ii) to square its elements, (iii) to find its mean value treating ϕ_m^2 as probabilities. This can be done by numerical diagonalization, truncating the matrix to a finite size and keeping track of the size dependence.

In some cases we can find the eigenvector analytically or, at least, we can decide whether ρ_c is finite or not. First, let us state that because $g(m, n) \geq 0$ for all m, n , from the Frobenius–Perron theorem all entries ϕ_m must be non-negative. Moreover, the eigenvector to the largest eigenvalue is non-degenerate. Since $g(m, n) = g(n, m)$, all eigenvalues are real, and eigenvectors to different eigenvalues are orthogonal. This means that all other eigenvectors must have some negative entries, thus if we find a semi-positive vector that obeys the equation:

$$\sum_n g(m, n) \phi_n = \lambda_{\max} \phi_m, \quad (21)$$

with some constant λ_{\max} , it has to be the one to the largest eigenvalue.

If $g(m, n) = g(M - m, M - n)$, it follows that $\phi_m = \phi_{M-m}$ and hence $\rho_c = M/2 \propto N$ tends to infinity for $N \rightarrow \infty$. This is the case of the SOS model, where $g(m, n) = e^{-J|m-n|}$; there is no condensation, regardless of how big the density ρ is. We see the importance of the U -term in (11). Its role is to break the symmetry in $g(m, n)$ under the change $(m, n) \rightarrow (M - m, M - n)$. When the symmetry is explicitly broken in the occupation number space, from (20) it follows that ϕ_m must decay faster than $\sim m^{-1}$ if the critical density has to be finite.

With the above remarks in mind, let us sketch the derivation of ρ_c for the weight (10). First, $z_c = 1$, so we do not need to shift the weights. As suggested by the functional form of $g(m, n)$, we assume that the eigenvector $\phi_m \propto e^{A\delta_{m,0} + Bm}$ with some constants A, B . Indeed, by inserting it into (21) with $g(m, n) = e^{-J|m-n| + U(\delta_{m,0} + \delta_{n,0})/2}$, one finds that the eigenvalue equation is fulfilled as long as

$$A = U/2, \quad (22)$$

$$B = -J - \ln(1 - e^{-U}). \quad (23)$$

One sees that when $J < J_0$, where $J_0 = U - \ln(e^U - 1)$, the eigenvector grows with m , therefore the critical density is infinite. This means that condensation is possible only for $J > J_0$. The unnormalized eigenvector reads $\phi_m \propto e^{U\delta_{m,0}/2 - (J - J_0)m}$, and hence the critical density becomes

$$\rho_c = \frac{\sum_{m=0}^{\infty} m \phi_m^2}{\sum_{m=0}^{\infty} \phi_m^2} = \frac{e^{J_0} - 1}{(e^{J_0} - e^{-2(J - J_0)})(e^{2(J - J_0)} - 1)}. \quad (24)$$

When $J \rightarrow J_0$, the above formula reduces to $\rho_c \approx 1/(e^{2(J - J_0)} - 1)$, the result obtained in [10] where the asymptotic form $\phi_m \sim e^{Bm}$ was assumed. As a byproduct we obtain the largest eigenvalue:

$$\lambda_{\max} = e^U + \frac{1}{e^{2J}(1 - e^{-U}) - 1}. \quad (25)$$

3.2. Extension of the condensate

When the density ρ exceeds ρ_c , the condensate emerges in the system. In figure 1 we present a snapshot of the condensate obtained from numerical simulations for $N = 1000$ sites, $J = U = 1$ and $\rho = 10$. The critical density is $\rho_c = 0.2397$, thus the system is deeply in the condensed phase. One sees that the condensate extends over many sites, in contrast to the zero-range process where it is always localized. In [10] the extension W , that is the number of sites occupied by the condensate, was estimated to be proportional to \sqrt{N} for sufficiently large systems. The argument was based on a similarity between the envelope of the condensate, and a trajectory of a random walker. Here we will employ another approach, which will allow us to determine not only the extension, but also the average shape of the condensate, its fluctuations and the distribution of particles $\pi(m)$ in the condensed phase.

Let us assume that $\rho > \rho_c$ and consider the weight P_n of a configuration, where n sites share the surplus of particles $M' = N(\rho - \rho_c)$ and where $N - n$ sites form a ‘background’ with occupation numbers of order ρ_c . For $n = 1$, namely for only one site carrying the

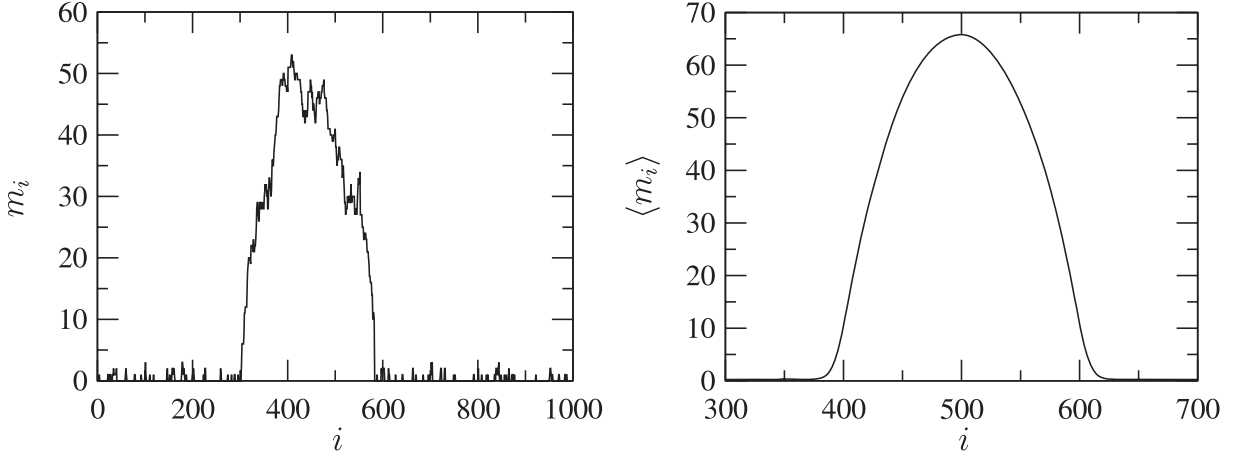


Figure 1. Left: a snapshot of the condensate for $N = 1000$ and $\rho = 10$ obtained in a MC simulation. Right: the shape of the condensate averaged over many realizations, each centered at $i = 500$.

condensate, P_1 is approximately

$$P_1 = NZ_{N-1}(1)K^2(M' - \rho_c)p(M' - \rho_c). \quad (26)$$

The first factor N stands for N possibilities of choosing the site occupied by the condensate. The second term accounts for $N - 1$ ‘background’ sites, it is just the partition function (17) for $N - 1$ sites. The third term stands for two ‘domain walls’ at the borders of the condensate. The fourth term accounts for the on-site potential. Since $p(m) \neq 1$ only for empty sites, the fourth term is simply equal to one. Hence the function $p(m)$ influences only Z_{N-1} via the maximal eigenvalue λ_{\max} . Using the definition (10) of $K(x)$ and neglecting ρ_c since it is small in comparison to $M' \rightarrow \infty$, we can write:

$$P_1 = N \exp[(N - 1) \ln \lambda_{\max} - 2JM']. \quad (27)$$

For $n = 2$, the condensate may occupy either two neighboring sites, or two condensates exist, separated by some number of background sites. The statistical weights for these two situations are, respectively,

$$P_{2,\text{single}} \approx NZ_{N-2}(1)K(xM')K(xM' - (1-x)M')K((1-x)M'), \quad (28)$$

$$P_{2,\text{double}} \approx N^2Z_{N-2}(1)K^2(xM')K^2((1-x)M'), \quad (29)$$

where x is the fraction of particles occupying the first site, and as before we dropped ρ_c in the arguments. These formulae can be rewritten as

$$P_{2,\text{single}} = P_1 e^{-\ln \lambda_{\max} + JM'(1-|2x-1|)}, \quad (30)$$

$$P_{2,\text{double}} = P_1 e^{-\ln \lambda_{\max} + \ln N}, \quad (31)$$

so that for any $0 < x < 1$, the two-site single condensate with probability $P_{2,\text{single}}$ dominates in the thermodynamic limit over (i) the one-site condensate P_1 and (ii) the case $P_{2,\text{double}}$ when the two sites carrying the condensate are separated.

Similarly, one can show that for $n = 3, 4, \dots$, P_n grows with n . The first observation is that the probability is higher when all n sites keep together. For such a case, assuming that the i th site takes a portion x_i of the condensate, we obtain

$$P_n \propto \exp \left(-n \ln \lambda_{\max} - JM' \left(\sum_{i=1}^{n-1} |x_{i+1} - x_i| + x_1 + x_n \right) \right), \quad (32)$$

where the factor independent of n is dropped, and $x_1 + \dots + x_N = 1$. The weight P_n has a maximum if

$$f(\vec{x}) = \sum_{i=1}^{n-1} |x_{i+1} - x_i| + x_1 + x_n \quad (33)$$

takes the minimum value. To find this minimum, let us observe that $f(\vec{x})$ can be always lowered if we order x_1, \dots, x_n so that x_1 and x_n are minimal, and both sequences x_1, x_2, \dots and x_n, x_{n-1}, \dots are increasing. Then, x_i has a maximum at some $i = i_0$. We have:

$$f(\vec{x}) = x_1 + \sum_{i=1}^{i_0-1} (x_{i+1} - x_i) + \sum_{i=i_0}^{n-1} (x_i - x_{i+1}) + x_n = 2x_{i_0}, \quad (34)$$

and, since all x_i must sum up to one, the smallest value is $x_{i_0} = 1/n$. Therefore, $P_n \sim \exp(-n \ln \lambda_{\max} - 2JM'/n)$ clearly grows with n as long as $n \ll \sqrt{N}$ and hence we see that the condensate must be spread over many sites. One could argue that we did not take into account the entropic contribution, namely in how many ways we can choose the set of x_1, \dots, x_n . But the entropic contribution cannot be larger than $\sim N^n = e^{n \ln N}$, so as long as n stays finite, it only changes the constant multiplying n in (32). But if n grows with N as $\sim \sqrt{N}$, the two terms $\sim n$ and $\sim M'/n$ become comparable. This could suggest that if they both were negative, P_n would take its maximum at $n \sim \sqrt{N}$ and the spatial extension of the condensate would be of order $\sim \sqrt{N}$. However, the ‘naive’ entropic term $n \ln N$ dominates over the $-n \ln \lambda_{\max}$ term for very large N and therefore configurations with growing n seem to be more probable. In other words, this argument would imply that the spatial extension should be $\sim N$, and there would be no condensation at all. The problem is caused by the fact that we overestimated the entropic contribution. In section 3.3 we will show that a more accurate calculation yields that the entropy $\sim n$ and the extension is indeed $\sim \sqrt{N}$.

3.3. Shape of the condensate

The interesting feature of the condensation discussed so far is that the condensate is extended over many sites. Apart from the question concerning the scaling of the extension with the system size, to which we will return soon, another interesting question concerns the shape of the condensate. By the shape we understand here the condensate’s envelope, that is we ask about the average occupation numbers $\langle m_i \rangle$ as a function of the index i . A numerical recipe for measuring the envelope in a Monte Carlo (MC) simulation is straightforward: one has to accumulate the shape over many realizations, each time shifting the condensation peak to a common origin. The only difficulty is to find the center of the condensate. This can be done in many ways which are equivalent in the thermodynamic limit. For any finite size, however, we found the following method to be

the best. First, one searches for the maximal occupation number, let it be the one at site i_{\max} . Then, starting from i_{\max} , one searches for the left i_{left} and the right i_{right} border of the condensate where the occupation numbers drop to the background level ρ_c . The center of the condensate is then defined as $i_0 = (i_{\text{left}} + i_{\text{right}})/2$. This method preserves the sharp borders of the condensate seen in figure 1, left, quite well, as can be seen in the same figure, right, where we show the envelope obtained as an average over 20 000 MC samples. Since steady-state properties of the model do not depend on the fact whether it is in equilibrium or not, to simulate the system we treated it as if it were in equilibrium with the probability of a microstate (11). The model can then be simulated using the standard MC method by picking up a random particle and moving it to a randomly chosen site, accepting or rejecting the move with Metropolis probability. This leads to different dynamic properties than the original model has, but considerably speeds up the convergence towards the steady state in the condensed phase.

A careful inspection of figure 1 shows that the envelope is neither circular (which would be true if particles behaved as in a droplet of water due to some surface tension) nor Gaussian (the most frequent distribution so that it may be naively expected to be observed), but has some parabolic-like shape. A simple argument shows that the shape cannot be determined from ‘energetic’ considerations alone, which do not take the entropy into account. Imagine that the condensate has a single maximum and falls monotonically on both sides. Then the ‘energy’ which is the logarithm of the steady-state probability (11) and reads approximately $-W \ln \lambda_{\max} - J \sum_i |m_{i+1} - m_i| = -W \ln \lambda_{\max} - 2JH$, depends only on the height H and the width W , but not on the shape. This means that the principle of energy minimization alone does not allow us to estimate the shape.

Therefore, to find the envelope of the condensate, one has to take the average over all possible shapes, weighted by $P(\vec{m})$. Below we will show how to do this. Since the whole reasoning is valid for more general weights (9), as long as $K(x)$ falls off faster than any power law, and $p(m) = 1$ for $m > m_{\max}$, we will keep the discussion quite general. At the end, we will present formulae for the special case (10).

Let us remark first that since the grand-canonical partition function (13) does not exist in the condensed state, a naive approach via evaluating (13) for the fugacity z determined by (14) does not work. One should in principle work with the canonical ensemble, but direct calculation of (12) is very hard. This is because in (12) we have $p(m)$ which depends on occupation numbers, as well as $K(|m - n|)$ which depends on their differences. This prevents us from decoupling terms for different sites and to perform the sum over $\{m_i\}$ directly. One can observe, however, that in the condensed state the system may be split into two parts: a condensate with W sites, and a background with $N - W$ sites. The background can be treated as being at the critical point, because the condensate absorbs all the surplus of balls. Summing over all configurations in the background we shall get the grand-canonical partition function for the system with $N - W$ sites and ρ_c balls per site on average, which reads $Z_{N-W}(1) = \lambda_{\max}^{N-W}$, as follows from (17). From the formulae

$$\langle m_i \rangle = \frac{z}{NZ_N(z)} \left. \frac{\partial Z_N(z)}{\partial z} \right|_{z=1}, \quad (35)$$

$$\langle m_i(m_i - 1) \rangle = \frac{z^2}{NZ_N(z)} \left. \frac{\partial^2 Z_N(z)}{\partial z^2} \right|_{z=1}, \quad (36)$$

applied to the background we obtain that the variance $\langle m_i^2 \rangle - \langle m_i \rangle^2$ is constant. Thus the total number of particles in the background fluctuates as $\sim \sqrt{N - W}$. This in turn means that M' , the number of particles in the condensate, also fluctuates as $\sim \sqrt{N - W} \ll M'$. We shall therefore neglect fluctuations of the condensate's mass, and assume that it always takes M' particles. The width and the height of the condensate can still fluctuate. Moreover, the condensate, from its definition, has occupation numbers growing with N . Since we assumed $p(m) = 1$ for large m , we can drop $p(m)$ in the condensate. On the borders, however, $m_i \approx \rho_c$ and since this is very small in comparison with other m_i , we can set it to zero. The statistical weight of the condensate extended over W sites is then

$$Z_c(W) \cong \sum_{\{m_k\}} \prod_{k=1}^{W+1} K(|m_k - m_{k-1}|) \delta \left[\sum_{k=1}^W m_k - M' \right], \quad (37)$$

where we assumed $m_0 = m_{W+1} = 0$ at the borders. Here the function $\delta[n]$ denotes the Kronecker delta $\delta_{n,0}$ and will be used from now on for notational convenience instead of $\delta_{n,0}$. The total weight of having the system composed of the condensate and the background will be assumed as a product of probabilities:

$$P(W) \cong Z_{N-W}(1) Z_c(W) = \exp((N - W) \ln \lambda_{\max} + \ln Z_c(W)). \quad (38)$$

To find the extension W , one has to find the maximum of (38) with respect to W .

We will focus now on calculating $Z_c(W)$ defined in (37). Due to the Kronecker delta constraint and the factors $K(|m_k - m_{k-1}|)$ which suppress large differences in neighboring occupation numbers, we expect that the majority of occupation numbers will be much greater than zero. Therefore, we assume that the summation over $\{m_i\}$ can be extended to negative values without changing $Z_c(W)$ too much. We will show later by calculating average occupation numbers and their variances that this assumption is fully justified, that is in fact almost all $m_k > 0$ in most probable configurations. We can now rewrite $Z_c(W)$ as

$$Z_c(W) \approx \sum_{d_1=-\infty}^{\infty} \cdots \sum_{d_{W+1}=-\infty}^{\infty} \delta \left[- \sum_{k=1}^{W+1} k d_k - M' \right] \delta \left[\sum_{k=1}^{W+1} d_k \right] \prod_{k=1}^{W+1} K(|d_k|), \quad (39)$$

where $d_k = m_k - m_{k-1}$. The first Kronecker delta in the above formula gives the conservation of particles. The second delta reflects fixed boundary conditions according to $m_0 = m_{W+1} = 0$, which gives $d_1 + \cdots + d_{W+1} = 0$. Let us introduce an auxiliary function

$$G(W, \vec{u}) = \sum_{d_1=-\infty}^{\infty} \cdots \sum_{d_W=-\infty}^{\infty} \delta \left[- \sum_{k=1}^W k d_k - M' \right] \delta \left[\sum_{k=1}^W d_k \right] \prod_{k=1}^W K(|d_k|) e^{d_k u_k}, \quad (40)$$

where $\vec{u} = \{u_1, \dots, u_W\}$ are auxiliary variables. We will use them later, let us now only observe that $Z_c(W) = G(W + 1, \vec{0})$ and hence $G(W, \vec{u})$ is a generating function for the moments of d_k :

$$\langle d_k^n \rangle = \left[G(W, \vec{u})^{-1} \frac{d^n}{d u_k^n} G(W, \vec{u}) \right]_{\vec{u}=0}. \quad (41)$$

Replacing both delta functions in (40) by their integral representations:

$$\delta[x] = \int_{-i\pi+\epsilon}^{i\pi+\epsilon} \frac{dz}{2\pi i} e^{xz}, \quad (42)$$

and performing the sum over $\{d_k\}$, we obtain:

$$G(W, \vec{u}) = \int_{-i\pi+\epsilon_1}^{i\pi+\epsilon_1} \frac{dz}{2\pi i} \int_{-i\pi+\epsilon_2}^{i\pi+\epsilon_2} \frac{dv}{2\pi i} e^{F(z,v,\vec{u})}, \quad (43)$$

where

$$F(z, v, \vec{u}) = -M'z + \sum_{k=1}^W \ln \tilde{K}(u_k + v - kz). \quad (44)$$

The function $\tilde{K}(x)$ is defined as

$$\tilde{K}(x) = \sum_{d=-\infty}^{\infty} K(|d|) e^{dx}. \quad (45)$$

The function $G(W, \vec{u})$ can be evaluated in the saddle point: $G(W, \vec{u}) \sim e^{F(z,v,\vec{u})}$, with $z = z(\vec{u})$, $v = v(\vec{u})$ being solutions to the saddle-point equation $\partial_z F(z, v, \vec{u}) = \partial_v F(z, v, \vec{u}) = 0$. Let us assume for a moment that we have solved this equation and have determined both $z(\vec{u})$ and $v(\vec{u})$. With help of (41), we can then write:

$$\langle d_k \rangle \cong \left[\frac{d}{du_k} F(z(\vec{u}), v(\vec{u}), \vec{u}) \right]_{\vec{u}=\vec{0}}, \quad (46)$$

$$\text{var}(d_k) = \langle d_k^2 \rangle - \langle d_k \rangle^2 \cong \left[\frac{d^2}{du_k^2} F(z(\vec{u}), v(\vec{u}), \vec{u}) \right]_{\vec{u}=\vec{0}}, \quad (47)$$

$$\text{cov}(d_j, d_k) = \langle d_j d_k \rangle - \langle d_j \rangle \langle d_k \rangle \cong \left[\frac{d}{du_j} \frac{d}{du_k} F(z(\vec{u}), v(\vec{u}), \vec{u}) \right]_{\vec{u}=\vec{0}}, \quad (48)$$

and similarly for higher-order correlation functions. We stress that in general all derivatives have to be taken for $z(\vec{u}), v(\vec{u})$ being functions of \vec{u} , and only at the end one can set $\vec{u} = \vec{0}$. But for the first moment we have

$$\langle d_k \rangle = \left[\frac{\partial F}{\partial z} \frac{\partial z}{\partial u_k} + \frac{\partial F}{\partial v} \frac{\partial v}{\partial u_k} + \frac{\partial F}{\partial u_k} \right]_{\vec{u}=\vec{0}}, \quad (49)$$

where the partial derivatives of F are taken in the saddle point and thus are zero, except of the last one. We thus obtain:

$$\langle d_k \rangle = \frac{\tilde{K}'(v - kz)}{\tilde{K}(v - kz)}. \quad (50)$$

The first observation that eliminates z without the need of solving the saddle-point equation is that the averaged peak must be symmetric around its center. This is true even if the system is not in equilibrium, that is when a current of particles flows through the system for $r \neq 1/2$. The symmetry of the shape implies that $\langle m_k \rangle = \langle m_{W-k} \rangle$, thus

$\langle d_k \rangle = -\langle d_{W-k} \rangle$ and hence $z = (2/W)v$. We obtain:

$$\langle d_k \rangle = \frac{\tilde{K}'(v(1 - (2k/W)))}{\tilde{K}(v(1 - (2k/W)))}. \quad (51)$$

In addition, because the total number of particles in the condensate equals M' , we have for large systems:

$$\sum_{k=0}^W k \langle d_k \rangle \cong \int_0^W k \frac{\tilde{K}'(v(1 - (2k/W)))}{\tilde{K}(v(1 - (2k/W)))} dk = -M'. \quad (52)$$

The last equation, after changing variables, reduces to

$$w = v \left[\frac{1}{2} \int_0^v \frac{x \tilde{K}'(x)}{\tilde{K}(x)} dx \right]^{-1/2}, \quad (53)$$

where neither M' nor W appear alone and where we defined the reduced extension $w \equiv W/\sqrt{M'}$. Equation (53) fixes w as a function of v . We can now write

$$\ln Z_c(W) \cong F \left(\frac{2v}{w(v)\sqrt{M'}}, v, \vec{0} \right), \quad (54)$$

and combining equations (54), (44) and (38) we obtain that the logarithm of the weight of a condensate extended over W sites is given by

$$\ln P(W(v)) = N \ln \lambda_{\max} + \sqrt{M'} \left[-w(v) \ln \lambda_{\max} - \frac{2v}{w(v)} + \frac{w(v)}{v} \int_0^v \ln \tilde{K}(x) dx \right]. \quad (55)$$

Since W does not appear in the above formula explicitly, but only through the relation $W = w(v)\sqrt{M'}$, the extension W is now determined by maximizing (55) with respect to v . Hence, W must be proportional to $\sqrt{M'} \sim \sqrt{N}$. Thus, we have proved that the extension grows with the square root of the system size.

One can go one step further and calculate v and w , and thus the exact, numeric value of the extension W . The probability (55) assumes a maximal value, when its derivative with respect to v is zero. Inserting (53) into (55) and taking the derivative d/dv we obtain:

$$\left(\ln \lambda_{\max} - \ln \tilde{K}(v) \right) \left[\frac{1}{v} - \frac{v \tilde{K}'(v)}{2 \tilde{K}(v)} \frac{1}{\int_0^v (x \tilde{K}'(x)/\tilde{K}(x)) dx} \right] = 0. \quad (56)$$

One can show (see appendix A) that the second term in the square brackets is never zero. Only the first term counts and gives:

$$v = \tilde{K}^{-1}(\lambda_{\max}). \quad (57)$$

Having determined v and w , it is now easy to obtain the shape of the condensate:

$$\langle m_n \rangle = \left\langle \sum_{k=1}^n d_k \right\rangle = \sum_{k=1}^n \langle d_k \rangle, \quad (58)$$

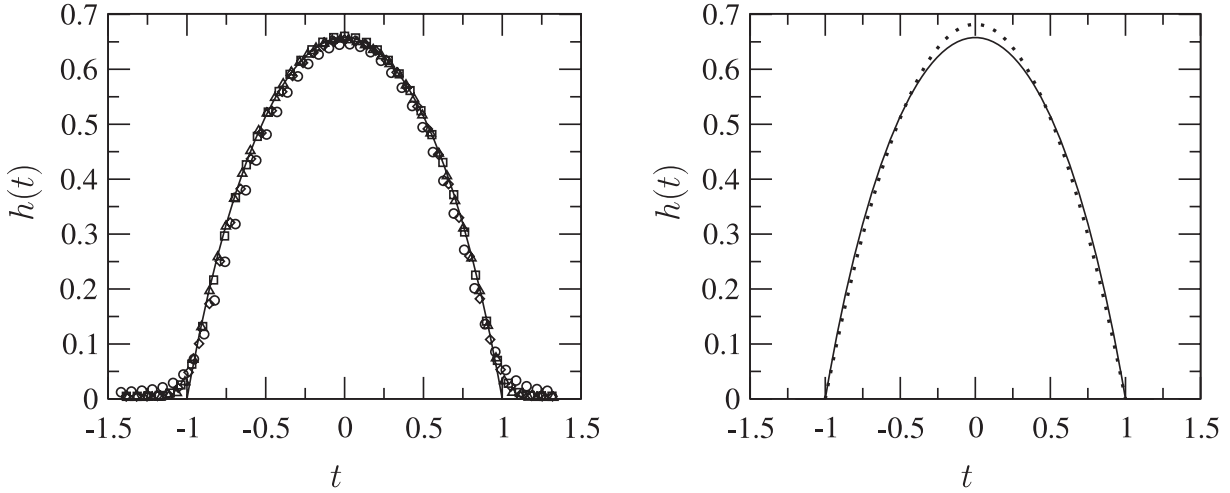


Figure 2. Left: comparison between the envelope of the condensate from (62) and rescaled MC data for the weights (10) with $J = U = 1$ for $N = 1000$, $\rho = M/N = 1$ and 3 (circles, squares) and $N = 4000$, $\rho = 1, 3$ (diamonds, triangles). Right: the envelope (62) (solid line) compared to the parabola (64) (dotted line).

where $\langle d_k \rangle$ is given by (50). If we define new variables:

$$h \equiv \langle m_n \rangle / \sqrt{M'}, \quad t = \frac{2n}{w\sqrt{M'}} - 1, \quad (59)$$

so that $t \in [-1, 1]$, we can write the formula for the envelope $h(t)$ as follows:

$$h(t) = \frac{w}{2v} \ln \frac{\tilde{K}(v)}{\tilde{K}(vt)}. \quad (60)$$

Thus the shape in reduced variables $t, h(t)$ becomes independent of the system size and the density of particles. The parameters v, w depend on the weights $K(x)$ and $p(m)$ only.

Let us now go back to the choice (10) of the weight functions. For this case, one obtains

$$\tilde{K}(x) = \frac{\sinh J}{\cosh J - \cosh x}, \quad (61)$$

and v , calculated from (57), assumes a very simple form: $v = J - J_0$. Inserting v obtained from the above equation into (53), we can calculate w numerically. The shape is given by (60) and reads

$$h(t) = \frac{w}{2v} \ln \left(\frac{\cosh J - \cosh vt}{\cosh J - \cosh v} \right), \quad (62)$$

with v, w being functions of the parameters J, U . For instance, for $J = U = 1$ one obtains $v = 0.5413$ and $w = 2.2005$. In figure 2 we show a comparison between (62) and MC simulations for different system sizes. One sees that the curves obtained for different N, ρ and plotted in the rescaled variables t, h agree with the theoretical line. Small deviations in tails can be attributed to finite-size effects and will be discussed in section 3.4, but

one sees already in figure 2 that slopes of the envelope become more steep for increasing system size N .

One may now ask the following question. The shape of $h(t)$ is determined by (39), which is just the partition function for a Brownian excursion: a 1D random walk with steps drawn from $K(x)$ starts at $m = 0$ and ends also at $m = 0$ after W steps. The only difference in comparison to ordinary excursions is that the area under the trajectory is fixed to some value M' . If this area is not fixed, it fluctuates around $\sim W^{3/2}$, the result known for a ‘free’ excursion [16]. Assuming $M' \propto W^{3/2}$ in our formulas we have $w = W/\sqrt{M'} \propto W^{1/4}$ and from (53) we obtain that

$$v = \frac{6\tilde{K}(0)}{\tilde{K}''(0)w^2} \rightarrow 0. \tag{63}$$

Thus, by applying the Taylor expansion to (60) one obtains a universal curve, independent of $K(x)$:

$$h(t) = (6/4w)(1 - t^2). \tag{64}$$

It must be stressed, however, that in our case M' scales always as W^2 , so the above argument does not hold, and the envelope is not universal. In figure 2, right, we see that despite a lack of universality in general, the exact form (62) does resemble the parabola from (64) very much, at least in a certain range of J, U .

3.4. Fluctuations

In this section we shall consider only the special choice (10). Assuming that the width W of the condensate is fixed, the fluctuations around the average envelope are measured by:

$$\text{var}(m_n)_W = \sum_{k=1}^n \sum_{j=1}^n (\langle d_k d_j \rangle - \langle d_k \rangle \langle d_j \rangle) = \sum_{k=1}^n \text{var}(d_k) + \sum_{k=1}^n \sum_{j \neq k}^n \text{cov}(d_k, d_j), \tag{65}$$

where the subscript W tells us that W is assumed to be constant. We need to calculate

$$\begin{aligned} \text{var}(d_k) &= \frac{\partial^2 F}{\partial u_k^2} + \frac{\partial^2 F}{\partial z^2} \left(\frac{\partial z}{\partial u_k} \right)^2 + \frac{\partial^2 F}{\partial v^2} \left(\frac{\partial v}{\partial u_k} \right)^2 \\ &+ 2 \frac{\partial^2 F}{\partial z \partial u_k} \frac{\partial z}{\partial u_k} + 2 \frac{\partial^2 F}{\partial v \partial u_k} \frac{\partial v}{\partial u_k} + 2 \frac{\partial^2 F}{\partial z \partial v} \frac{\partial z}{\partial u_k} \frac{\partial v}{\partial u_k} \end{aligned} \tag{66}$$

and

$$\begin{aligned} \text{cov}(d_k, d_j) &= \frac{\partial^2 F}{\partial z^2} \frac{\partial z}{\partial u_k} \frac{\partial z}{\partial u_j} + \frac{\partial^2 F}{\partial v^2} \frac{\partial v}{\partial u_k} \frac{\partial v}{\partial u_j} + \frac{\partial^2 F}{\partial z \partial v} \left(\frac{\partial v}{\partial u_k} \frac{\partial z}{\partial u_j} + \frac{\partial v}{\partial u_j} \frac{\partial z}{\partial u_k} \right) \\ &+ \frac{\partial^2 F}{\partial z \partial u_k} \frac{\partial z}{\partial u_j} + \frac{\partial^2 F}{\partial z \partial u_j} \frac{\partial z}{\partial u_k} + \frac{\partial^2 F}{\partial v \partial u_k} \frac{\partial v}{\partial u_j} + \frac{\partial^2 F}{\partial v \partial u_j} \frac{\partial v}{\partial u_k}, \end{aligned} \tag{67}$$

where all partial derivatives of $F(z, v, \vec{u})$ defined in (44) are taken in the saddle point with $\vec{u} = 0$. Other derivatives, which in principle may appear in the above formula, are zero and have been dropped. Since this is rather technical, we postpone the derivation to

appendix B. Here we shall only present the final result:

$$\text{var}(m_n)_W = \frac{w\sqrt{M'}}{2}\sigma^2 \left(\frac{2n}{w\sqrt{M'}} - 1 \right), \quad (68)$$

where the size-independent variance $\sigma^2(y)$ is given by

$$\sigma^2(y) = I_0(y) - \frac{I_0^2(y)}{I_0(1)} - \frac{I_1^2(y)}{I_2(1)}, \quad (69)$$

with

$$I_m(t) \equiv \int_{-1}^t y^m \frac{\cosh J \cosh(vy) - 1}{(\cosh J - \cosh(vy))^2} dy \quad (70)$$

derived in appendix B. The formula (68) would give the fluctuations of the envelope if the width W was kept fixed. In reality, the width also fluctuates around the mean $W = w\sqrt{M'}$. To take this into account we have to sum over all possible widths, each having the probability $P(W)$ given by (55),

$$\text{var}(m_n) = \sum_W P(W) \langle m_n^2 \rangle_W - \left(\sum_W P(W) \langle m_n \rangle_W \right)^2 \quad (71)$$

$$= \sum_W P(W) [\text{var}(m_n)_W + (\langle m_n \rangle_W - \langle m_n \rangle)^2]. \quad (72)$$

Here $\langle m_n \rangle_W$ means the envelope for given W , while $\langle m_n \rangle$ means the envelope averaged over W . Since the probability $P(W)$ is concentrated around $W = w\sqrt{M'}$ for large systems, the latter can be approximated by $\sqrt{M'}h(t)$, where $t = 2n/(w\sqrt{M'}) - 1$. Also the variance $\text{var}(m_n)_W$ summed with weights $P(W)$ is well approximated by $\frac{1}{2}w\sqrt{M'}\sigma^2(t)$. Then, from the above equation it follows that

$$\text{var}(m(t)) \approx \sqrt{M'} \left[\frac{w}{2}\sigma^2(t) + \sqrt{M'} \int P(W) \left(h \left(t \frac{w\sqrt{M'}}{W} \right) \frac{w\sqrt{M'}}{W} - h(t) \right)^2 dW \right], \quad (73)$$

so there is an additional contribution to the variance from fluctuations of the condensate's width. The probability $P(W)$ can be well approximated by a Gaussian distribution with the mean $w\sqrt{M'}$ and variance $\text{var}(W) \equiv \sqrt{M'}s^2$ with

$$s^2 = w \frac{(w^2/4)\tilde{K}'(v) - \tilde{K}(v)}{v\tilde{K}'(v)}. \quad (74)$$

This formula comes from expanding the probability $P(W(v))$ around the minimum for $v = \tilde{K}^{-1}(\lambda_{\max})$, and is true for any $\tilde{K}(x)$. For our special case (10) it reads

$$s^2 = w \frac{w^2/4 + (\cosh(v) - \cosh(J))/\sinh(v)}{v}. \quad (75)$$

Then, equation (73) can be rewritten as

$$\text{var}(m(t)) \approx \sqrt{M'} \left[\frac{w}{2}\sigma^2(t) + \frac{wM'^{3/4}}{\sqrt{2\pi s^2}} \int e^{-(w^2\sqrt{M'}/2s^2)(x-1)^2} \left(\frac{1}{x}h \left(\frac{t}{x} \right) - h(t) \right)^2 dx \right]. \quad (76)$$

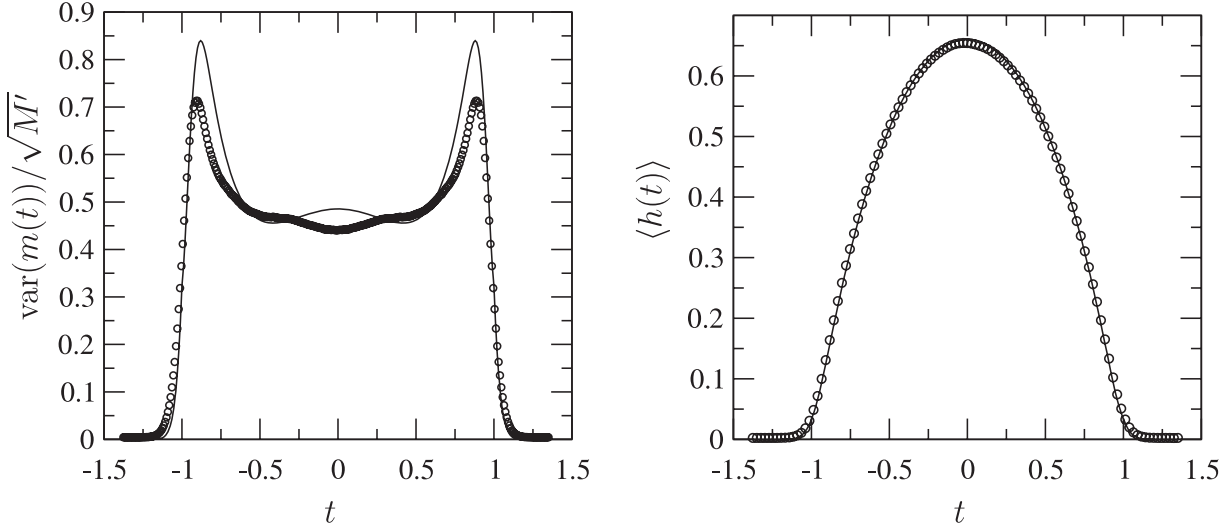


Figure 3. Left: the variance $\text{var}(m_n)$ obtained from MC simulations for $N = 4000, \rho = 3$ (circles), compared to (76) evaluated numerically (thin line). Right: the averaged envelope $\langle h(t) \rangle$ from (78), taking into account small fluctuations of M' .

In figure 3, left, we compare the result (76) and the one obtained by means of MC simulations. To a good approximation, the variance $\text{var}(m_n)$ grows linearly with $W \sim \sqrt{M'}$. This means that in the thermodynamic limit the number of occupation numbers being less than zero becomes negligible, because typically $m_n = \langle m_n \rangle \pm \sqrt{\text{var}(m_n)} \sim W \pm \sqrt{W}$, so that extending the summation in (39) to negative values was fully allowed because it leads to self-consistent results. At the center of the condensate, for $t = 0$, the formula (76) simplifies to

$$\text{var}(m(0)) \approx \sqrt{M'} \left(\frac{w}{2} \sigma^2(0) + \frac{s^2}{w^2} h^2(0) \right). \quad (77)$$

In section 3.5 we shall see that this formula is very important for the problem of the distribution of particles.

As a byproduct we obtained the variance $\sqrt{M'} s^2$ of the distribution of the width W . This allows us to calculate the envelope $h(t)$ averaged over fluctuations of W :

$$\langle h(t) \rangle \approx \frac{w M'^{1/4}}{\sqrt{2\pi s^2}} \int e^{-(w^2 \sqrt{M'} / 2s^2)(x-1)^2} \frac{1}{x} h\left(\frac{t}{x}\right) dx. \quad (78)$$

In figure 3, right, we show the curve obtained from the above equation. It fits to the MC data points even better than equation (62), especially in the tails. This indicates that smooth tails are mainly due to the fluctuations of the condensate's width. This effect clearly vanishes with increasing system size, i.e., is a typical finite-size effect.

It turns out that not only $\langle m_n \rangle$ and $\text{var}(m_n)$ but also the distribution of m_n for any fixed n can be predicted. Since m_n is a sum of many (almost independent) random variables d_k , each of them having finite variance, the distribution tends to a Gaussian, according to the Central Limit Theorem, provided that $1 \ll n \ll W$, that is we are not

too close to the borders. In the thermodynamic limit we can safely assume that all m_n have Gaussian distributions.

3.5. Distribution of particles

Let us denote by $\pi_b(m), \pi_c(m)$ the probability of finding m particles at a background site and a condensate site, respectively. The distribution $\pi_b(m) \propto \phi_m^2$, as follows from the partition function (13) at the critical point. In particular, for the previous weights one obtains $\pi_b(m) \propto [e^{-J}/(1 - e^{-U})]^{2m} e^{U\delta_{m,0}}$. The distribution in the condensate, $\pi_c(m)$, is given by summing Gaussian distributions with mean $\langle m_n \rangle$ and variance $\text{var}(m_n)$ over all $n = 1, \dots, W$ condensate sites:

$$\pi_c(m) \propto \sum_{n=1}^W \frac{1}{\sqrt{2\pi \text{var}(m_n)}} \exp \left[-\frac{(m - \langle m_n \rangle)^2}{2\text{var}(m_n)} \right], \quad (79)$$

which gives

$$\pi_c(m) \propto \int_0^1 \frac{dt}{\sqrt{2\pi \text{var}(m(t))}} \exp \left[-\frac{(m - \sqrt{M'}h(t))^2}{2\text{var}(m(t))} \right]. \quad (80)$$

By changing variables $t \rightarrow y = h(t)$ one obtains

$$\pi_c(m) \propto \int_0^{h(0)} \frac{dy}{\sqrt{2\pi \text{var}(m(h^{-1}(y)))}} (h^{-1})'(y) \exp \left[-\frac{(m - \sqrt{M'}y)^2}{2\text{var}(m(h^{-1}(y)))} \right]. \quad (81)$$

Hence the probability $\pi_c(m)$ may be viewed as a convolution of $(h^{-1})'(y)$ with a Gaussian distribution which smears the thermodynamic-limit distribution $\pi_c(m) \sim (h^{-1})'(m/\sqrt{M'})$. Since the variance grows like $\sqrt{M'}$, the smearing acts on distances $\sim M'^{1/4}$ and influences the profile of $\pi_c(m)$ only at $y \approx h(0)$, because $(h^{-1})'(y)$ is narrow only at $y = h(0)$. We can therefore assume $\text{var}(m(t)) = \text{var}(m(0))$ because it affects the distribution only at $t = 0$, and write

$$\pi_c(m) \propto \int_0^1 \frac{dt}{\sqrt{2\pi \text{var}(m(0))}} \exp \left[-\frac{(m - \sqrt{M'}h(t))^2}{2\text{var}(m(0))} \right], \quad (82)$$

with $\text{var}(m(0))$ given by (77). Since the main contribution to $\pi_c(m)$ comes from the flat region in $h(t)$, the condensate's peak tends to a Gaussian for large systems.

The probability $\pi(m)$ of finding m particles at any site is additively composed according to

$$\pi(m) = \frac{N - W}{N} \pi_b(m) + \frac{W}{N} \pi_c(m). \quad (83)$$

In figure 4 we show a comparison between $\pi(m)$ obtained from (83) after inserting $\pi_b(m)$ and $\pi_c(m)$ normalized to probabilities, with MC data. The agreement is very good.

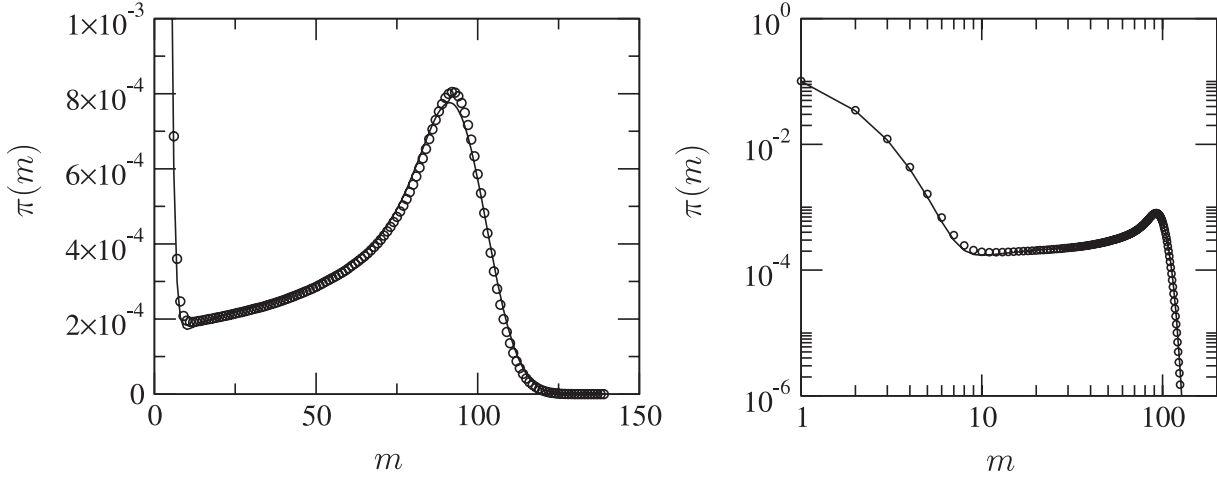


Figure 4. Left: the distribution of particles $\pi(m)$ obtained from (83) (thin line) and from MC simulations for $N = 8000, \rho = 3$ (circles). Right: the same on a log-log scale.

4. The case when $p(m) \neq 1$ for all m

As long as $K(x)$ falls off faster than any power law, and $p(m) = \text{const}$ for m larger than some m_{max} , the method from section 3 directly applies. This means that one can calculate the envelope of the condensate, its fluctuations and the distribution of particles. Since $p(m)$ can be multiplied by any factor without changing ‘physical’ quantities, the condition on $p(m)$ can be assumed to be $p(m > m_{\text{max}}) = 1$.

However, if $p(m)$ does not tend to one for large m , the partition function cannot be approximated by (37). An important example is

$$K(x) \sim e^{-a|x|^\beta}, \quad p(m) \sim e^{-bm^\gamma}, \quad (84)$$

for $a, b, \beta, \gamma > 0$, with both $K(x)$ and $p(m)$ decaying fast with x and m , respectively. In order to have the condensation, the grand-canonical partition function must have a finite radius of convergence. This means that the function

$$\sum_m g(m, n) z^m = e^{-bn^\gamma} \sum_m e^{-a|m-n|^\beta - bm^\gamma} z^m \quad (85)$$

must diverge for z larger than some critical z_c . It is easy to check that z_c is finite for $\gamma < 1$ and there is condensation above some critical density. For $\gamma > 1$, the radius of convergence is infinite, so there is no condensation for $\gamma > 1$. The value of β is not important here and can be arbitrarily large.

Next we have to determine for which values of β, γ the condensate is extended. Let us calculate the contribution from the condensate having $\sim N$ particles at a single site. This is the same approach as in section 3.2:

$$P_1 \approx N c^{N-1} p(N) K^2(N) \approx N c^{N-1} e^{-bN^\gamma - 2aN^\beta}. \quad (86)$$

As before, N stands for N possible condensate’s positions, c^{N-1} with some constant c accounts for background sites, $p(N)$ is the on-site potential at the condensed site and

$K^2(N)$ stands for the two domain walls. The contribution from the condensate residing on two adjacent sites, having respectively $N/2 + \epsilon$ and $N/2 - \epsilon$ particles, reads

$$P_2 \approx Nc^{N-2}K(N/2 + \epsilon)K(N/2 - \epsilon)K(2\epsilon)p(N/2 + \epsilon)p(N/2 - \epsilon) \quad (87)$$

$$\approx Nc^{N-2}e^{-a(N/2+\epsilon)^\beta - a(N/2-\epsilon)^\beta - a|2\epsilon|^\beta - b(N/2+\epsilon)^\gamma - b(N/2-\epsilon)^\gamma}. \quad (88)$$

The contribution is large only for $\epsilon \approx 0$. We can then write:

$$\ln P_1/P_2 \approx -\ln c - 2a(1 - 2^{-\beta})N^\beta - b(1 - 2^{1-\gamma})N^\gamma. \quad (89)$$

Since $\gamma < 1$, $1 - 2^{1-\gamma} < 0$. On the other hand, $1 - 2^{-\beta} > 0$ and hence

$$\ln P_1/P_2 \sim -N^\beta + N^\gamma, \quad (90)$$

where the proportionality factors have been skipped. For $\gamma > \beta$ we thus obtain $P_1/P_2 \rightarrow \infty$, so the condensate must be located on a single site. For $\gamma < \beta$, however, the contribution from two sites is larger than from a single one. Similarly, one can show that the contribution grows for 3, 4, 5, ... sites. Therefore the condensate must be extended, but we still do not know, for how much.

4.1. Fixed-envelope approximation

We have learned that fluctuations of the occupation numbers can be neglected if $p(m) \rightarrow 1$. This is also the case here, because the fluctuations are determined by the variance of $K(x)$, which is again finite. We shall thus assume that, for sufficiently large systems, the envelope of the condensate is essentially fixed, modulo some small fluctuations around it. Then, the logarithm of the weight of the condensate extended over W sites is

$$\ln P(W) \approx -W \ln \lambda_{\max} + Ws + \sum_k \ln K(\langle m_{k+1} - m_k \rangle) + \sum_k \ln p(\langle m_k \rangle). \quad (91)$$

Here s is some ‘entropic’ factor due to small fluctuations of the m_k . Values of $\langle m_k \rangle$ are equal to $Hh(2k/W - 1)$, where $H = M'/W$ is the height of the condensate and $h(t)$ is some fixed ‘envelope’ having the same meaning as in section 3.3, but its functional form may be unknown. The differential term $\langle m_{k+1} - m_k \rangle$ is more difficult. It can behave in two distinct ways. First, if we assume that $h(t)$ is a ‘smooth’ (differentiable) function, the difference can be rewritten as $(2M'/W^2)h'(t)$ with $t = 2k/W - 1$. A second possibility is that $h(t)$ has a rectangular shape, $h'(t)$ does not exist at $t = \pm 1$. Thus, we may have in principle two extended condensates: the one whose shape resembles that in the previous sections, with smooth $h(t)$, which we will call ‘smooth’, and a ‘rectangular’ one where $h(t)$ looks like a step function in the limit of large N . We can write

$$\ln P_{\text{smooth}}(W) \approx W \left[c + \int_0^1 \ln K \left(\frac{2M'}{W^2} h'(t) \right) dt + \int_0^1 \ln p(Hh(t)) dt \right], \quad (92)$$

$$\ln P_{\text{rect}}(W) \approx Wc + 2 \ln K \left(h(1) \frac{M'}{W} \right) + W \int_0^1 \ln p(Hh(t)) dt, \quad (93)$$

where $c = s - \ln \lambda_{\max}$ is some unknown constant. To find the extension W , we must find the larger of the maximum values of the two $P(W)$. Assuming now the weights (84), with

$0 < \gamma < 1$, $\beta > \gamma$, we have:

$$\ln P_{\text{smooth}}(W) \approx W \left[c - a \left(\frac{2M'}{W^2} \right)^\beta \int_0^1 |h'(t)|^\beta dt - b \left(\frac{M'}{W} \right)^\gamma \int_0^1 |h(t)|^\gamma dt \right], \quad (94)$$

$$\ln P_{\text{rect}}(W) \approx Wc - a \left(\frac{M'}{W} \right)^\beta 2|h(1)|^\beta - bW \left(\frac{M'}{W} \right)^\gamma \int_0^1 |h(t)|^\gamma dt. \quad (95)$$

The integrals are just some constants. The value c must be smaller than zero, because otherwise $P(W)$ has no maximum which is in contradiction to the fact that the condensate is extended. To estimate the extension, we can drop constants (remembering that they are all positive, except c) and write

$$\ln P_{\text{smooth}}(W) \sim -W - W \left(\frac{M'}{W^2} \right)^\beta - W \left(\frac{M'}{W} \right)^\gamma, \quad (96)$$

$$\ln P_{\text{rect}}(W) \sim -W - \left(\frac{M'}{W} \right)^\beta - W \left(\frac{M'}{W} \right)^\gamma. \quad (97)$$

Let us consider $P_{\text{smooth}}(W)$ first. If $\beta > 1/2$, it has a maximum for $d \ln P_{\text{smooth}}/dW = 0$ which reads

$$(1 - 2\beta) \left(\frac{M'}{W^2} \right)^\beta + \left(\frac{M'}{W} \right)^\gamma = -1, \quad (98)$$

and hence

$$W \sim M'^{(\beta-\gamma)/(2\beta-\gamma)}, \quad (99)$$

because the exponents of M' in both terms must be equal. We also obtain the maximal value:

$$\ln P_{\text{smooth}}(W) \sim -M'^{(\gamma\beta+\beta-\gamma)/(2\beta-\gamma)}. \quad (100)$$

If $\beta < 1/2$, the above equation has no solution, and $\ln P_{\text{smooth}}(W)$ takes its maximal value at $W \sim 1$, which would mean ‘delta-like’. This is impossible, so this solution is obviously not valid for $\beta < 1/2$, where the condensate must therefore be rectangular and described by (97), which has the maximum for

$$\beta M'^\beta W^{-\beta-1} - (1 - \gamma) M'^\gamma W^{-\gamma} = 1. \quad (101)$$

From the above equation we obtain the extension of the rectangular condensate

$$W \sim M'^{(\beta-\gamma)/(\beta-\gamma+1)}, \quad (102)$$

as well as

$$\ln P_{\text{rect}}(W) \sim -M'^{\beta/(\beta-\gamma+1)}. \quad (103)$$

If $\beta > 1/2$, both equations (96) and (97) are valid. But for $M' \rightarrow \infty$, it turns out that $\ln P_{\text{rect}}(W) > \ln P_{\text{smooth}}(W)$ for $1/2 < \beta < 1$. Therefore, collecting everything together we see that the rectangular-shaped condensate dominates for $\beta < 1$, and the smooth condensate dominates for $\beta > 1$. In figure 5 we show a phase diagram which summarizes all the above results. One sees there that the extension $W \sim M'^\alpha$ can be tuned to any $0 \leq \alpha < 1/2$, but cannot be larger than $1/2$.

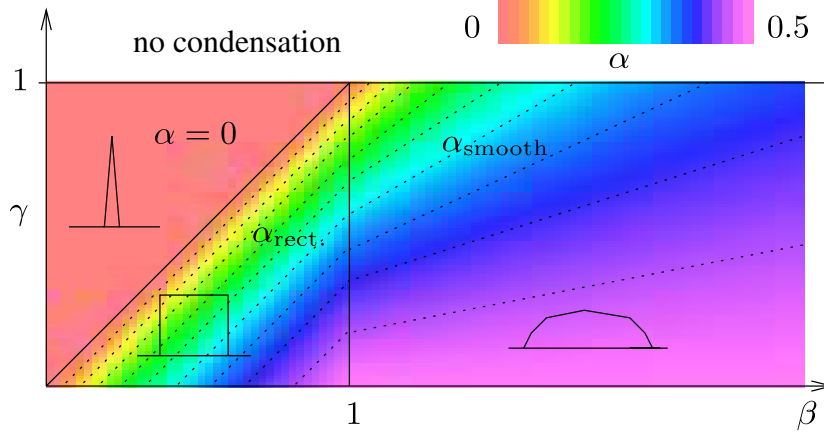


Figure 5. Phase diagram for $K(x) \sim e^{-|x|^\beta}$ and $p(m) \sim e^{-m^\gamma}$. Values of the exponent α in the extension $W \sim M^\alpha$, $\alpha_{\text{rect}} = (\beta - \gamma)/(\beta - \gamma + 1)$ for the rectangular and $\alpha_{\text{smooth}} = (\beta - \gamma)/(2\beta - \gamma)$ for the smooth condensate, are represented by the color (gray) code. Dotted lines mark $\alpha = 0.05, 0.1, \dots, 0.45$.

5. The case when $K(x)$ has heavy tails

In this section we consider $K(x)$ decaying as a power law: $K(x) \sim x^{-\nu}$. For this case, the analytic method from section 3 cannot be applied even if $p(m)$ is constant above some m_{max} . The reason is that the series (45) does not converge for $K(x)$ decaying more slowly than exponentially. On the other hand, the assumption that fluctuations vanish in the thermodynamical limit, a crucial step in the fixed-envelope approximation, is also not well justified *a priori*, because $K(x)$ may have infinite variance. Fortunately, it will turn out to be quite easy to show that the condensate is never extended over more than some finite number of sites, so there will be no need to use any of the two methods mentioned.

5.1. Double power law: $K(x) \sim x^{-\nu}$ and $p(m) \sim m^{-b}$

Let us first discuss the criterion for condensation. For $b > 0$, the eigenvector ϕ_m has the following form for large m :

$$\phi_m \sim m^{-b/2-\nu}, \quad (104)$$

which can be checked by inserting it into (21). The critical density

$$\rho_c \sim \sum_m m^{-b-2\nu+1} \quad (105)$$

is finite when $\nu > 1 - b/2$. The above formula is valid also for $\nu < 0$, that is when configurations with large differences between neighboring occupation numbers are favored. For $\nu = 0$, that is when $K(x) \sim \text{const}$, we recover the ZRP case: the critical density is finite for $b > 2$. If $b < 0$, the solution (104) is not valid, but then ϕ_m has to grow with m . The critical density is therefore infinite for $b < 0$.

To estimate the extension, let us first calculate the contribution from the condensate having $\sim N$ particles at a single site. Following the same method as before one obtains

$$P_1 \approx Nc^{N-1}p(N)K^2(N) \approx c^{N-1}N^{1-b-2\nu}. \quad (106)$$

The contribution from the condensate residing on two adjacent sites, taking respectively x and $1-x$ fractions of particles, is:

$$P_2 \approx Nc^{N-2}p(xN)f((1-x)N)K(xN)K((1-x)N)K(|1-2x|N). \quad (107)$$

For fixed x , in the limit $N \rightarrow \infty$, we obtain

$$P_2 \approx c^{N-2}N^{1-2b-3\nu}(x(1-x))^{-b-\nu}|1-2x|^{-\nu}. \quad (108)$$

Clearly $P_2 \sim P_1N^{-b-\nu}$ and since $-b-\nu < -b/2-1$, it vanishes in comparison to P_1 in the thermodynamic limit. Even if we assume that the difference ϵ in occupation numbers of both sites is small ($x \approx 1/2$), we have

$$P_2 \approx c^{N-2}N^{1-2b-2\nu}K(\epsilon), \quad (109)$$

which is still by a factor $\sim N^{-b}$ smaller than P_1 . Performing also the sum over ϵ gives an additional factor not larger than $N^{1-\nu}$. Thus $P_2 \sim P_1N^{1-b-\nu} < P_1N^{-b/2}$ and the ratio P_2/P_1 tends always to zero. The conclusion is therefore that the condensate, if it exists, is always localized at a single site for any b, ν , precisely as for the ZRP.

This localization is, however, not easy to observe in numerical simulations. If the system is initially prepared in a state with a uniform distribution of particles and the simulation starts, after even a very long time one usually sees the condensate occupying more than one site. This is caused by a small probability to decay from the extended to a single-site condensate. Once the condensate becomes localized on a few sites, there are large energetic barriers between condensates of decreasing sizes. On the other hand, if one starts the simulation from the state with only one site occupied by all particles, one observes that this state does not decay to an extended condensate in the course of the simulation.

5.2. Mixed decay law: $K(x) \sim x^{-\nu}$ and $p(m) = e^{U\delta_{m,0}}$

Let us finally briefly discuss the case where $K(x) \sim x^{-\nu}$ follows a power law as in section 5.1 but $p(m) = e^{U\delta_{m,0}}$ is of short-range nature as in (10). Again, we shall first estimate the probabilities P_1, P_2 of having the condensate on one or two sites:

$$P_1 \sim Nc^{N-1}N^{-2\nu}, \quad (110)$$

$$P_2 \sim Nc^{N-2}N^{-2\nu}4^\nu K(\epsilon). \quad (111)$$

One sees that both terms depend on N in the same way and none becomes negligible for $N \rightarrow \infty$. Also higher P_n share the same feature:

$$P_n \sim Nc^{N-n}N^{-2\nu}4^\nu \left[\sum_{\epsilon} K(\epsilon) \right]^{n-2}. \quad (112)$$

This suggests that the extension is determined by the ratio $\sum_x K(x)/c$ (or more precisely: by the inverse of its logarithm) and is constant if this ratio is smaller than one. If

$\sum_x K(x) = \infty$, that is for $\nu \leq 1$, or when $\sum_x K(x)/c > 1$, P_n grows with $n \rightarrow \infty$ and the condensation is impossible. This means that the prefactor in $K(x) \sim x^{-\nu}$ does matter—consider for instance $K(x) = 1/(1+x/a)^\nu$, which behaves as $a^\nu x^{-\nu}$ for large x . When a is small, $\sum_x K(x)$ is large and there is no condensation.

We can get the same prediction by applying naively the fixed-envelope approach, developed in section 4, and forgetting about problems caused by (perhaps) large fluctuations of neighboring occupation numbers. We use equations (92) and (93) with $K(x) \sim x^{-\nu}$ and $p(m) = 0$, because $p(m)$ vanishes now in the condensate. We find that $\ln P_{\text{smooth}} \sim -M^{1/2}$ with $W \sim M^{1/2}$ and $\ln P_{\text{rect}} \sim -\text{const}$ with $W \sim \text{const}$, so that $\ln P_{\text{smooth}} < \ln P_{\text{rect}}$ and the condensate is rectangular and of fixed extension.

6. Conclusions and outlook

We have shown how to tune the shape of the condensate during spontaneous symmetry breaking in mass transport models on a ring topology. The shape of the condensate and the scaling of its extension with the system size are non-universal. They depend on the competition between the ultralocal and local contributions (in occupation number space) to the weight factors whose product over all pairs of sites determines the stationary state. Analytical predictions were possible even for the condensed phase above the critical mass density for which the grand-canonical partition function is no longer convergent; a partition function that factorizes over the condensate and the critical background turns out to be the appropriate approximation scheme, as the excellent agreement with numerical simulations of the shape and the single-site mass distribution has demonstrated.

Some of the results presented in this paper we have already used in [15] for predicting the onset of condensation in case of anisotropic hopping in two dimensions, where the two-dimensional system could be dimensionally reduced to an effectively one-dimensional ZRP. Moreover, in [15] we generalized the topology on which certain classes of hopping rates lead to PFSS from a one-dimensional ring to arbitrary connected and undirected graphs. From the theoretical point of view it remains challenging to derive the phase structure from these known PFSS in higher dimensions that cannot be reduced to effectively one-dimensional processes.

Vice versa, from experimental observations of the shape and the scaling of the width with the system size, one may trace back the class of hopping interactions that are compatible with the observations. When atoms condense on a crystal surface, they can migrate and build extended islands. As experiments on the deposition of clusters [17] or fabrication of quantum dots [18] show, the islands can be extended in the direction perpendicular to the surface. Currently it is open as to whether hopping rates leading to PFSS are able to reproduce the shape and the typical size of the islands of atoms obtained in such experiments.

Acknowledgments

BW and WJ thank the EU RTN Network ENRAGE under grant No. MRTN-CT-2004-005616 and the Alexander von Humboldt Foundation's grant No. 3.4-Fokoop-DEU/1117877 for support. BW would like to thank the International Center for Transdisciplinary Studies (ICTS) at Jacobs University for its hospitality and support of several visits during this collaboration.

Appendix A

We will show that

$$Q \equiv \frac{1}{v} - \frac{v \tilde{K}'(v)}{2 \tilde{K}(v)} \frac{1}{\int_0^v (x \tilde{K}'(x) / \tilde{K}(x)) dx}, \tag{A.1}$$

entering (56) is never equal to zero. From the definition of $\tilde{K}(x)$,

$$\tilde{K}(x) = \sum_{d=-\infty}^{\infty} K(|d|) e^{dx}, \tag{A.2}$$

we see that $\tilde{K}'(x) + \tilde{K}(0) > \tilde{K}(x)$ for $x > 0$. This in turn means that $\tilde{K}'(x)/\tilde{K}(x)$ grows with x . Therefore, $x\tilde{K}'(x)/\tilde{K}(x)$ is convex for $x > 0$. The area under a convex function $f(x)$ over a range $(0, v)$ is smaller than the area of a trapezoid whose parallel edges are placed at $x = 0, v$ and range from zero to $f(0), f(v)$, respectively. This leads to the following inequality:

$$\int_0^v \frac{x \tilde{K}'(x)}{\tilde{K}(x)} dx < \frac{v^2 \tilde{K}'(v)}{2 \tilde{K}(v)}, \tag{A.3}$$

and finally gives

$$Q < \frac{1}{v} - \frac{v \tilde{K}'(v)}{2 \tilde{K}(v)} \frac{1}{(v^2/2)(\tilde{K}'(v)/\tilde{K}(v))} = 0, \tag{A.4}$$

so that $Q < 0$ for all $v > 0$.

Appendix B

We shall evaluate the formulae (66) and (67) in order to find the variance $\text{var}(m_n)_W$ from (65). They can be simplified if one observes that by differentiating the saddle point equation with respect to u_j one obtains:

$$\frac{\partial^2 F}{\partial z^2} \frac{\partial z}{\partial u_j} + \frac{\partial^2 F}{\partial z \partial v} \frac{\partial v}{\partial u_j} = - \frac{\partial^2 F}{\partial z \partial u_j}, \tag{B.1}$$

$$\frac{\partial^2 F}{\partial v^2} \frac{\partial v}{\partial u_j} + \frac{\partial^2 F}{\partial z \partial v} \frac{\partial z}{\partial u_j} = - \frac{\partial^2 F}{\partial v \partial u_j}. \tag{B.2}$$

One can insert equations (B.1) and (B.2) into (66) and (67), and rewrite (65) as

$$\text{var}(m_n)_W = \sum_{k=1}^n \frac{\partial^2 F}{\partial u_k^2} + \sum_{k=1}^n \sum_{j=1}^n \left(\frac{\partial^2 F}{\partial z \partial u_k} \frac{\partial z}{\partial u_j} + \frac{\partial^2 F}{\partial v \partial u_k} \frac{\partial v}{\partial u_j} \right). \tag{B.3}$$

This formula is far less complicated, but still needs derivatives of z, v taken at the saddle point with $\vec{u} = 0$. Instead of solving the saddle point equation with $\vec{u} \neq 0$ directly, and differentiating the solution over u_j , we can use the following trick: we will calculate the derivatives of F and obtain $\partial z/\partial u_j, \partial v/\partial u_j$ from equations (B.1) and (B.2). Let us

calculate first:

$$\frac{\partial^2 F}{\partial u_k^2} = c(k), \quad (\text{B.4})$$

$$\frac{\partial^2 F}{\partial z \partial u_k} = -\frac{k}{z} c(k), \quad (\text{B.5})$$

$$\frac{\partial^2 F}{\partial v \partial u_k} = \frac{1}{v} c(k), \quad (\text{B.6})$$

$$\frac{\partial^2 F}{\partial z^2} = \frac{1}{z^2} \sum_k k^2 c(k) \cong \frac{1}{z^2} \frac{W^3}{8} (I_0(1) + I_2(1)), \quad (\text{B.7})$$

$$\frac{\partial^2 F}{\partial v^2} = \frac{1}{v^2} \sum_k c(k) \cong \frac{1}{v^2} \frac{W}{2} I_0(1), \quad (\text{B.8})$$

$$\frac{\partial^2 F}{\partial z \partial v} = -\frac{1}{zv} \sum_k k c(k) \cong -\frac{1}{zv} \frac{W^2}{4} I_0(1), \quad (\text{B.9})$$

where we defined

$$c(k) \equiv \frac{\cosh J \cosh [v(1 - (2k/W))] - 1}{[\cosh J - \cosh (v(1 - (2k/W)))]^2}, \quad (\text{B.10})$$

$$I_m(t) \equiv \int_{-1}^t y^m \frac{\cosh J \cosh(vy) - 1}{(\cosh J - \cosh(vy))^2} dy. \quad (\text{B.11})$$

From (B.1) and (B.2) one sees that it must be also $(\partial z / \partial u_j) \sim c(j)$ and $(\partial v / \partial u_j) \sim c(j)$. The proportionality factors in these formulas must be functions of type $A + Bj$, because the only dependence on j in equations (B.1) and (B.2), after dividing by $c(j)$, is linear in j . Inserting $(\partial z / \partial u_j) = (A_z + B_z j)c(j)$ and $(\partial v / \partial u_j) = (A_v + B_v j)c(j)$ into (B.1) and (B.2) and solving for A_v, A_z, B_v, B_z , we obtain

$$\frac{\partial z}{\partial u_j} = -z \frac{4}{I_2(1)W^2} \left(1 - \frac{2j}{W}\right) c(j), \quad (\text{B.12})$$

$$\frac{\partial v}{\partial u_j} = v \frac{2}{I_2(1)W} \left(-1 - \frac{I_2(1)}{I_0(1)} + \frac{2j}{W}\right) c(j). \quad (\text{B.13})$$

Inserting this into (B.3), we finally obtain the variance for fixed width W :

$$\text{var}(m_n)_W = \frac{w\sqrt{M'}}{2} \sigma^2 \left(\frac{2n}{w\sqrt{M'}} - 1\right), \quad (\text{B.14})$$

where we defined a size-independent variance σ^2 as follows:

$$\sigma^2(y) = I_0(y) - \frac{I_0^2(y)}{I_0(1)} - \frac{I_1^2(y)}{I_2(1)}. \quad (\text{B.15})$$

This is precisely equation (69).

Note added in proof. After completing this work we learned about the reference [19], in which a partition function very similar to equation (39) was already evaluated and some of the subsequent formulas were derived.

References

- [1] Chowdhury D, Santen L and Schadschneider A, 2006 *Phys. Rep.* **329** 199
- [2] Liu C-H, Nagel S R, Schechter D A, Coppersmith S N, Majumdar S, Narayan O and Witten T A, 1995 *Science* **269** 513
Coppersmith S N, Liu C-H, Majumdar S, Narayan O and Witten T A, 1996 *Phys. Rev. E* **53** 4673
- [3] Majumdar S N, Krishnamurthy S and Barma M, 1998 *Phys. Rev. Lett.* **81** 3691
Majumdar S N, Krishnamurthy S and Barma M, 2000 *J. Stat. Phys.* **99** 1
- [4] Krapivsky P L, Redner S and Leyvraz F, 2000 *Phys. Rev. Lett.* **85** 4629
- [5] Bialas P, Burda Z, Petersson B and Tabaczek J, 1997 *Nucl. Phys. B* **495** 463
de Bakker B V, 1996 *Phys. Lett. B* **389** 238
- [6] Bialas P, Burda Z and Johnston D A, 1997 *Nucl. Phys. B* **493** 505
- [7] Evans M R, 2000 *Braz. J. Phys.* **30** 42
- [8] Evans M R, Majumdar S N and Zia R K P, 2004 *J. Phys. A: Math. Gen.* **37** L275
- [9] Evans M R, Majumdar S N and Zia R K P, 2006 *J. Phys. A: Math. Gen.* **39** 4859
- [10] Evans M R, Hanney T and Majumdar S N, 2006 *Phys. Rev. Lett.* **97** 010602
- [11] Chui S T and Weeks J D, 1981 *Phys. Rev. B* **23** 2438
Burkhardt T W, 1981 *J. Phys. A: Math. Gen.* **14** L63
- [12] van Leeuwen J M L and Hilhorst H J, 1981 *Physica A* **107** 319
- [13] Chechkin A V, Metzler R, Gonchar V Y, Klafter J and Tanatarow L V, 2003 *J. Phys. A: Math. Gen.* **36** L537
- [14] Waclaw B, Sopik J, Janke W and Meyer-Ortmanns H, 2009 *Phys. Rev. Lett.* **103** 080602
- [15] Waclaw B, Sopik J, Janke W and Meyer-Ortmanns H, 2009 *J. Phys. A: Math. Theor.* **42** 315003
- [16] Majumdar S N, 2005 *Curr. Sci.* **89** 2076 <http://www.ias.ac.in/currsci/dec252005/2076.pdf>
- [17] Bréchnignac C, Cahuzac P, Carlier F, Colliex C, Leroux J, Masson A, Yoon B and Landman U, 2002 *Phys. Rev. Lett.* **88** 196103
Yoon B, Akulin V M, Cahuzac P, Carlier F, de Frutos M, Masson A, Mory C, Colliex C and Bréchnignac C, 1999 *Surf. Sci.* **443** 76
- [18] Nötzel R, 1996 *Semicond. Sci. Technol.* **11** 1365
- [19] De Coninck J, Dunlop F and Rivasseau V, 1989 *Commun. Math. Phys.* **121** 40