MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

SFG ZÁVĚREČNÁ ZPRÁVA

František Fröde

Fázové přechody v souborech podmíněných součtů náhodných proměnných

Katedra Makromolekulární Fyziky

Vedoucí projektu: RNDr. Artem Ryabov, Ph.D.
Studijní program: Matematické modelování
Studijní obor: Matematické modelování

Praha 2023

FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

SFG PROJECT REPORT

František Fröde

Phase transitions in systems of random variables conditioned by sum

Department of Macromolecular Physics

Supervisor of the project: RNDr. Artem Ryabov, Ph.D.
Study programme: Mathematical modelling
Study branch: Mathematical modelling

Prague 2023

Contents

Introduction

Modelling microscopic transport processes far from thermodynamic equilibrium lies at the heart of non-equilibrial statistical physics. Important tools are lattice gas models, in which particles hop between lattice sites based on predetermined rules. One class of such models is the zero-range process (ZRP), which can be studied analytically.

An exhaustive summary of ZRP is provided by Evans and Hanney [2005]. The most basic ZRP model involves a one-dimensional lattice with periodic boundary conditions, populated by a fixed number of particles. The particles hop to the site on their right, with the hop rate depending only on the number of particles at the current site. This rule is restrictive enough, to allow analytical study of such a model; z-transform links the overall probability distribution of the particles with the transformed single-site probability weights. Using this transform, it may be shown, that in some cases, there exists a critical density, above which the behaviour of the model changes. Using other means, one can show, that in such cases, most of the particles accumulate at a single site.

The properties of this transform are understood for a limited class of models, for which the condensation is comparable to an equivalent phenomenon arising when independent identically distributed random variables are conditioned on their sum. The goal of this work is to establish a similar link between ZRP generalized to multiple particle types and discrete random vectors.

1. Systems of random variables conditioned by sum

Let us consider L independent identically distributed random variables $\{x_l\}_{l=1}^{L}$ with non-negative integer values. Let $f(n)$ denote the probability of a given random variable having value n. Let us further condition the sum of the random variables to:

$$\sum_{l=1}^{L} x_l = N \tag{1.1}$$

The probability of a state $x_1 = n_1 \wedge \cdots \wedge x_L = n_L$ is:

$$\begin{aligned}
\mathcal{P}(n_1; \ldots; n_L) &:= \mathcal{P}\left[\bigwedge_{l=1}^{L} x_l = n_l | \sum_{l=1}^{L} x_l = N\right] \\
&= \mathcal{P}\left[\bigwedge_{l=1}^{L} x_l = n_l \wedge \sum_{l=1}^{L} x_l = N\right] / \mathcal{P}\left[\sum_{l=1}^{L} x_l = N\right] \\
&= \left(\prod_{l=1}^{L} f(n_l)\right) \delta\left(\sum_{l=1}^{L} n_l = N\right) / \mathcal{P}\left[\sum_{l=1}^{L} x_l = N\right] \\
&=: \left(\prod_{l=1}^{L} f(n_l)\right) \delta\left(\sum_{l=1}^{L} n_l = N\right) / Z_L(N)
\end{aligned} \tag{1.2}$$

where δ is defined as 1 if its argument is true, and 0 otherwise. The denominator can be calculated recursively. For $L = 1$, equation (1.2) directly yields $Z_1(N) = f(N)$. For larger values of L, multiplying (1.2) by $Z_L(N)$ and summing over all possibilities leads to:

$$\begin{aligned}
Z_L(N) &= Z_L(N) \sum_{n_1=0}^{N} \cdots \sum_{n_L=0}^{N} \mathcal{P}(n_1; \ldots; n_L) \\
&= \sum_{n_1=0}^{N} \cdots \sum_{n_L=0}^{N} \delta\left(\sum_{l=1}^{L} n_l = N\right) \prod_{l=1}^{L} f(n_l) \\
&= \sum_{n_1=0}^{N} f(n_1) \sum_{n_2=0}^{N} \cdots \sum_{n_L=0}^{N} \delta\left(\sum_{l=2}^{L} n_l = N - n_1\right) \prod_{l=2}^{L} f(n_l) \\
&= \sum_{n=0}^{N} f(n) Z_{L-1}(N - n) \\
&=: (f * Z_{L-1})(N)
\end{aligned} \tag{1.3}$$

This recurrence can be generalised to a system of random vectors, that is, $\{\vec{x}_l\}_{l=1}^{L} \subset \mathbb{N}_0^K$ with their sum conditioned to be $\vec{N}$. Just as before, the probability of a given state is:

$$\mathcal{P}(\vec{n}_1; \ldots; \vec{n}_L) = \left(\prod_{l=1}^{L} f(\vec{n}_l)\right) \delta\left(\sum_{l=1}^{L} \vec{n}_l = \vec{N}\right) / Z_L(\vec{N}) \tag{1.4}$$

with a notable complication in how $Z_L(\vec{N})$ is computed – we must now sum over all vectors, such that each of their components is less than or equal to the corresponding component of $\vec{N}$. If we introduce the notation $\sum_{\vec{n}}^{\vec{N}}$ for such summation, the formula for Z_L can be written as:

$$
\begin{aligned}
Z_L(\vec{N}) &= Z_L(\vec{N}) \sum_{\vec{n}_1}^{\vec{N}} \cdots \sum_{\vec{n}_L}^{\vec{N}} \mathcal{P}(\vec{n}_1; \ldots; \vec{n}_L) \\
&= \sum_{\vec{n}_1}^{\vec{N}} \cdots \sum_{\vec{n}_L}^{\vec{N}} \delta\left(\sum_{l=1}^{L} \vec{n}_l = \vec{N}\right) \prod_{l=1}^{L} f(\vec{n}_l) \\
&= \sum_{\vec{n}_1}^{\vec{N}} f(\vec{n}_1) \sum_{\vec{n}_2}^{\vec{N}} \cdots \sum_{\vec{n}_L}^{\vec{N}} \delta\left(\sum_{l=2}^{L} \vec{n}_l = \vec{N} - \vec{n}_1\right) \prod_{l=2}^{L} f(\vec{n}_l) \\
&= \sum_{\vec{n}}^{\vec{N}} f(\vec{n}) Z_{L-1}(\vec{N} - \vec{n}) \\
&=: (f * Z_{L-1})(\vec{N})
\end{aligned}
\tag{1.5}
$$

Thanks to associativity of discrete convolution (see appendix 4.1 for a proof), the reccurences (1.3) and (1.5) can be further refined into:

$$
Z_{2L} = f * Z_{2L-1} = (f * f) * Z_{2L-2} = \cdots = \underbrace{(f * \cdots * f)}_{L} * Z_{2L-L} = Z_L * Z_L \tag{1.6}
$$

This formula is particularly useful for evaluating Z_L numerically.

As we shall see later, the equations for the probabilities (1.2) and (1.4) are identical to the factorisations of the ZRP steady-state probabilities (2.2) and (3.2) respectively.

1.1 Distribution of the maximum

In the one-dimensional case, the probability of the maximum being less than or equal to t is given as:

$$
\begin{aligned}
\mathcal{P}\left[\max_{l \in \{1,\ldots,L\}} x_l \leq t\right] &= \sum_{n_1=0}^{t} \cdots \sum_{n_L=0}^{t} \mathcal{P}(n_1; \ldots; n_L) \\
&= \sum_{n_1=0}^{t} \cdots \sum_{n_L=0}^{t} \delta\left(\sum_{l=1}^{L} n_l = N\right) Z_L(N)^{-1} \prod_{l=1}^{L} f(n_l) \\
&= Z_L(N)^{-1} \sum_{n_1=0}^{N} \cdots \sum_{n_L=0}^{N} \delta\left(\sum_{l=1}^{L} n_l = N\right) \prod_{l=1}^{L} f_t(n_l) \\
&= Z_L(N)^{-1} \cdot Z_{t,L}(N)
\end{aligned}
\tag{1.7}
$$

where

$$
f_t(n) := \begin{cases} f(n) & n \leq t \\ 0 & \text{otherwise} \end{cases} \tag{1.8}
$$

and $Z_{t,L}$ the result of (1.3) with f replaced by f_t.

2. Single particle type fully asymmetric ZRP

An overview of ZRP can be found in Evans and Hanney [2005].

The fully asymmetric zero range process (ZRP) is described as follows. There is a periodic 1D-lattice of L sites (site 1 follows site L). Among those sites, there is a total of N identical particles. Each site may contain an arbitrary number n_l of particles. The particles hop to the next site with a "hop rate" $u(n_l)$ depending only on the number of particles present at the departure site.

An advantage of the ZRP is, that the steady-state probability distribution factorises into single-site weights. Let $\mathcal{P}(n_1;\ldots;n_L)$ denote the probability of site 1 containing n_1 praticles, site 2 n_2, etc. In the steady state, the probability must satisfy the equation:

$$\begin{aligned}0 = \mathcal{J} = \sum_{l=1}^{L} &\mathcal{P}(n_1;\ldots;n_l+1;n_{l+1}-1;\ldots;n_L)u(n_l+1) \\ &- \mathcal{P}(n_1;\ldots;n_l;n_{l+1};\ldots;n_L)u(n_l)\end{aligned} \tag{2.1}$$

Writing $\mathcal{P}$ in the form:

$$\mathcal{P}(n_1;\ldots;n_L) = Z_L(N)^{-1}\left(\prod_{l=1}^{L} f(n_l)\right)\delta\left(\sum_{l=1}^{L} n_l = N\right) \tag{2.2}$$

where

$$f(n) = c_0 \prod_{\nu=1}^{n} \frac{c_1}{u(\nu)} \tag{2.3}$$

and $Z_L(N)$ is a normalisation constant satisfies (2.1); after cancelling common terms, each term of the sum evaluates to:

$$\begin{aligned}&f(n_l+1)f(n_{l+1}-1)u(n_l+1) - f(n_l)f(n_{l+1})u(n_l) \\ &= f(n_l)f(n_{l+1}-1) - f(n_l)f(n_{l+1}-1) \\ &= 0\end{aligned} \tag{2.4}$$

and the entire sum is thus zero. Note, that if f is normalised (with the help of c_0), such that $\sum_{n=0}^{N} f(n) = 1$, equation (2.2) is identical to (1.2), and $Z_L(N)$ can be calculated using (1.3). For other values of c_0, $Z_L(N)$ can be calculated using the same recurrence, and more, the value of $\mathcal{P}$ does not depend on c_0 and c_1 (for proof, see appendix 4.2). As this is the case, the form of f in literature usually corresponds to $c_0 = c_1 = 1$.

2.1 Condensation

It is a well-known property of ZRP, that for certain functional forms of u, condensation can occur. When taking the limit $N, L \to \infty, N/L \to \rho$, if the density ρ exceeds a critical density ρ_c, particles may accumulate at a single site. This behaviour is discussed in Evans and Hanney [2005] and subsequent references.

The critical density ρ_c can be calculated by applying a z-transform. Let

$$\mathcal{Z}_L(z) := \sum_{n=0}^{\infty} z^n Z_L(n) \tag{2.5}$$

$$F(z) := \sum_{n=0}^{\infty} z^n f(n) \tag{2.6}$$

In case the first sum has a non-zero radius of convergence, $Z_L(N)$ can be recovered from $\mathcal{Z}$ using residue theorem:

$$Z_L(N) = \oint \frac{\mathrm{d}z}{2\pi i} z^{-1-N} \mathcal{Z}_L(z) \tag{2.7}$$

where the contour integral is around $z = 0$. Furthermore,

$$\begin{aligned} F(z)^L &= \left(\sum_{n=0}^{\infty} z^n f(n) \right)^L \\ &= \sum_{n_1=0}^{\infty} \cdots \sum_{n_L=0}^{\infty} \prod_{l=1}^{L} z^{n_l} f(n_l) \\ &= \sum_{n_1=0}^{\infty} \cdots \sum_{n_L=0}^{\infty} \sum_{N=0}^{\infty} \delta \left(\sum_{l=1}^{L} n_l = N \right) z^N \prod_{l=1}^{L} f(n_l) \\ &= \sum_{N=0}^{\infty} \delta \left(\sum_{l=1}^{L} n_l = N \right) z^N \sum_{n_1=0}^{N} \cdots \sum_{n_L=0}^{N} \prod_{l=1}^{L} f(n_l) \\ &= \sum_{N=0}^{\infty} z^N Z_L(N) \\ &= \mathcal{Z}_L(z) \end{aligned} \tag{2.8}$$

which allows us to rewrite (2.7) to

$$Z_L(N) = \oint \frac{dz}{2\pi i} z^{-1-N} \mathcal{Z}_L(z) = \oint \frac{dz}{2\pi i} z^{-1-N} F(z)^L \tag{2.9}$$

For large values of L, N we get:

$$\begin{aligned} Z_L(N) &= \oint \frac{dz}{2\pi i} \left[z^{-\frac{1-N}{L}} F(z) \right]^L \approx \oint \frac{dz}{2\pi i} \left[z^{-\rho} F(z) \right]^L \\ &= \oint \frac{dz}{2\pi i} e^{L(-\rho \ln z + \ln F(z))} =: \oint \frac{dz}{2\pi i} e^{L\phi(z)} \end{aligned} \tag{2.10}$$

If ϕ has a saddle point s within the radius of convergence of F, it is possible to approximate the integral as:

$$Z_L(N) \approx \frac{\mathcal{Z}_L(s)}{s^{N+1} \sqrt{2\pi L |\phi''(s)|}} \tag{2.11}$$

The condition for the saddle point is:

$$0 = \phi'(s) = -\frac{\rho}{s} + \frac{F'(s)}{F(s)} \implies \frac{sF'(s)}{F(s)} = \rho \tag{2.12}$$

Taking the s-derivative, we get:

$$\left(\frac{sF'(s)}{F(s)}\right)' = \frac{F'(s)F(s) + zF''(s)F(s) - zF'(s)^2}{F(s)^2} \tag{2.13}$$

with the denominator being positive for real values of s.

$$\begin{aligned} & F'(s)F(s) + zF''(s)F(s) - zF'(s)^2 \\ &= \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} s^{a+b-1} f(a)f(b)[a + a(a-1) - ab] \\ &= \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} s^{a+b-1} f(a)f(b)(a^2 - ab) \end{aligned} \tag{2.14}$$

The terms for which $a = b$ are all zero (including the only term containing s^{-1}). For the other terms, we can sum terms $a = \tilde{a}, b = \tilde{b}$ and $a = \tilde{b}, b = \tilde{a}$ leading to:

$$s^{\tilde{a}+\tilde{b}-1} f(\tilde{a})f(\tilde{b})(\tilde{a}^2 - 2\tilde{a}\tilde{b} + \tilde{b}^2) = s^{\tilde{a}+\tilde{b}-1} f(\tilde{a})f(\tilde{b})(\tilde{a} - \tilde{b})^2 \tag{2.15}$$

If we restrict s to be real and positive, the above quantity is positive, leading to $sF'(s)/F(s)$ being increasing in s. Excluding the case $f(0) = 0$,

$$\lim_{s \to 0^+} \frac{sF'(s)}{F(s)} = 0 \tag{2.16}$$

Let β denote the radius of convergence of F and let us define

$$\rho_c := \lim_{s \to \beta^-} \frac{sF'(s)}{F(s)} \tag{2.17}$$

We see, that for $\rho \in [0, \rho_c)$ it is possible to invert (2.12), which provides us with a saddle point. For $\rho \geq \rho_c$ a different approach to computing $Z_L(N)$ needs to be taken.

Note, that the only restriction for f we required was $f(0) \neq 0$. For ZRP, (2.3) leads to $f(n) \neq 0$ for all values of n. If we return to independent random variables discussed in the first section, f can be chosen more freely. Let us consider the case $f(n) = 0$ for all $n > m$, and $f(m) \neq 0$. The series in (2.6) effectively becomes a finite sum, and as such has an infinite radius of convergence. The critical density is in this case

$$\rho_c = \lim_{s \to \infty} \frac{sF'(s)}{F(s)} = \lim_{s \to \infty} \frac{sf(m)(s^m)'}{f(m)s^m} = m \tag{2.18}$$

which happens iff (almost) all sites are fully saturated.

3. ZRP generalized to multiple particle types

Let us consider a generalization of ZRP with multiple (K) particle types. We will show, that under certain conditions for the hop rates, this generalized model also allows for factorization into single-site weights.

The number of particles at individual sites can be represented by K-dimensional vectors $\{\vec{x}_l\}_{l=0}^{L} \subset \mathbb{N}_0^K$. Let $x_{l,k}$ denote the number of particles of the k-th type at site l, $u_k(\vec{x})$ the hop-rate of k-th particle type when the total numbers of particles at a given site are $\vec{x}$, and $\vec{e}_k$ a vector with 1 as k-th component and 0 as the other components. The steady state condition (an analogue of (2.1)) now becomes:

$$\begin{aligned} 0 = \mathcal{J} = \sum_{l=1}^{L}\sum_{k=1}^{K} & \mathcal{P}(\vec{n}_1;\dots;\vec{n}_l+\vec{e}_k;\vec{n}_{l+1}-\vec{e}_k;\dots;\vec{n}_L)u_k(\vec{n}_l+\vec{e}_k) \\ & - \mathcal{P}(\vec{n}_1;\dots;\vec{n}_l;\vec{n}_{l+1};\dots;\vec{n}_L)u_k(\vec{n}_l) \end{aligned} \tag{3.1}$$

Let us assume it is possible to factorise $\mathcal{P}$ as:

$$\mathcal{P}(\vec{n}_1;\dots;\vec{n}_L) = Z_L(\vec{N})^{-1}\left(\prod_{l=1}^{L} f(\vec{n}_l)\right)\delta\left(\sum_{l=1}^{L}\vec{n}_l = \vec{N}\right) \tag{3.2}$$

If we further assume, that $f(\vec{n}) > 0$ for all values of n, equation (3.1) can be rewritten to (after factoring out the common factor $\mathcal{P}(\vec{n}_1;\dots;\vec{n}_L)$):

$$0 = \sum_{l=1}^{L}\sum_{k=1}^{K}\frac{f(\vec{n}_l+\vec{e}_k)}{f(\vec{n}_l)}\frac{f(\vec{n}_{l+1}-\vec{e}_k)}{f(\vec{n}_{l+1})}u_k(\vec{n}_l+\vec{e}_k) - u_k(\vec{n}_l) \tag{3.3}$$

The steady state condition has to hold for any configuration $\{\vec{n}_l\}_{l=1}^{L}$, special case beiing:

$$\vec{n}_l = \begin{cases} \vec{n}_2 & l = 2 \\ 0 & \text{otherwise} \end{cases} \tag{3.4}$$

Since if $n_{l,k} = 0$ then $u_k(\vec{n}_l) = 0$ (a particle may only hop from a site where there are particles), and if $n_{l,k} < 0$ then $f(\vec{n}_l) = 0$, equation (3.3) simplifies to:

$$0 = \sum_{k=1}^{K}\frac{f(\vec{e}_k)}{f(\vec{0})}\frac{f(\vec{n}_2-\vec{e}_k)}{f(\vec{n}_2)}u_k(\vec{e}_k) - u_k(\vec{n}_2) \tag{3.5}$$

We may examine the above case further, but it will be more useful to study a different special case:

$$\vec{n}_l = \begin{cases} \vec{n}_2 & l = 2 \\ \vec{n}_3 & l = 3 \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

Equation (3.3) now becomes:

$$\begin{aligned} 0 = \sum_{k=1}^{K} & \frac{f(\vec{e}_k)}{f(\vec{0})}\frac{f(\vec{n}_2-\vec{e}_k)}{f(\vec{n}_2)}u_k(\vec{e}_k) - u_k(\vec{n}_2) \\ & + \frac{f(\vec{n}_2+\vec{e}_k)}{f(\vec{n}_2)}\frac{f(\vec{n}_3-\vec{e}_k)}{f(\vec{n}_3)}u_k(\vec{n}_2+\vec{e}_k) - u_k(\vec{n}_3) \end{aligned} \tag{3.7}$$

The first part being zero as per above further simplifies this into:

$$0 = \sum_{k=1}^{K} \frac{f(\vec{n}_2 + \vec{e}_k)}{f(\vec{n}_2)} \frac{f(\vec{n}_3 - \vec{e}_k)}{f(\vec{n}_3)} u_k(\vec{n}_2 + \vec{e}_k) - u_k(\vec{n}_3) \tag{3.8}$$

This again has to hold for all values of $\vec{n}_3$. Fixing it to $\vec{n}_3 = \vec{e}_j$ results in:

$$0 = \frac{f(\vec{n}_2 + \vec{e}_j)}{f(\vec{n}_2)} \frac{f(\vec{0})}{f(\vec{e}_j)} u_j(\vec{n}_2 + \vec{e}_j) - u_j(\vec{e}_j) \tag{3.9}$$

which after substituting $\vec{n}_l$ for $\vec{n}_2 + \vec{e}_j$ gives us a recursive formula for $f(\vec{n})$:

$$f(\vec{n}_l) = f(\vec{n}_l - \vec{e}_j) \frac{f(\vec{e}_j)}{f(\vec{0})} \frac{u_j(\vec{e}_j)}{u_j(\vec{n}_l)} \tag{3.10}$$

Using this equation to find the relation between $f(\vec{n}_l)$ and $f(\vec{n}_l - \vec{e}_i - \vec{e}_j)$ imposes certain restriction to the hop rates u_i, u_j: if $\vec{e}_i$ is subtracted first, we get:

$$f(\vec{n}_l) = f(\vec{n}_l - \vec{e}_i) \frac{f(\vec{e}_i)}{f(\vec{0})} \frac{u_i(\vec{e}_i)}{u_i(\vec{n}_l)} = f(\vec{n}_l - \vec{e}_i - \vec{e}_j) \frac{f(\vec{e}_j)}{f(\vec{0})} \frac{u_j(\vec{e}_j)}{u_j(\vec{n}_l - \vec{e}_i)} \frac{f(\vec{e}_i)}{f(\vec{0})} \frac{u_i(\vec{e}_i)}{u_i(\vec{n}_l)} \tag{3.11}$$

whereas subtracting $\vec{e}_j$ first leads to:

$$f(\vec{n}_l) = f(\vec{n}_l - \vec{e}_i - \vec{e}_j) \frac{f(\vec{e}_i)}{f(\vec{0})} \frac{u_i(\vec{e}_i)}{u_i(\vec{n}_l - \vec{e}_j)} \frac{f(\vec{e}_j)}{f(\vec{0})} \frac{u_j(\vec{e}_j)}{u_j(\vec{n}_l)} \tag{3.12}$$

Since the order, in which individual terms are decreased shouldn't matter, setting the right-hand sides equal and cancelling common terms leads to:

$$\begin{aligned} \frac{1}{u_j(\vec{n}_l - \vec{e}_i)} \frac{1}{u_i(\vec{n}_l)} &= \frac{1}{u_i(\vec{n}_l - \vec{e}_j)} \frac{1}{u_j(\vec{n}_l)} \\ \frac{u_j(\vec{n}_l)}{u_j(\vec{n}_l - \vec{e}_i)} &= \frac{u_i(\vec{n}_l)}{u_i(\vec{n}_l - \vec{e}_j)} \end{aligned} \tag{3.13}$$

Note, that the condition is automatically fulfilled for $i = j$.

Equation (3.10) can be iterated to obtain:

$$\begin{aligned} f(\vec{n}_l) &= f(\vec{n}_l - n_{l,j}\vec{e}_j) \prod_{\nu_j=1}^{n_{l,j}} \frac{f(\vec{e}_j)}{f(\vec{0})} \frac{u_j(\vec{e}_j)}{u_j(\vec{n}_l - (\nu_j - 1)\vec{e}_j)} \\ &= f(\vec{n}_l - n_{l,j}\vec{e}_j) \left(\frac{f(\vec{e}_j) u_j(\vec{e}_j)}{f(\vec{0})} \right)^{n_{l,j}} \prod_{\nu_j=1}^{n_{l,j}} \frac{1}{u_j(\nu_j\vec{e}_j + \vec{n}_l - n_{l,j}\vec{e}_j)} \\ &= f(\vec{n}_l - n_{l,j}\vec{e}_j) c_j^{n_{l,j}} \prod_{\nu_j=1}^{n_{l,j}} \frac{1}{u_j(\nu_j\vec{e}_j + \vec{n}_l - n_{l,j}\vec{e}_j)} \end{aligned} \tag{3.14}$$

The value of $\mathcal{P}$ calculated from those rates does not depend on the value of c_j. (For proof, see appendix 4.2). For brevity, we will set its value to $c_j = 1$. By repeatedly using (3.14) to eliminate all components of $\vec{n}_l$, $f(\vec{n}_l)$ can be written in terms of hop rates u_j and $f(\vec{0})$ only:

$$f(\vec{n}_l) = f(\vec{0}) \prod_{k=1}^{K} \left(\prod_{\nu_k=1}^{n_{l,k}} \frac{1}{u_k(\nu_k\vec{e}_k + \vec{n}_l - \sum_{j=1}^{k} n_{l,j}\vec{e}_j)} \right) \tag{3.15}$$

Now, that we inferred the form of $f(\vec{n}_l)$ from special cases, we need to check that the steady state condition (3.3) is satisfied. Substiting for $f(\vec{n}_l + \vec{e}_k)$ and $f(\vec{n}_{l+1})$ from (3.10), the corresponding term of the sum evaluates to:

$$\begin{aligned}
&\frac{f(\vec{n}_l + \vec{e}_k)}{f(\vec{n}_l)} \frac{f(\vec{n}_{l+1} - \vec{e}_k)}{f(\vec{n}_{l+1})} u_k(\vec{n}_l + \vec{e}_k) - u_k(\vec{n}_l) \\
&= \frac{f(\vec{n}_l) \frac{f(\vec{e}_k)}{f(\vec{0})} \frac{u_k(\vec{e}_k)}{u_k(\vec{n}_l + \vec{e}_k)}}{f(\vec{n}_l)} \frac{f(\vec{n}_{l+1} - \vec{e}_k)}{f(\vec{n}_{l+1} - \vec{e}_k) \frac{f(\vec{e}_k)}{f(\vec{0})} \frac{u_k(\vec{e}_k)}{u_k(\vec{n}_{l+1})}} u_k(\vec{n}_l + \vec{e}_k) - u_k(\vec{n}_l) \\
&= u_k(\vec{n}_{l+1}) - u_k(\vec{n}_l)
\end{aligned} \tag{3.16}$$

Since the sum is over all sites l, it evaluates to 0.

3.1 Connection to random vectors

The factorized form (3.2) is identical to the probability (1.4) a collection of independent random vectors $\{\vec{x}_l\}_{l=1}^{L} \subset \mathbb{N}_0^K$ sharing the probability density f[1] takes on the corresponding values $\{\vec{n}_l\}_{l=1}^{L}$. In other words, the probability the many-particle-type ZRP is in a given state $\mathcal{P}(\vec{n}_1; \ldots; \vec{n}_L)$ can be thought of as the probability of $\vec{x}_l = \vec{n}_l$ for all values of l, provided, that the sum of the random vectors is conditioned to $\sum_{l=1}^{L} \vec{x}_l = \vec{N}$.

A natural question to ask is, when is it possible to go in the opposite direction? Let f denote the shared probability density of some discrete random vectors $\{\vec{x}_l\}_{l=1}^{L}$ and let us try to find a multi-particle-type ZRP with the same f as its single site weights. Equation (3.10) provides us with an idea to define the hop rates as:

$$u_k(\vec{n}) = \frac{f(\vec{n} - \vec{e}_k)}{f(\vec{n})} \tag{3.17}$$

which satisfies said equation. This choice immediately requires $f(\vec{n})$ to be positive if all components of $\vec{n}$ are non-negative. If the k-th component of $\vec{n}$ is zero, $u_k(\vec{n})$ evaluates to zero, which is consistent with the idea of particles being able to hop only from occupied sites. The last thing required is to check, whether the factorization condition (3.13) is satisfied:

$$\begin{aligned}
\frac{u_j(\vec{n}_l)}{u_j(\vec{n}_l - \vec{e}_i)} &= \frac{f(\vec{n}_l - \vec{e}_j)}{f(\vec{n}_l)} / \frac{f(\vec{n}_l - \vec{e}_i - \vec{e}_j)}{f(\vec{n}_l - \vec{e}_i)} \\
&= \frac{f(\vec{n}_l - \vec{e}_i)}{f(\vec{n}_l)} / \frac{f(\vec{n}_l - \vec{e}_j - \vec{e}_i)}{f(\vec{n}_l - \vec{e}_j)} \\
&= \frac{u_i(\vec{n}_l)}{u_i(\vec{n}_l - \vec{e}_j)}
\end{aligned} \tag{3.18}$$

As we can see, it is possible to find such ZRP under the condition $f(\vec{n})$ is positive when all components of $\vec{n}$ are non-negative.

[1] By f we mean the single site weights of the ZRP. In general, it would need to be normalized, such that $\sum_{\vec{n}}^{\vec{N}} f(\vec{n}) = 1$. The normalization does not change the properties of the studied model, as per 4.2.

3.2 Condensation in many-particle-type ZRP

Just as in the single-type ZRP, there is a possibility for a phase transition to occur at high enough densities. Unlike before, however, there is no single critical density, but rather a critical curve. The shape of the curve depends on the hop rates u_k.

The normalisation factor $Z_L(\vec{N})$ in the assumed factorisation (3.2) is:

$$Z_L(\vec{N}) = \sum_{\vec{n}_1}^{\vec{N}} \cdots \sum_{\vec{n}_L}^{\vec{N}} \delta\left(\vec{N} = \sum_{l=1}^{L} \vec{n}_l\right) \prod_{l=1}^{L} f(\vec{n}_l) \tag{3.19}$$

In order to study its behaviour for large values of L, we may wish to find an analogue to (2.8). Let us define

$$\vec{z}^{\vec{n}} := \prod_{k=1}^{K} z_k^{n_k} \tag{3.20}$$

$$\mathcal{Z}_L(\vec{z}) := \sum_{\vec{n}\in\mathbb{Z}_0^K} \vec{z}^{\vec{n}} Z_L(\vec{n}) \tag{3.21}$$

$$F(\vec{z}) := \sum_{\vec{n}\in\mathbb{Z}_0^K} \vec{z}^{\vec{n}} f(\vec{n}) \tag{3.22}$$

Similarly to (2.8), it is the case that:

$$\begin{aligned}
F(\vec{z})^L &= \left(\sum_{\vec{n}\in\mathbb{Z}_0^K} \vec{z}^{\vec{n}} f(\vec{n})\right)^L \\
&= \sum_{\vec{n}_1\in\mathbb{Z}_0^K} \cdots \sum_{\vec{n}_L\in\mathbb{Z}_0^K} \prod_{l=1}^{L} \vec{z}^{\vec{n}_l} f(\vec{n}_l) \\
&= \sum_{\vec{n}_1\in\mathbb{Z}_0^K} \cdots \sum_{\vec{n}_L\in\mathbb{Z}_0^K} \sum_{\vec{N}\in\mathbb{Z}_0^K} \delta\left(\vec{N} = \sum_{l=1}^{L} \vec{n}_l\right) \vec{z}^{\vec{N}} \prod_{l=1}^{L} f(\vec{n}_l) \\
&= \sum_{\vec{N}\in\mathbb{Z}_0^K} \vec{z}^{\vec{N}} \sum_{\vec{n}_1\in\mathbb{Z}_0^K} \cdots \sum_{\vec{n}_L\in\mathbb{Z}_0^K} \delta\left(\vec{N} = \sum_{l=1}^{L} \vec{n}_l\right) \prod_{l=1}^{L} f(\vec{n}_l) \\
&= \sum_{\vec{N}\in\mathbb{Z}_0^K} \vec{z}^{\vec{N}} Z_L(\vec{N}) \\
&= \mathcal{Z}_L(\vec{z})
\end{aligned} \tag{3.23}$$

Yet again, $Z_L(\vec{N})$ can be recovered with the help of residue theorem:

$$\begin{aligned}
Z_L(\vec{N}) &= \oint \frac{z_1^{-1-N_1}\mathrm{d}z_1}{2\pi i} \cdots \oint \frac{z_K^{-1-N_K}\mathrm{d}z_K}{2\pi i} \mathcal{Z}_L(\vec{z}) \\
&= \oint \frac{z_1^{-1-N_1}\mathrm{d}z_1}{2\pi i} \cdots \oint \frac{z_K^{-1-N_K}\mathrm{d}z_K}{2\pi i} F(\vec{z})^L
\end{aligned} \tag{3.24}$$

with all the integrals being taken around the origin. Each integral can be approximated by the saddle point method. For the k-th integral, the saddle point condition reads:

$$0 = \frac{\partial}{\partial s_k} \ln(s_k^{-1-N_k} F(\vec{s})^L) = \frac{-1 - N_k}{s_k} + L \frac{\frac{\partial}{\partial s_k} F(\vec{s})}{F(\vec{s})} \tag{3.25}$$

$$s_k \frac{\frac{\partial}{\partial s_k} F(\vec{s})}{F(\vec{s})} = \frac{1}{L} + \frac{N_k}{L} \approx \frac{N_k}{L} = \rho_k \tag{3.26}$$

F can be thought of as a power series in terms of s_k, with non-negative coefficients. Thus by equations (2.14) and the argument thereafter, the saddle point condition can be inverted to obtain the saddle point s_k. Note however, that the radius of convergence, used in the calculation of the critical density for the single-type ZRP, depends on the values of other components of $\vec{s}$.

Conclusion

In this work, we have studied the link between the ZRP model, including its generalisation to multiple particle types, to a system of independent identically distributed random variables conditioned by their sum. Given the connections of ZRP models to other significant models of nonequilibrium statistical physics (such as the totally asymmetric simple exclusion process model, surface growth models or renewal processes), the results may be widely applicable in the study of stochastic systems.

In the first chapter, we have derived a recursive formula (1.6) for calculating the probability the sum of random variables evaluates to a given value. This probability is crucial for the study of systems conditioned by sum.

In the second chapter, we have discussed the factorisation of ZRP into single-site weights and reviewed the procedure for obtaining the critical density for possible condensation.

In the third chapter, we have derived a necessary and sufficient condition for the factorisation of many-particle-type ZRP. We proceeded to illuminate the link between this generalisation of ZRP and a system of independent identically distributed random discrete vectors conditioned by their sum. We have also briefly discussed the calculation of the critical line – an analogue to the critical density.

The invariance discussed in appendix 4.2, which, when applied to ZRP, we believe to have the meaning of individual time scale for different particle types, asks for further research.

4. Appendix

4.1 Appendix 1: Associativity of discrete convolution

Definition 1 (Discrete convolution)**.** *Let $f, g : \mathbb{N}_0^K \to \mathbb{R}$. We define their discrete convolution $f * g$ as:*

$$(f * g)(\vec{N}) := \sum_{\vec{n}}^{\vec{N}} f(\vec{n}) \cdot g(\vec{N} - \vec{n}) \tag{4.1}$$

where $\sum_{\vec{n}}^{\vec{N}}$ is defined as the sum over all vectors $\vec{n} \in \mathbb{N}_0^K$, such that all of $\vec{n}$'s components are less than or equal to the corresponding component of $\vec{N}$. For brevity, we will henceforth omit the arrows.

Using $\delta(prop)$ defined as 1 if $prop$ holds, 0 otherwise, the definition can be straight-forwardly rewritten to:

$$(f * g)(n) = \sum_{k_1}^{n} \sum_{k_2}^{n} \delta(k_1 + k_2 = n) \cdot f(k_1) \cdot g(k_2) \tag{4.2}$$

With this, the associativity of discrete convolution can be proven:

Claim 1. *Discrete convolution is associative. That is $(f * g) * h = f * (g * h)$.*

Proof. By computation. The third equality relies on the fact that all components of k_1 are less than or equal to the corresponding component of n, and if either of k_3 and k_4 exceeds k_1 in any component, the equality $k_3 + k_4 = k_1$ never holds. The fifth equality is based on there being exactly one value of k_1 for which $k_3 + k_4 = k_1$ holds, thanks to the equality $k_3 + k_4 + k_2 = n$.

$$\begin{aligned}
&((f * g) * h)(n) \\
&= \sum_{k_1=0}^{n} \sum_{k_2=0}^{n} \delta(k_1 + k_2 = n) \cdot (f * g)(k_1) \cdot h(k_2) \\
&= \sum_{k_1=0}^{n} \sum_{k_2=0}^{n} \delta(k_1 + k_2 = n) \cdot \left(\sum_{k_3=0}^{k_1} \sum_{k_4=0}^{k_1} \delta(k_3 + k_4 = k_1) \cdot f(k_3) \cdot g(k_4) \right) \cdot h(k_2) \\
&= \sum_{k_1=0}^{n} \sum_{k_2=0}^{n} \delta(k_1 + k_2 = n) \cdot \left(\sum_{k_3=0}^{n} \sum_{k_4=0}^{n} \delta(k_3 + k_4 = k_1) \cdot f(k_3) \cdot g(k_4) \right) \cdot h(k_2) \\
&= \sum_{k_1=0}^{n} \sum_{k_2=0}^{n} \sum_{k_3=0}^{n} \sum_{k_4=0}^{n} \delta(k_3 + k_4 + k_2 = n) \cdot \delta(k_3 + k_4 = k_1) \cdot f(k_3) \cdot g(k_4) \cdot h(k_2) \\
&= \sum_{k_2=0}^{n} \sum_{k_3=0}^{n} \sum_{k_4=0}^{n} \delta(k_3 + k_4 + k_2 = n) \cdot f(k_3) \cdot g(k_4) \cdot h(k_2) \\
&= \sum_{a=0}^{n} \sum_{b=0}^{n} \sum_{c=0}^{n} \delta(a + b + c = n) \cdot f(a) \cdot g(b) \cdot h(c)
\end{aligned}$$

Analogically:

$$
\begin{aligned}
&(f*(g*h))(n)\\
&=\sum_{k_1=0}^{n}\sum_{k_2=0}^{n}\delta(k_1+k_2=n)\cdot f(k_1)\cdot(g*h)(k_2)\\
&=\sum_{k_1=0}^{n}\sum_{k_2=0}^{n}\delta(k_1+k_2=n)\cdot f(k_1)\sum_{k_3=0}^{k_2}\sum_{k_4=0}^{k_2}\delta(k_3+k_4=k_2)\cdot g(k_3)\cdot h(k_4)\\
&=\sum_{k_1=0}^{n}\sum_{k_2=0}^{n}\delta(k_1+k_2=n)\cdot f(k_1)\sum_{k_3=0}^{n}\sum_{k_4=0}^{n}\delta(k_3+k_4=k_2)\cdot g(k_3)\cdot h(k_4)\\
&=\sum_{k_1=0}^{n}\sum_{k_2=0}^{n}\sum_{k_3=0}^{n}\sum_{k_4=0}^{n}\delta(k_1+k_3+k_4=n)\cdot\delta(k_3+k_4=k_2)\cdot f(k_1)\cdot g(k_3)\cdot h(k_4)\\
&=\sum_{k_1=0}^{n}\sum_{k_3=0}^{n}\sum_{k_4=0}^{n}\delta(k_1+k_3+k_4=n)\cdot f(k_3)\cdot g(k_4)\cdot h(k_2)\\
&=\sum_{a=0}^{n}\sum_{b=0}^{n}\sum_{c=0}^{n}\delta(a+b+c=n)\cdot f(a)\cdot g(b)\cdot h(c)
\end{aligned}
$$

□

4.2 Appendix 2: Invariance of $\mathcal{P}$ under certain transformations of f

Lemma 2. *Let $f, g : \mathbb{N}_0^K \to \mathbb{R}$, $c \in \mathbb{R}$. Then*

$$((cf)*g)(\vec{n}) = (f*(cg))(\vec{n}) = c(f*g)(\vec{n}) \tag{4.3}$$

Proof. Directly follows from the definition. □

Definition 2. *Let $f : \mathbb{N}_0^K \to \mathbb{R}$, $\{c_k\}_{k=1}^K \subset \mathbb{R}$. Then we define $\underline{f}$ as:*

$$\underline{f}(\vec{n}) := f(\vec{n}) \cdot \prod_{k=1}^{K} c_k^{n_k} \tag{4.4}$$

Lemma 3.

$$(\underline{f} * \underline{g})(\vec{N}) = \underline{(f*g)}(\vec{N}) \tag{4.5}$$

Proof. By computation:

$$
\begin{aligned}
(\underline{f} * \underline{g})(\vec{N}) &= \sum_{\vec{n}}^{\vec{N}} \underline{f}(\vec{n}) \cdot \underline{g}(\vec{N}-\vec{n})\\
&= \sum_{\vec{n}}^{\vec{N}} \left(\prod_{k=1}^{K} c_k^{n_k}\right) f(\vec{n}) \cdot \left(\prod_{k=1}^{K} c_k^{(N-n)_k}\right) g(\vec{N}-\vec{n})\\
&= \left(\prod_{k=1}^{K} c_k^{N_k}\right) \sum_{\vec{n}}^{\vec{N}} f(\vec{n}) \cdot g(\vec{N}-\vec{n})\\
&= \underline{(f*g)}(\vec{N})
\end{aligned}
\tag{4.6}
$$

□

Definition 3 ($\mathcal{P}$). *Let*

$$\mathcal{P}\{f\}(\vec{n}_1;\ldots;\vec{n}_L) := Z_L\{f\}(\vec{N})^{-1} \cdot \delta\left(\sum_{l=1}^{L} \vec{n}_l = \vec{N}\right) \cdot \prod_{l=1}^{L} f(\vec{n}_l) \tag{4.7}$$

where

$$Z_L\{f\}(\vec{N}) = \begin{cases} f(\vec{N}) & L = 1 \\ (f * Z_{L-1}\{f\})(\vec{N}) & \textit{otherwise} \end{cases} \tag{4.8}$$

Claim 4.

$$\mathcal{P}\{c_0\underline{f}\} = \mathcal{P}\{f\} \tag{4.9}$$

Proof. First, we shall prove by induction, that $Z_L\{c_0\underline{f}\} = c_0^L Z_L\{\underline{f}\}$. The case $L = 1$ follows from the definition. For $L > 1$, assuming this holds for $L - 1$, we get (using lemma 2):

$$\begin{aligned} Z_L\{c_0\underline{f}\} &= (c_0\underline{f}) * Z_{L-1}\{c_0\underline{f}\} = c_0\left(\underline{f} * (c_0^{L-1} Z_{L-1}\{\underline{f}\})\right) \\ &= c_0^L \left(\underline{f} * Z_{L-1}\{\underline{f}\}\right) = c_0^L Z_L\{\underline{f}\} \end{aligned} \tag{4.10}$$

The next step is to prove $Z_L\{\underline{f}\} = \underline{Z_L\{f\}}$, again by induction. The base case $L = 1$ is trivial, and the case $L > 1$ is a straightforward application of lemma 3. Finally:

$$\begin{aligned} \mathcal{P}\{c_0\underline{f}\}(\vec{n}_1;\ldots;\vec{n}_L) &:= Z_L\{c_0\underline{f}\}(\vec{N})^{-1} \cdot \delta\left(\sum_{l=1}^{L} \vec{n}_l = \vec{N}\right) \cdot \prod_{l=1}^{L}\left(c_0\underline{f}(\vec{n}_l)\right) \\ &= \left(c_0^L \underline{Z_L\{f\}(\vec{N})}\right)^{-1} \cdot \delta\left(\sum_{l=1}^{L} \vec{n}_l = \vec{N}\right) \cdot c_0^L \prod_{l=1}^{L} \underline{f}(\vec{n}_l) \\ &= \left(Z_L(\vec{N}) \prod_{l=1}^{L} c_l^{N_l}\right)^{-1} \cdot \delta\left(\sum_{l=1}^{L} \vec{n}_l = \vec{N}\right) \cdot \prod_{l=1}^{L}\left(f(\vec{n}_l) \prod_{k=1}^{L} c_l^{n_{l,k}}\right) \\ &= Z_L(\vec{N})^{-1} \cdot \delta\left(\sum_{l=1}^{L} \vec{n}_l = \vec{N}\right) \cdot \prod_{l=1}^{L} f(\vec{n}_l) \\ &= \mathcal{P}\{f\}(\vec{n}_1;\ldots;\vec{n}_L) \end{aligned} \tag{4.11}$$

□

Bibliography

M R Evans and T Hanney. Nonequilibrium statistical mechanics of the zero-range process and related models. *Journal of Physics A: Mathematical and General*, 38(19):R195, apr 2005. doi: 10.1088/0305-4470/38/19/R01. URL https://dx.doi.org/10.1088/0305-4470/38/19/R01.