Limits to positional information in boundary-driven systems

Prashant Singh and Karel Proesmans[1, *]

[1] *Niels Bohr International Academy, Niels Bohr Institute,*
University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark

Chemical gradients can be used by a particle to determine its position. This *positional information* is of crucial importance, for example in developmental biology in the formation of patterns in an embryo. The central goal of this paper is to study the fundamental physical limits on how much positional information can be stored inside a system. To achieve this, we study positional information for general boundary-driven systems, and derive, in the near-equilibrium regime, a universal expression involving only the chemical potential and density gradients of the system. We also conjecture that this expression serves as an upper bound on the positional information of boundary driven systems beyond linear response. To support this claim, we test it on a broad range of solvable boundary-driven systems.

Introduction: Nature endows biological matter with the astounding ability to self-organize fascinating spatio-temporal patterns that pervade across several length scales [1]. For instance, one often sees insects with their bodies divided into segments of repetitive patterns or birds and animals with unique patterns of spots and stripes. One classic mechanism of pattern formation in reaction-diffusion systems is Turing pattern [2]. While naively, one expects diffusion to generate uniform concentration of chemicals, Turing showed that two diffusing chemical species with distinct diffusion coefficients and activation-inhibition interplay can, under suitable condition, spontaneously break the symmetry of homogeneous concentrations and generate recurring structures such as stripes, spots, or even more complex patterns.

Another important mechanism commonly studied in the context of developing embryos is the the French-flag model [3, 4]. In order to build complex body structures, individual cells have to take decisions and adopt fates suitable for their positions. Yet cells do not have any direct way to measure their positions. The key idea in the French-flag model is that during the early developmental stage of an embryo, some specific chemicals, called 'morphogens' are deposited locally on one side of the embryo. Following diffusion, these chemicals are translated inside the embryo thereby establishing a concentration gradient [5, 6]. Unlike in Turing patterns, the spatial symmetry is broken due to the presence of graded concentration of morphogens. Wolpert proposed that cells could determine their positions from the local morphogen concentrations within these graded profiles and take up fates correlated with their positions [3, 7]. Therefore, the graded-concentration profile of signalling morphogen provides 'positional information' to the cells, see Figure 1. Cells, in turn, read out these positional cues in a threshold-dependent manner and give rise to spatial patterns with distinct boundaries.

Although the idea of positional information was originally introduced in the context of developmental biology, its applications reach beyond this. For example, synthetic versions of the French-Flag model have been con-structed and positional information can play a crucial role in the construction of self-assembling soft materials, where individual components could determine their position through the chemical concentration of a signalling molecule [8–11].

Recently a mathematical framework has been developed to quantitatively study positional information [12–18]. Essentially, one defines the positional information as the amount of information one gets about the position x, given that one measures a concentration n of signal molecules. Information theory dictates that this amount of information is equal to

$$\mathcal{I}_{(n,x)}\left(\{\rho\}\right) = \int dn \; dx \; P(x,n) \log_2\left[\frac{P(x,n)}{P_x(x)P_n(n)}\right], \; (1)$$

where $P(x,n)$ is the joint probability distribution of position x and concentration n and $P_x(x)$ and $P_n(n)$ are the associated marginals. Without any prior knowledge, one generally sets $P_x(x)$ to be uniform while other probabilities are uniquely determined through the system's dynamics. We remark that if n and x take only discrete values, then the integrations in Eq. (1) need to be replaced appropriately by their summations. Also, we have written positional information as a function of model parameters $\{\rho\}$ whose precise definition will be given later.

This framework has been successfully applied to biological systems. For example, experiments reveal that the four gap genes in the Drosophila embryo carry approximately ~ 4.2 bits of information. This amount of information enables cells to know their positions within an errorbar of $\sim 1\%$ of the total embryo length [12]. On the other hand, there are, to the best of our knowledge, no results on the fundamental physical limits of how much positional information can be stored inside a system.

The central goal of this paper is to fill this gap by determining the maximal amount of positional information stored in boundary-driven systems, in terms of their distance from thermodynamic equilibrium [19–25]. In particular, we will look at systems that are in contact with two particle reservoirs with chemical potentials μ_L and μ_R [and $\Delta\mu = (\mu_L - \mu_R)$ being their difference]. Our
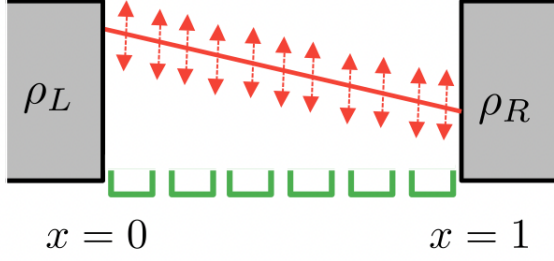
FIG. 1. Lattice sites depicted in green are in contact with two distinct reservoirs characterized by average densities ρ_L and ρ_R. Due to this coupling, the system attains a steady-state with average density profile shown by the red solid line. However, due to noise, there are fluctuations around this mean value which are indicated by dashed arrows. Positional information of a lattice site is obtained by measuring the local particle density and employing the formula in Eq. (1).

study reveals that in the near-equilibrium regime, positional information takes a universal form in the steady-state and is given in terms of the chemical potentials and average densities at the left and right reservoirs (denoted respectively by ρ_L and ρ_R):

$$\mathcal{I}_{(n,x)}\left(\bar{\rho}, \Delta\rho\right) \simeq \frac{\Delta\mu \; \Delta\rho}{12 \; k_B T \ln 2}, \tag{2}$$

where $\bar{\rho} = (\rho_L + \rho_R)/2$ and $\Delta\rho = (\rho_L - \rho_R)/2$. Here k_B is the Boltzmann constant and T is the absolute temperature at which the system is kept. In the far-from-equilibrium regime, we numerically demonstrate that the right hand side of Eq. (2) provides an upper bound to the positional information for a broad class of systems:

$$\mathcal{I}_{(n,x)}\left(\bar{\rho}, \Delta\rho\right) \leq \frac{\Delta\mu \; \Delta\rho}{12 \; k_B T \ln 2}. \tag{3}$$

This inequality provides a quantitative link between the positional information and the drive to maintain non-equilibrium. And for a given drive, it tells us that there is a limit on how much positional information can be generated. Since we will assume the temperature to be constant, we set $k_B T = 1$ in the remainder of our letter.

Positional information near equilibrium: We begin with a one-dimensional lattice system consisting of N sites represented by index i with $1 \leq i \leq N$. In bulk where $1 < i < N$, a particle can jump to either of its neighbouring sites with an arbitrary rate (the rate can also be a function of the number of particles in the sites). On the other hand, at the two end sites ($i = 1$ and $i = N$), the system is in contact with two particle reservoirs characterised by the average densities ρ_L and ρ_R. Without any loss of generality, we will take $\rho_L \geq \rho_R$ throughout this paper. Due to the coupling

with the reservoirs, the system eventually reaches a non-equilibrium steady-state.

Our starting point is to note that even though the system as a whole is driven out-of-equilibrium, local regions can still be described, to first order in gradients, by an equilibrium measure with parameters that vary slowly across the system [26, 27]. This *local thermodynamic equilibrium* then allows us to identify thermodynamic quantities such as chemical potential locally even in out-of-equilibrium set-ups. With this idea in mind, one can introduce a scaled continuous variable $x = i/N$ and take the limit $N \to \infty$, so that $x \in [0,1]$ and write the distribution to observe a local density $n(x)$ at position x in the steady-state as [23]

$$P(n|x) \sim \exp\left[-\left(\mathcal{G}_{\mu(x)}\left(n\right) - \mathcal{G}_{\mu(x)}\left(\rho(x)\right)\right)\right], \tag{4}$$

where $\mathcal{G}_{\mu(x)}\left(n\right)$ is given in terms of the Helmholtz free energy $a(n)$ as

$$\mathcal{G}_{\mu(x)}(n) = a(n) - \mu(x)n, \quad \text{and} \quad \mu(x) = \left.\frac{\partial a(n)}{\partial n}\right|_{n=\rho(x)}. \tag{5}$$

Here $\mu(x)$ stands for the local chemical potential which is given in terms of the local average density $\rho(x) = \langle n(x) \rangle$ such that $\rho(0) = \rho_L$ and $\rho(1) = \rho_R$.

In general, the form of the density $\rho(x)$ can be both linear or non-linear in x, and depends on the specific model. However, close to equilibrium, one generally expects it to take a linear form $\rho(x) \simeq \bar{\rho} - (2x - 1)\Delta\rho$ according to the diffusion equation. This is also in agreement with experimental observations for certain morphogens [28]. Plugging this in Eq. (4) then gives us the conditional distribution $P(n|x)$. Recall that we also need the joint probability distribution $P(n, x)$ in Eq. (1). To obtain this, we assume that there is no information about the position before the measurement, and therefore a flat prior distribution, $P_x(x) = 1$ [12]. This implies $P(x, n) = P(n|x)$ and we now have all quantities needed in Eq. (1). We show in the supplemental materials [29] that under these assumptions, the positional information is given by

$$\mathcal{I}_{(n,x)}\left(\bar{\rho}, \Delta\rho\right) \simeq \frac{\Delta\rho^2}{6 \ln 2 \; \sigma_2(\bar{\rho})}, \tag{6}$$

where $\sigma_k(\bar{\rho}) = \langle (n - \bar{\rho})^k \rangle_{\text{eq}}$ is the k-th central moment of the density at equilibrium ($\Delta\rho = 0$). As expected, $\mathcal{I}_{(n,x)}\left(\bar{\rho}, \Delta\rho\right)$ goes to zero at equilibrium since the density remains constant across the system and therefore the chemical concentration does not correlate with position.

We now establish the connection of positional information with the chemical potential difference $\Delta\mu = (\mu_L - \mu_R)$. Noting Eq. (5), this can be expressed in terms of the derivative of the free energy. At the leading order in $\Delta\rho$, we find $\Delta\mu \sim a_2(\bar{\rho}) \, \Delta\rho$, where $a_2(\bar{\rho})$ stands for the second derivative of $a(\bar{\rho})$. We next use the fluctuation-response relation to write $a_2(\bar{\rho}) = 1/\sigma_2(\bar{\rho})$; see [29] for a
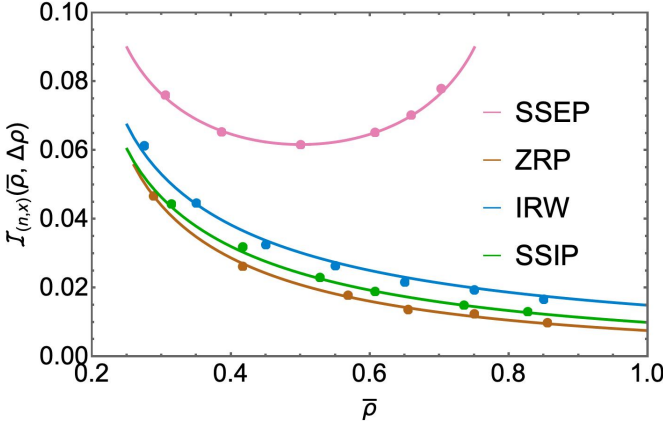
FIG. 2. Plot of the positional information as a function of $\bar{\rho}$ for four different boundary-driven processes and its comparison with the numerical simulations. In each case, solid lines illustrate the analytical expression, c.f., table I, while symbols denote simulation data. We have fixed the bias to $\Delta\rho = 0.25$.

proof. This leads us to the expression of chemical potential difference as

$$\Delta\mu \simeq \frac{2\Delta\rho}{\sigma_2(\bar{\rho})}. \qquad (7)$$

Combining this with Eq. (6), we arrive at the universal form of the positional information quoted in Eq. (2). This is first main result of our letter. It is valid in the near-equilibrium regime for any 1-dimensional boundary-driven system.

Far-from-equilibrium regime: Up to now, our analysis has focused on the near-equilibrium situation where local thermodynamic equilibrium enabled us to derive a general expression. An immediate question that now follows is - what happens to Eq. (2) in the far-from-equilibrium situations? Although we are unable to prove it generally, numerical results on a broad range of models suggest that the near-equilibrium result serves as an upper bound on the positional information far from equilibrium. We will now show this for four different models, namely symmetric simple exclusion processes (SSEP), zero-range processes (ZRP), independent random walkers (IRW), and simple symmetric inclusion processes (SSIP).

Example I: SSEP- Consider the SSEP model where every lattice site can either be vacant ($n_i = 0$) or be occupied by a single particle ($n_i = 1$). Any particle in the bulk can jump to one of its neighbouring sites with some rate p, as long as the target site is vacant. However, dynamics at the end sites are modified due to the presence of particle reservoirs, see [29] for details. Since the occupation number n for this model is a binary variable, one can write the conditional probability $P(n|x) = \rho(x)\delta_{n,1} + [1 - \rho(x)]\,\delta_{n,0}$ with density $\rho(x) = \bar{\rho} - (2x - 1)\Delta\rho$ [20–23]. If we now use this in Eq. (1), the positional information can be calculated.
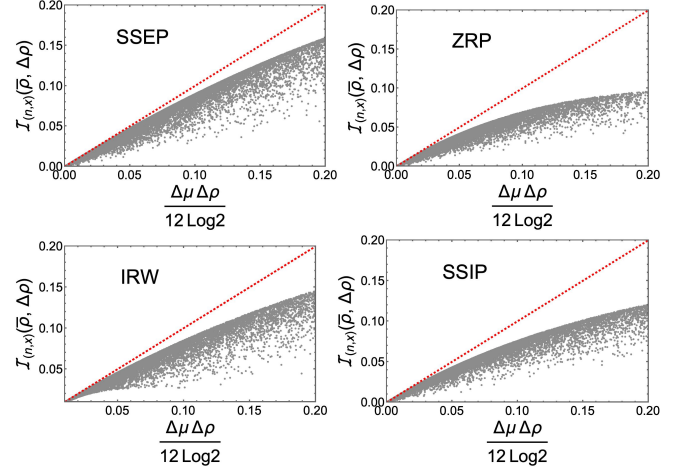

FIG. 3. Upper bound on the positional information in terms of the differences $\Delta\mu$ and $\Delta\rho$ for four different boundary-driven processes. For all cases, the gray symbols are the analytical results in Table (I) for different values of $\bar{\rho}$ and $\Delta\rho$, while the red line corresponds to the upper bound in (3). Clearly, the bound is saturated in the near-equilibrium limit. We have set $k_B T = 1$.

We have presented this expression in Table (I), and validated it against numerical simulations in Figure 2. It is also consistent with our general result in Eq. (2) in the limit $\Delta\rho \to 0$, as one can show that $\sigma_2(\bar{\rho}) = \bar{\rho}(1 - \bar{\rho})$. Our objective now is to compare this expression with the chemical potential difference $\Delta\mu$ quoted in Table (I). In Figure 3, we have plotted both $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ and $\Delta\mu\Delta\rho/12\ln 2$ for several values of $\bar{\rho}$ and $\Delta\rho$. The red dashed line represents the bound in (3). Clearly all positional information values lie below this line. In fact, for SSEP, we have rigorously proved in [29] that the bound is valid across all parameter values. Hence, our upper bound (3) holds for the SSEP arbitrarily far from the equilibrium.

Example II: ZRP- As a second example, we look at the boundary-driven zero-range process [24]. Unlike in SSEP, the lattice sites are now capable of accommodating an arbitrary number of particles. In bulk, a particle can jump to any of its neighbouring side with a rate of pu_{n_i}, where u_{n_i} is a non-negative function of n_i. At the boundaries, the dynamics are modified allowing for the addition or removal of particles. For this model, the steady-state density profile is non-linear. Moreover, for the case of constant rate $u_n = 1$, the positional information can be calculated as shown in Table (I). In Figure 2, we have compared this expression with numerical simulations and found a good agreement between them. Next we employ this result in conjunction with $\Delta\mu$ to test the upper bound (3). As depicted in Figure 3, positional information values are situated below the red line even for this model. This observation holds true across all parameter regimes. The bound is satu-

Model	$\Delta\mu$	$\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$
SSEP	$\ln\left(\frac{\bar{\rho}+\Delta\rho}{1-\bar{\rho}-\Delta\rho}\right) - \ln\left(\frac{\bar{\rho}-\Delta\rho}{1-\bar{\rho}+\Delta\rho}\right)$	$\frac{1}{\ln 16\ \Delta\rho}\Big[\mathcal{Y}(\bar{\rho},\Delta\rho) + \mathcal{Y}(1-\bar{\rho},\Delta\rho) - 2\Delta\rho\Big] - (1-\bar{\rho})\log_2(1-\bar{\rho}) - \bar{\rho}\log_2\bar{\rho}$
ZRP	$\ln\left(\frac{\bar{\rho}+\Delta\rho}{1+\bar{\rho}+\Delta\rho}\right) - \ln\left(\frac{\bar{\rho}-\Delta\rho}{1+\bar{\rho}-\Delta\rho}\right)$	$\frac{1}{\ln 2(c_+ - c_-)}\Big[\mathrm{Li}_2(1-c_+) - \mathrm{Li}_2(1-c_-) + c_-\ln c_- - (1-c_+)\ln(1-c_+)$ $+ (1-c_-)\ln(1-c_-) - c_+\ln c_+ - \sum_{n=0}^{\infty} f(n)\ln f(n)\Big] + \log_2(c_+ - c_-)$
IRW	$\ln(\bar{\rho}+\Delta\rho) - \ln(\bar{\rho}-\Delta\rho)$	$\frac{1}{\ln 16\ \Delta\rho}\Big[\mathcal{Y}(\bar{\rho},\Delta\rho) - 6\bar{\rho}\ \Delta\rho\Big] - \sum_{n=0}^{\infty}\frac{g(n)}{n!}\log_2 g(n)$
SSIP	$\ln\left(\frac{\bar{\rho}+\Delta\rho}{m+\bar{\rho}+\Delta\rho}\right) - \ln\left(\frac{\bar{\rho}-\Delta\rho}{m+\bar{\rho}-\Delta\rho}\right)$	$\frac{1}{\ln 16\ \Delta\rho}\Big[\mathcal{Y}(\bar{\rho},\Delta\rho) - \mathcal{Y}(m+\bar{\rho},\Delta\rho) + 2m\ \Delta\rho\Big] + (m-1)\log_2 m$ $- \sum_{n=0}^{\infty}\frac{m\Gamma(m+n)}{\Gamma(n+1)\Gamma(m)}\mathcal{S}(n)\log_2\mathcal{S}(n)$

TABLE I. Expressions of the positional information and chemical potential difference for different boundary-driven systems studied in this letter. Derivation of these expressions along with the forms of different functions used to represent them are provided in [29].

rated in the near-equilibrium regime where both $\Delta\mu$ and $\Delta\rho$ approach zero as expected. Furthermore, in this example also, $\mathcal{I}_{(n,x)}(\bar{\rho},\Delta\rho)$ is bounded by (3) even in the far-from equilibrium conditions.

Example III: IRW- Another choice of rate $u_n = n$ corresponds to the independent, unbiased random walkers [24] and turns out to be analytically solvable. Once again, we use the positional information and chemical potential difference $\Delta\mu$ from Table (I) to check our bound for this model. Plotting $\mathcal{I}_{(n,x)}(\bar{\rho},\Delta\rho)$ and $(\Delta\mu\Delta\rho)/(12\ln 2)$ in Figure 3, we have demonstrated that the upper bound is satisfied for this model. This model presents a third solvable example where the upper bound is valid in all parameter regimes.

Example IV: SSIP- As a final example, we investigate the simple symmetric inclusion process [30, 31]. For this model also, a lattice site can accommodate an arbitrary number of particles. However unlike the previous examples, the jump rate from $i \to j$ [with $j = (i\pm 1)$] depends on the both occupation numbers n_i and n_j. We choose the form of this rate as $pn_i(n_j + m)$ where $m\ (> 0)$ is a parameter in the model. Furthermore, the system is coupled with two reservoirs at the boundaries which drive it to a non-equilibrium steady-state. In this steady-state, we calculate the positional information and $\Delta\mu$ as presented in Table (I). These results are plotted in the bottom panel of Figure 3. From this, it is evident that the upper bound (3) is satisfied by this model. This further reinforces the conjecture that our bound holds for general boundary-driven system. Therefore, in all the examples investigated in this letter, we observe that there exists a quantitative constraint on the amount of positional in-

formation given in terms of the chemical potential and density gradients driving the system.

Conclusions and outlook: In conclusion, we established a link between positional information and non-equilibrium statistical physics. More specifically, we derived a universal expression for the positional information $\mathcal{I}_{(n,x)}(\bar{\rho},\Delta\rho)$ near equilibrium, involving only the chemical potential difference driving the system, and the difference in average densities. Furthermore, our analysis on several solvable models suggested that this relation turns into an upper bound in the far-from-equilibrium conditions: $\mathcal{I}_{(n,x)}(\bar{\rho},\Delta\rho) \leq \Delta\mu\ \Delta\rho/12\ln 2$. This means that there is a limit on how much positional information the morphogen particles can provide, depending on how far the system is from equilibrium.

Our work paves the way for several future directions. While we examined specific models to demonstrate the bound, it remains an open problem to prove it generically. Moreover, throughout this letter, we have focused only on lattice models that are driven out-of-equilibrium from the boundary. Extending this study for other models with bulk drive is an important future direction. This is especially relevant because, in some experiments, morphogen molecules experience degradation effects inside the embryo [5, 6]. Under this circumstance, the local equilibrium assumption is no longer valid, and one needs to develop new theoretical methodologies to tackle the problem [32].

––––––––––

* prashant.singh@nbi.ku.dk; karel.proesmans@nbi.ku.dk

[1] H. Meinhardt, *Models of Biological Pattern Formation (Academic Press, London, 1982)* (1982).

[2] A. M. Turing, The chemical basis of morphogenesis, Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **237**, 37 (1952).

[3] L. Wolpert, Positional information and the spatial pattern of cellular differentiation, Journal of Theoretical Biology **25**, 1 (1969).

[4] J. Sharpe, Wolpert's French Flag: what's the problem?, Development **146**, dev185967 (2019).

[5] T. Gregor, W. Bialek, R. R. de Ruyter van Steveninck, D. W. Tank, and E. F. Wieschaus, Diffusion and scaling during early embryonic pattern formation, Proceedings of the National Academy of Sciences **102**, 18403 (2005).

[6] V. Grieneisen, B. Scheres, P. Hogeweg, and A. Maree, Morphogengineering roots: Comparing mechanisms of morphogen gradient formation, BMC Systems Biology **6**, 37 (2012).

[7] L. Wolpert, Positional information and pattern formation, Current Topics in Developmental Biology **117**, 597 (2016), (Essays on Developmental Biology, Edited by P. M. Wassarman).

[8] A. Zadorin, Y. Rondelez, G. Gines, V. Dilhas, A. Zambrano, J.-C. Galas, and A. Estevez-Torres, Synthesis and materialization of a reaction-diffusion french flag pattern, Nature Chemistry **9**, 990 (2017).

[9] A. Baccouche, K. Montagne, A. Padirac, T. Fujii, and Y. Rondelez, Dynamic dna-toolbox reaction circuits: A walkthrough, Methods **67**, 234 (2014).

[10] S. Toda, W. L. McKeithan, T. J. Hakkinen, P. Lopez, O. D. Klein, and W. A. Lim, Engineering synthetic morphogen systems that can program multicellular patterning, Science **370**, 327 (2020).

[11] A. Dupin, L. Aufinger, I. Styazhkin, F. Rothfischer, B. K. Kaufmann, S. Schwarz, N. Galensowske, H. Clausen-Schaumann, and F. C. Simmel, Synthetic cell–based materials extract positional information from morphogen gradients, Science Advances **8**, eabl9228 (2022).

[12] J. O. Dubuis, G. Tkačik, E. F. Wieschaus, T. Gregor, and W. Bialek, Positional information, in bits, Proceedings of the National Academy of Sciences **110**, 16301 (2013).

[13] T. Gregor, D. Tank, E. Wieschaus, and W. Bialek, Probing the limits to positional information, Cell **130**, 153 (2007).

[14] W. Bialek, Biophysics: Searching for principles, Biophysics: Searching for Principles (2012).

[15] G. Tkačik, J. O. Dubuis, M. D. Petkova, and T. Gregor, Positional Information, Positional Error, and Read-out Precision in Morphogenesis: A Mathematical Framework, Genetics **199**, 39 (2014).

[16] P. Hillenbrand, U. Gerland, and G. Tkačik, Beyond the french flag model: Exploiting spatial and gene regulatory interactions for positional information, PLoS ONE **11**, 1 (2016).

[17] G. Tkačik and T. Gregor, The many bits of positional information, Development **148**, dev176065 (2021).

[18] K. S. Iyer, C. Prabhakara, S. Mayor, and M. Rao, Cellular compartmentalisation and receptor promiscuity as a strategy for accurate and robust inference of position during morphogenesis, eLife **12**, e79257 (2023).

[19] R. A. Blythe and M. R. Evans, Nonequilibrium steady states of matrix-product form: a solver's guide, Journal of Physics A: Mathematical and Theoretical **40**, R333 (2007).

[20] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, Exact solution of a 1d asymmetric exclusion model using a matrix formulation, Journal of Physics A: Mathematical and General **26**, 1493 (1993).

[21] B. Derrida, J. L. Lebowitz, and E. R. Speer, Free energy functional for nonequilibrium systems: An exactly solvable case, Physical Review Letters **87**, 150601 (2001).

[22] B. Derrida, J. L. Lebowitz, and E. R. Speer, Large deviation of the density profile in the steady state of the open symmetric simple exclusion process, Journal of Statistical Physics **107**, 599 (2002).

[23] B. Derrida, Non-equilibrium steady states: fluctuations and large deviations of the density and of the current, Journal of Statistical Mechanics: Theory and Experiment , P07023 (2007).

[24] E. Levine, D. Mukamel, and G. M. Schütz, Zero-range process with open boundaries, Journal of Statistical Physics **120**, 759 (2005).

[25] G. Schütz, Exactly solvable models for many-body systems far from equilibrium (Academic Press, 2001).

[26] S. R. d. Groot and P. Mazur, *Non-equilibrium Thermodynamics (Dover Publications, New York, 1984)* (1984).

[27] S. Sasa, Derivation of hydrodynamics from the hamiltonian description of particle systems, Physical Review Letters **112**, 100602 (2014).

[28] M. J. Thompson, C. A. Young, V. Munnamalai, and D. M. Umulis, Early radial positional information in the cochlea is optimized by a precise linear bmp gradient and enhanced by sox2, Scientific Reports **13**, 8567 (2023).

[29] P. Singh and K. Proesmans, Supplementary material for " limits to positional information in boundary-driven systems" (2024).

[30] K. Vafayi and M. H. Duong, Weakly nonequilibrium properties of a symmetric inclusion process with open boundaries, Physical Review E **90**, 052143 (2014).

[31] C. Franceschini, P. Gonçalves, and F. Sau, Symmetric inclusion process with slow boundary: Hydrodynamics and hydrostatics, Bernoulli **28**, 1340 (2022).

[32] G. L. Eyink, J. L. Lebowitz, and H. Spohn, Hydrodynamics and fluctuations outside of local equilibrium: Driven diffusive systems, Journal of Statistical Physics **83**, 385 (1996).

SUPPLEMENTARY MATERIAL: LIMITS TO POSITIONAL INFORMATION IN BOUNDARY-DRIVEN SYSTEMS

Prashant Singh and Karel Proesmans
Niels Bohr International Academy, Niels Bohr Institute,
University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark

In this supplementary note, we will present an extensive derivation of the results which were quoted in the main text of our letter. To begin with, it is useful to recall the mathematical framework of positional information developed in [1]

I. MATHEMATICAL FRAMEWORK FOR POSITIONAL INFORMATION

Wolpert's idea is that although cells do not have any direct way to measure their positions, they can still acquire positional information by reading out the local concentration of the signalling morphogen molecules [2]. During the initial stage of development, cells inside an embryo have similar structures, and if one does not measure the concentration of the morphogen, it could lie anywhere in the embryo. In the language of probabilities, this means that the position x of a cell is drawn from a prior probability distribution $P_x(x)$, which we choose to be uniform, *i.e.* $P_x(x) = 1$. Now if one observes a certain certain morphogen concentration n, then the position of a cell can be more accurately specified. However, due to the fact that the cells are in a noisy environment, the position of a cell after measurement is still drawn from a probability distribution $P(x|n)$ but conditioned on the value n. Observe that this conditional distribution is always narrower than the prior $P_x(x)$, and the degree of this narrowness represents the amount of positional information gained by the cells. For instance, if the distribution $P(x|n)$ is still flat, then not much information is gained about the position of the cell. On the other hand, if $P(x|n)$ is highly peaked at some value of x, then the position of a cell is determined with more precision.

The two distributions $P_x(x)$ and $P(x|n)$ are the main ingredients in this framework and the information gained by measuring the concentration n is given by

$$\mathbb{I}_{n \to x} = S[P_x(x)] - S[P(x|n)], \tag{S1}$$

where $S[P_x(x)]$ and $S[P(x|n)]$ are the entropies associated with distributions $P_x(x)$ and $P(x|n)$

$$S[P_x(x)] = -\int dx \ P_x(x) \log_2 [P_x(x)], \tag{S2}$$

$$S[P(x|n)] = -\int dx \ P(x|n) \log_2 [P(x|n)]. \tag{S3}$$

The fact that $P(x|n)$ is narrower than $P_x(x)$ implies that $S[P(x|n)]$ is smaller than $S[P_x(x)]$. Hence $\mathbb{I}_{n \to x}$ in Eq. (S1) can take only non-negative values. Moreover, we will assume that both x and n are continuous variables. When they take discrete values, the integrations over x and n need to be replaced by their summations.

Now if we randomly choose a cell, then the morphogen concentration will be distributed as $P_n(n)$ and taking the average of Eq. (S1) with this distribution gives us

$$\mathcal{I}_{(n,x)}(\{\rho\}) = \int dn \ P_n(n) \Big[S[P_x(x)] - S[P(x|n)] \Big], \tag{S4}$$

$$= \int dn \ dx \ P(x,n) \log_2 \left[\frac{P(x,n)}{P_x(x)P_n(n)} \right]. \tag{S5}$$

Here $P(x,n) = P(x|n)P_n(n) = P(n|x)P_x(x)$ stands for the joint distribution of n and x. Moreover, we have written positional information as a function of model parameters $\{\rho\}$ whose precise definition will be given later. Interestingly, the expression in Eq. (S5) emphasizes that the average information is mutual, *i.e.* the information gained about the position of a cell by measuring the morphogen concentration is on average same as the information gained about the morphogen concentration by measuring the position of a cell. We therefore have

$$\mathcal{I}_{(n,x)}(\{\rho\}) = \int dx \ P_x(x) \Big[S[P_n(n)] - S[P(n|x)] \Big], \tag{S6}$$

with entropies measured in bits as

$$S[P_n(n)] = -\int dn \ P_n(n) \log_2 [P_n(n)], \tag{S7}$$

$$S[P(n|x)] = -\int dn \ P(n|x) \log_2 [P(n|x)]. \tag{S8}$$

In some of our calculations below, Eq. (S6) turns out to be more useful to calculate the positional information.

II. A PERTURBATIVE APPROACH FOR POSITIONAL INFORMATION IN BOUNDARY-DRIVEN SYSTEMS

In this section, we develop a perturbative approach to obtain the positional information for general boundary driven systems. Let us consider a one-dimensional lattice system consisting of N sites represented by the index i with $1 \leq i \leq N$. These sites can either accommodate an arbitrary number of particles (as observed in systems like ZRP or IRW) or can have at most a fixed number of particles (as seen in SSEP). In bulk where $1 < i < N$, a particle can jump to either of its neighbouring sites with an arbitrary rate (the rate can also depend on the occupation numbers of these sites). On the other hand, at the two end sites ($i = 1$ and $i = N$), the system is in contact with two particle reservoirs characterised by the average densities ρ_L and ρ_R and chemical potentials μ_L and μ_R. However, both reservoirs have the same temperature T. Without any loss of generality, we will take $\rho_L \geq \rho_R$ and $k_B T = 1$ (k_B is the Boltzmann constant) throughout our discussion. Due to the coupling with the reservoirs, we assume that the system eventually reaches a non-equilibrium steady-state.

Our approach relies on the fact that even though there are significant departures from equilibrium in the system as a whole, local regions can still be described, to first order in gradients, by an equilibrium measure with parameters that vary slowly across the system This *local thermodynamic equilibrium* allows one to identify thermodynamic quantities such as chemical potential locally even in out-of-equilibrium set-ups. With this idea in mind, we now define a scaled continuous variable $x = i/N$ ($N \gg 1$) with $x \in [0, 1]$ and write the probability distribution to observe a local density n at position x as

$$P(n|x) \sim \exp\left[-\left(\mathcal{G}_{\mu(x)}(n) - \mathcal{G}_{\mu(x)}(\rho(x))\right)\right], \tag{S9}$$

where $\mathcal{G}_{\mu(x)}(n)$ is given in terms of the Helmholtz free energy $a(n)$ as

$$\mathcal{G}_{\mu(x)}(n) = a(n) - \mu(x)n \ \text{ and } \ \mu(x) = \left.\frac{\partial a(n)}{\partial n}\right|_{n=\rho(x)}. \tag{S10}$$

Here $\mu(x)$ is the local chemical potential which is given in terms of the local average density $\rho(x) = \langle n(x) \rangle$ such that $\rho(0) = \rho_L$ and $\rho(1) = \rho_R$. Although the precise form of $\rho(x)$ depends on the specific model, it turns out useful to expand $\rho(x)$ as a series in $\Delta\rho = (\rho_L - \rho_R)/2$ as

$$\rho(x) = \bar{\rho} + (1 - 2x)\,\Delta\rho + \sum_{k=2}^{\infty} \mathcal{L}_k(x, \bar{\rho})\,\Delta\rho^k, \tag{S11}$$

where $\bar{\rho} = (\rho_L + \rho_R)/2$. For computational convenience, we have written the $k = 1$ term separately in the above expression and taken it to be $\mathcal{L}_1(x, \bar{\rho}) = (1 - 2x)$. This form is sensible because the average density should remain invariant under the transformation $\rho_L \leftrightarrow \rho_R$ and $x \to (1 - x)$. By the same symmetry argument, we must also have $\mathcal{L}_k(x, \bar{\rho}) = (-1)^k \mathcal{L}_k(1 - x, \bar{\rho})$ for all values of k. Apart from this symmetry, we do not make any assumption on $\mathcal{L}_k(x, \bar{\rho})$ and their specific forms will depend on the model.

Using the expansion in Eq. (S11), the local chemical potential $\mu(x)$ and the free energy in Eq. (S10) can also be expanded in $\Delta\rho$ as

$$\mu(x) = \bar{\mu} + \sum_{k=1}^{\infty} \frac{a_{k+1}(\bar{\rho})}{k!} \ [\mathcal{K}(x, \bar{\rho})]^k, \tag{S12}$$

$$\mathcal{G}_{\mu(x)}(n) = \mathcal{G}_{\bar{\mu}}(n) - n\sum_{k=1}^{\infty} \frac{a_{k+1}(\bar{\rho})}{k!} [\mathcal{K}(x, \bar{\rho})]^k, \tag{S13}$$

$$\mathcal{G}_{\mu(x)}(\rho(x)) = \mathcal{G}_{\bar{\mu}}(\bar{\rho}) - \bar{\rho}\sum_{k=1}^{\infty} \frac{a_{k+1}(\bar{\rho})}{k!} [\mathcal{K}(x, \bar{\rho})]^k - \sum_{k=2}^{\infty} \frac{(k-1)a_k(\bar{\rho})}{k!} [\mathcal{K}(x, \bar{\rho})]^k, \tag{S14}$$

where we use the notation

$$\bar{\mu} = \frac{\partial a(\bar{\rho})}{\partial \bar{\rho}}, \qquad a_k(\bar{\rho}) = \frac{d^k a(\bar{\rho})}{d\bar{\rho}^k}, \quad \mathcal{K}(x, \bar{\rho}) = \rho(x) - \bar{\rho}. \tag{S15}$$

Plugging the expansions in Eq. (S9) yields

$$P(n|x) = P_{\text{eq}}(n) \ \exp\left[(n - \bar{\rho}) \sum_{k=1}^{\infty} \frac{a_{k+1}(\bar{\rho})}{k!} \left[\mathcal{K}(x, \bar{\rho})\right]^k - \sum_{k=2}^{\infty} \frac{(k-1)a_k(\bar{\rho})}{k!} \left[\mathcal{K}(x, \bar{\rho})\right]^k \right], \tag{S16}$$

where $P_{\text{eq}}(n) \sim e^{-(\mathcal{G}_{\bar{\mu}}(n) - \mathcal{G}_{\bar{\mu}}(\bar{\rho}))}$ is the equilibrium measure with average density $\bar{\rho}$ and chemical potential $\bar{\mu}$. The idea now is to use this series expansion in Eq. (S5) to obtain a perturbation expansion for the positional information. Let us calculate the first term in the expansion.

A. First term

It turns out, as also demonstrated later, that the first term is of the order $\sim \Delta\rho^2$. Hence, we include all terms up to this order in the density expansion in Eq. (S11) and truncate $\mathcal{K}(x, \bar{\rho})$ as

$$\mathcal{K}(x, \bar{\rho}) \simeq (1 - 2x)\ \Delta\rho + \mathcal{L}_2(x, \bar{\rho})\ \Delta\rho^2. \tag{S17}$$

From Eq. (S16), it then follows

$$P(n|x) \simeq P_{\text{eq}}(n) \left[1 + \Delta\rho\ (1 - 2x)\ a_2(\bar{\rho})\ (n - \bar{\rho}) + \frac{\Delta\rho^2}{2} \left\{ -a_2(\bar{\rho})\ (1 - 2x)^2 + a_2(\bar{\rho})^2 (1 - 2x)^2 (n - \bar{\rho})^2 \right. \right.$$
$$\left. \left. + \left(2a_2(\bar{\rho})\ \mathcal{L}_2(x, \bar{\rho}) + a_3(\bar{\rho})\ (1 - 2x)^2 \right)(n - \bar{\rho}) \right\} \right]. \tag{S18}$$

By noting that the prior is chosen to be uniform, $P_x(x) = 1$, the joint distribution $P(x, n)$ is then equal to the conditional probability $P(n|x)$. Moreover, the marginal distribution $P_n(n) = \int_0^1 dx\ P(n|x)$ can be calculated to be

$$P_n(n) \simeq P_{\text{eq}}(n) \left[1 - \frac{\Delta\rho^2}{6} \left\{ a_2(\bar{\rho})^2\ (n - \bar{\rho})^2 + \left(a_3(\bar{\rho}) + 6a_2(\bar{\rho}) \int_0^1 dx\ \mathcal{L}_2(x, \bar{\rho}) \right) (n - \bar{\rho}) \right\} \right]. \tag{S19}$$

We now have all quantities required to compute the positional information in Eq. (S5). Inserting these distributions, the first term in $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ can be written as

$$\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho) \simeq \frac{\Delta\rho^2}{6 \ln 2} a_2(\bar{\rho})^2\ \sigma_2(\bar{\rho}) \simeq \frac{\Delta\rho^2}{6 \ln 2\ \sigma_2(\bar{\rho})}, \tag{S20}$$

where $\sigma_k(\bar{\rho}) = \langle (n - \bar{\rho})^k \rangle_{\text{eq}}$ denotes the k-th central moment. In writing Eq. (S20), we have used the fluctuation-response relation $a_2(\bar{\rho}) = 1/\sigma_2(\bar{\rho})$. For completeness, the proof of this relation is provided in Section VI. To sum up, we have derived the first term in the series expansion of $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ in terms of the second central moment of the density at equilibrium.

B. Second term

In order to obtain the second order term in the expansion, we have to consider higher order terms in the expansion of density in Eq. (S11). The second term is of the order $\sim \Delta\rho^4$ and we therefore take

$$\mathcal{K}(x, \bar{\rho}) \simeq (1 - 2x)\ \Delta\rho + \mathcal{L}_2(x, \bar{\rho})\ \Delta\rho^2 + \mathcal{L}_3(x, \bar{\rho})\ \Delta\rho^3 + \mathcal{L}_4(x, \bar{\rho})\ \Delta\rho^4. \tag{S21}$$

We now proceed exactly as before but keeping terms up to order $\sim \Delta\rho^4$ in the analysis. The positional information can then be obtained to be

$$\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho) \simeq \frac{\Delta\rho^2}{6 \ln 2\ \sigma_2(\bar{\rho})} + \frac{8\ \Delta\rho^4}{45 \ln 2} \left[\frac{4a_2(\bar{\rho})a_4(\bar{\rho}) - 5\left(a_3(\bar{\rho})^2 + a_2(\bar{\rho})^3 \right)}{32 a_2(\bar{\rho})} + \frac{45 a_3(\bar{\rho})}{16} \int_0^1 dx\ \mathcal{L}_2(x, \bar{\rho})\ (4x^2 - 1) \right.$$
$$\left. - \frac{45}{16} a_2(\bar{\rho}) \left\{ \left(\int_0^1 dx\ \mathcal{L}_2(x, \bar{\rho}) \right)^2 - \int_0^1 dx\ \mathcal{L}_2(x, \bar{\rho})^2 - 2 \int_0^1 dx\ \mathcal{L}_3(x, \bar{\rho})(1 - 2x) \right\} \right]. \tag{S22}$$

For later comparison, it is useful to write this expression in terms of the $\sigma_k(\bar{\rho})$. To achieve this, we use the following set of relations

$$a_2(\bar{\rho}) = \frac{1}{\sigma_2(\bar{\rho})}, \quad a_3(\bar{\rho}) = -\frac{\sigma_3(\bar{\rho})}{\sigma_2(\bar{\rho})^3}, \quad a_4(\bar{\rho}) = -\frac{\sigma_4(\bar{\rho})}{\sigma_2(\bar{\rho})^4} + \frac{3}{\sigma_2(\bar{\rho})^5}\big[\sigma_3(\bar{\rho})^2 + \sigma_2(\bar{\rho})^3\big], \tag{S23}$$

which have been derived in Section VI. The above expression now becomes

$$\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho) \simeq \frac{\Delta\rho^2}{6\ln 2\; \sigma_2(\bar{\rho})} + \frac{8\;\Delta\rho^4}{45\ln 2\; \sigma_2(\bar{\rho})}\left[\frac{7\big(\sigma_3(\bar{\rho})^2 + \sigma_2(\bar{\rho})^3\big) - 4\sigma_2(\bar{\rho})\sigma_4(\bar{\rho})}{32\;\sigma_2(\bar{\rho})^4} + \frac{45\sigma_3(\bar{\rho})}{16\sigma_2(\bar{\rho})^2}\int_0^1 dx\; \mathcal{L}_2(x,\bar{\rho})\left(4x^2 - 1\right)\right.$$

$$\left. -\frac{45}{16}\left\{\left(\int_0^1 dx\; \mathcal{L}_2(x,\bar{\rho})\right)^2 - \int_0^1 dx\; \mathcal{L}_2(x,\bar{\rho})^2 - 2\int_0^1 dx\; \mathcal{L}_3(x,\bar{\rho})\big(1 - 2x\big)\right\}\right]. \tag{S24}$$

This gives the first two terms in the series expansion of the positional information. One can also obtain the higher order terms in the same way. In summary, we have developed a perturbative approach to calculate $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ for a general boundary-driven system. Our approach relies on the local equilibrium assumption and requires only the knowledge of the average density $\rho(x)$. Positional information is then obtained as an expansion in $\Delta\rho$ with coefficients depending on the equilibrium central moments of the density (or equivalently the derivatives of the free energy).

C. Connection of $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ with the chemical potential difference

Having developed a methodology to compute positional information, we are now in a position to establish its connection with the chemical potential difference driving the system, $\Delta\mu = \mu_L - \mu_R$. Following Eq. (S9), this can be written as

$$\Delta\mu = a_1(\bar{\rho} + \Delta\rho) - a_1(\bar{\rho} - \Delta\rho) \simeq \frac{2\Delta\rho}{\sigma_2(\bar{\rho})}. \tag{S25}$$

Combining this with the expression of $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ in Eq. (S20), we obtain for the leading order in $\Delta\rho$

$$\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho) \simeq \frac{\Delta\mu\; \Delta\rho}{12\ln 2}. \tag{S26}$$

Reinstating the $k_B T$ term, the expression becomes

$$\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho) \simeq \frac{\Delta\mu\; \Delta\rho}{12\; k_B T \ln 2}. \tag{S27}$$

This relation quantitatively gives the link between the positional information and non-equilibrium nature of the system. At least with only the first term in $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ and $\Delta\mu$, Eq. (S27) tells us how positional information increases on increasing the non-equilibrium drive.

Beyond linear response regime, writing any universal expression is difficult due to the dependence of higher-order terms in positional information on the specific model, as illustrated in Eq. (S24). In absence of a general formulation, we have studied several solvable boundary-driven systems. For each of these models, we find that the equality in Eq. (S27) turns into an upper bound in the far-from equilibrium conditions and we obtain a fundamental limit on the positional information in terms of the system's distance from the equilibrium

$$\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho) \leq \frac{\Delta\mu\; \Delta\rho}{12\; k_B T \ln 2}. \tag{S28}$$

Below, we demonstrate this relation for SSEP followed by a number of other models.

III. POSITIONAL INFORMATION FOR THE BOUNDARY-DRIVEN SSEP

One of the models that we discussed in the main text is the open symmetric simple exclusion process (SSEP). The model consists of N lattice sites in one dimension represented by the index i that runs from 1 to N. Each lattice site can either be vacant ($n_i = 0$) or occupied by a single particle ($n_i = 1$). The dynamics of the particles are as follows: Within the bulk, where $1 < i < N$, a particle can jump to either of its neighbouring sites with a rate p,

provided that these neighbouring sites are empty. On the other hand, the two boundary sites ($i = 1$ and $i = N$) are in contact with two different particle reservoirs, and their dynamics are modified accordingly. At $i = 1$, a particle can be added (removed) with a rate α_L (β_L) if the site is vacant (occupied). Conversely, at $i = N$, a particle can be added (removed) with a rate α_R (β_R) if it is vacant (occupied). At any small time interval $[t, t + dt]$, the occupancy variable $n_i(t)$ for each lattice site evolves according to the following update rule [3, 4]

$$n_i(t + \Delta t) - n_i(t) = \begin{cases} n_{i+1}(t), & \text{w.p.} \, [1 - n_i(t)] \, p\Delta t, \\ n_{i-1}(t), & \text{w.p.} \, [1 - n_i(t)] \, p\Delta t, \\ -n_i(t), & \text{w.p.} \, [1 - n_{i+1}(t)] \, p\Delta t, \quad , \quad \text{for } i \neq 1, \;\; i \neq N \\ -n_i(t), & \text{w.p.} \, [1 - n_{i-1}(t)] \, p\Delta t, \\ 0, & \text{otherwise,} \end{cases} \tag{S29}$$

$$n_1(t + \Delta t) - n_1(t) = \begin{cases} -n_1(t), & \text{w.p.} \, [1 - n_2(t)] \, p\Delta t, \\ n_2(t), & \text{w.p.} \, [1 - n_1(t)] \, p\Delta t, \\ 1, & \text{w.p.} \, [1 - n_1(t)] \, \alpha_L \Delta t, \\ -1, & \text{w.p.} \quad n_1(t) \, \beta_L \Delta t, \\ 0, & \text{otherwise,} \end{cases} \tag{S30}$$

$$n_N(t + \Delta t) - n_N(t) = \begin{cases} -n_N(t), & \text{w.p.} \, [1 - n_{N-1}(t)] \, p\Delta t, \\ n_{N-1}(t), & \text{w.p.} \, [1 - n_N(t)] \, p\Delta t, \\ 1, & \text{w.p.} \, [1 - n_N(t)] \, \alpha_R \Delta t, \\ -1, & \text{w.p.} \quad n_N(t) \, \beta_R \Delta t, \\ 0, & \text{otherwise,} \end{cases} \tag{S31}$$

where "w.p" is the short-hand notation for "with probability". Our aim is to calculate the positional information for this model. As evident from its definition, it is then necessary to compute the conditional probability $P(n|i)$ of having n particles at i-th site. Later, we show that this probability can be expressed in terms of the mean density $\rho_i(t) = \langle n_i(t) \rangle$. Therefore, in what follows, we will first calculate $\rho_i(t)$.

A.   Average density $\rho_i(t) = \langle n_i(t) \rangle$

From the update rules written above, one can write the differential equations for $\rho_i(t)$ as

$$\frac{\partial \rho_i(t)}{\partial t} = p\left[\rho_{i+1}(t) + \rho_{i-1}(t) - 2\rho_i(t)\right], \qquad (\text{for } i \neq 1, \;\; i \neq N) \tag{S32}$$

$$\frac{\partial \rho_1(t)}{\partial t} = \alpha_L - (p + \alpha_L + \beta_L)\rho_1(t) + p \, \rho_2(t), \tag{S33}$$

$$\frac{\partial \rho_N(t)}{\partial t} = \alpha_R - (p + \alpha_R + \beta_R)\rho_1(t) + p \, \rho_{N-1}(t). \tag{S34}$$

Since we are interested in the steady-state properties, we replace the time derivatives on the left hand sides of these equations by zero and obtain the solution as

$$\rho_i(t \to \infty) = \mathcal{A} + (i - 1) \, \mathcal{B}, \tag{S35}$$

$$\text{with } \mathcal{A} = \frac{\alpha_L \, (\alpha_R + \beta_R) \, (N - 1) + p \, (\alpha_L + \alpha_R)}{(\alpha_L + \beta_L) \, (\alpha_R + \beta_R) \, (N - 1) + p \, (\alpha_L + \beta_L + \alpha_R + \beta_R)}, \tag{S36}$$

$$\mathcal{B} = \frac{\alpha_R \beta_L - \alpha_L \beta_R}{(\alpha_L + \beta_L) \, (\alpha_R + \beta_R) \, (N - 1)}. \tag{S37}$$

For large N, we introduce a rescaled (continuous) variable $x = i/N$ where $x \in [0, 1]$ and rewrite $\rho(x) = \rho_i(t \to \infty)$

$$\rho(x) \simeq \rho_L - (\rho_L - \rho_R) \, x, \quad \text{with } \rho_L = \frac{\alpha_L}{\alpha_L + \beta_L}, \quad \rho_R = \frac{\alpha_R}{\alpha_R + \beta_R}. \tag{S38}$$

We have used ρ_L and ρ_R to represent the average densities of the left and the right reservoir respectively and are expressed in terms of the parameters of our model. Furthermore, the approximate equality in Eq. (S39) is used to
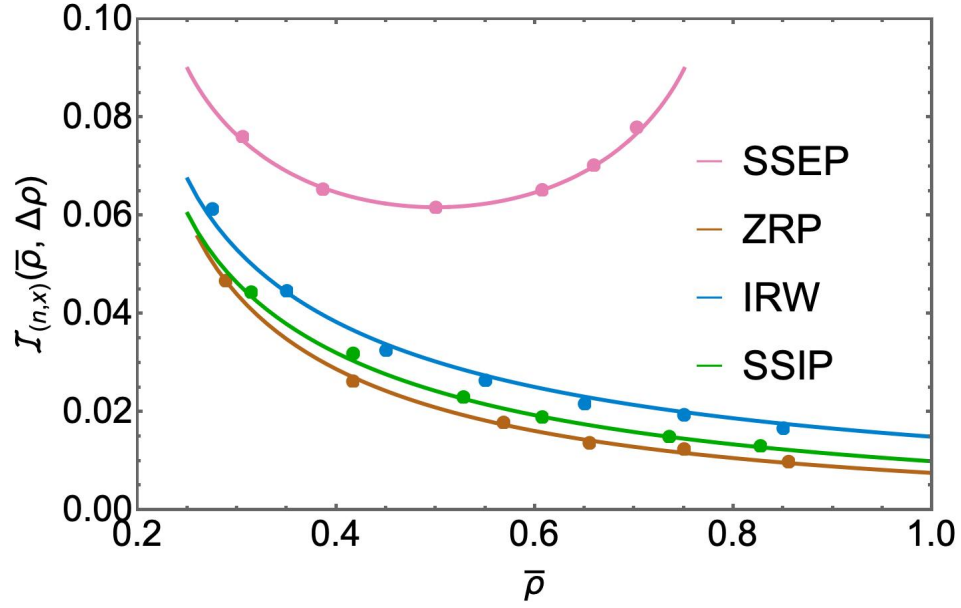
FIG. S1. Plot of the positional information as a function of $\bar{\rho}$ for four different boundary-driven processes and its comparison with the numerical simulations. In each case, solid lines illustrate the analytical expression, while symbols denote simulation data. We have fixed the bias to $\Delta\rho = 0.25$.

indicate that we are working in the large N limit. Finally, we express $\rho(x)$ in terms of the variables $\bar{\rho} = (\rho_L + \rho_R)/2$ and $\Delta\rho = (\rho_L - \rho_R)/2$

$$\rho(x) \simeq \bar{\rho} + (1 - 2x)\,\Delta\rho, \tag{S39}$$

and utilize this form to derive the positional information.

B. Positional information

In order to obtain the positional information, we will use its definition in Eq. (S6) and calculate the entropies $S[P(n|x)]$ and $S[P_n(n)]$. Recall that $n(x)$ is a binary variable that can take values either 0 or 1 depending on whether the site is vacant or occupied. We therefore have $\rho(x) = 1 \times P(n = 1|x) + 0 \times P(n = 0|x) = P(n = 1|x)$. The complementary probability will then simply be $P(n = 0|x) = 1 - \rho(x)$.

$$P(n = 1|x) = \rho(x) = \bar{\rho} + (1 - 2x)\,\Delta\rho, \tag{S40}$$
$$P(n = 0|x) = 1 - \rho(x) = 1 - \bar{\rho} - (1 - 2x)\,\Delta\rho. \tag{S41}$$

From these, we get

$$S[P(n|x)] = -\sum_{n=\{0,1\}} P(n|x) \log_2[P(n|x)], \tag{S42}$$
$$= -[\bar{\rho} + (1 - 2x)\,\Delta\rho] \log_2[\bar{\rho} + (1 - 2x)\,\Delta\rho] - [1 - \bar{\rho} - (1 - 2x)\,\Delta\rho] \log_2[1 - \bar{\rho} - (1 - 2x)\,\Delta\rho]. \tag{S43}$$

For the average information in Eq. (S6), we also have to perform averaging of $S[P(n|x)]$ over the prior distribution $P_x(x)$ which we have chosen to be uniform. This yields

$$
\begin{aligned}
\langle S[P(n|x)]\rangle_x &= \int_0^1 dx \ P_x(x) \ S[P(n|x)], \\
&= -\int_0^1 dx \left[[\bar{\rho} + (1-2x)\Delta\rho] \log_2 [\bar{\rho} + (1-2x)\Delta\rho] + [1 - \bar{\rho} - (1-2x)\Delta\rho] \log_2 [1 - \bar{\rho} - (1-2x)\Delta\rho] \right], \\
&= -\frac{1}{\ln 16 \ (\Delta\rho)} \left[(\bar{\rho} + \Delta\rho)^2 \ln (\bar{\rho} + \Delta\rho) - (\bar{\rho} - \Delta\rho)^2 \ln (\bar{\rho} - \Delta\rho) - 2\Delta\rho + (1 - \bar{\rho} + \Delta\rho)^2 \ln (1 - \bar{\rho} + \Delta\rho) \right. \\
&\qquad\qquad \left. - (1 - \bar{\rho} - \Delta\rho)^2 \ln (1 - \bar{\rho} - \Delta\rho) \right].
\end{aligned}
$$

$$(S44)$$

We next compute the other entropy, $S[P_n(n)]$, in the definition of $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$. We can write $P_n(n)$ by integrating over x in the joint probability $P(n, x) = P(n|x)$ and then using Eqs. (S40) and (S41) for $P(n|x)$. This gives

$$
P_n(n) = \int_0^1 dx \ P(n|x) = \bar{\rho} \ \delta_{n,1} + (1 - \bar{\rho}) \ \delta_{n,0}. \tag{S45}
$$

Here $\delta_{n,1}$ is the Kronecker delta which takes value one if $n = 1$ and zero otherwise (similarly for $\delta_{n,0}$). Using this expression above, we find

$$
S[P_n(n)] = -\sum_{\{n=0,1\}} P_n(n) \log_2 [P_n(n)] = -\bar{\rho} \log_2 \bar{\rho} - (1 - \bar{\rho}) \log_2 (1 - \bar{\rho}). \tag{S46}
$$

We now have all quantities essential for calculating the positional information. Substituting Eqs. (S44) and (S46) in Eq. (S6), the final expression is

$$
\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho) = \frac{1}{\ln 16 \ \Delta\rho} \left[\mathcal{Y}(\bar{\rho}, \Delta\rho) + \mathcal{Y}(1 - \bar{\rho}, \Delta\rho) - 2\Delta\rho \right] - (1 - \bar{\rho}) \log_2 (1 - \bar{\rho}) - \bar{\rho} \log_2 \bar{\rho}, \tag{S47}
$$

$$
\text{where } \mathcal{Y}(\bar{\rho}, \Delta\rho) = (\bar{\rho} + \Delta\rho)^2 \ln (\bar{\rho} + \Delta\rho) - (\bar{\rho} - \Delta\rho)^2 \ln (\bar{\rho} - \Delta\rho). \tag{S48}
$$

This result has been quoted in the main text. Eq. (S47) is also consistent with our general result in Eq. (S24) in the $\Delta\rho \to 0$ limit. In Figure S1, we have plotted $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ as a function of $\bar{\rho}$ and compared it with numerical simulations. We see an excellent agreement between our theoretical formula and the numerics. Notice that for a given $\Delta\rho$, $\bar{\rho}$ can vary between $\Delta\rho$ and $(1 - \Delta\rho)$. Within this range, we find that the positional information for SSEP changes in a non-monotonic manner with $\bar{\rho}$. To understand this heuristically, let us see what happens when $\bar{\rho} \to (1 - \Delta\rho)$ or equivalently $\rho_L \to 1$. In terms of the jump rates, this means that $\alpha_L \gg \beta_L$ and sites near the left boundary are more likely to be occupied with particles than those near the right boundary. Hence, if we pick a site randomly but is occupied by a particle, it is more likely to be closer to the left boundary than the right one. From the probabilistic perspective, this would imply that the distribution $P(x|n = 1)$ is sharply peaked at $x = 0$, while the complementary probability $P(x|n = 0)$ peaks at $x = 1$. The resulting entropy $\langle S[P(x|n)]\rangle_n$ associated with these peaked distributions, as expressed in Eq. (S1), is small which leads to a larger value of $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$. Similarly, for $\bar{\rho} \to \Delta\rho$, one can argue that the probabilities $P(x|n = 1)$ and $P(x|n = 0)$ are narrow functions of x which again give large $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$. In between these large values, $P(x|n = 1)$ and $P(x|n = 0)$ are broadest at some $\bar{\rho}$ for which the entropy $\langle S[P(x|n)]\rangle_n$ is highest. This in turn leads to the smallest positional information via Eq. (S1).

C. Upper bound on $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$

We will now use the derived expression of $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ to prove the upper bound in (S28). Observe that the chemical potential difference $\Delta\mu$ is expressed in terms of the first derivative of the associated free energy in Eq. (S25). To obtain this, we use the fluctuation-response relation in Eq. (S23) as

$$
a_2(\bar{\rho}) = \frac{1}{\sigma_2(\bar{\rho})} = \frac{1}{\bar{\rho}(1 - \bar{\rho})}, \implies a_1(\bar{\rho}) = \ln \left(\frac{\bar{\rho}}{1 - \bar{\rho}} \right). \tag{S49}
$$

where $\sigma_2(\bar{\rho}) = \bar{\rho}(1 - \bar{\rho})$ follows from Eqs. (S40) and (S41) with $\Delta\rho = 0$ and $k_B T = 1$ is assumed. Using Eq. (S25), we now get

$$
\Delta\mu = \ln \left(\frac{\bar{\rho} + \Delta\rho}{1 - \bar{\rho} - \Delta\rho} \right) - \ln \left(\frac{\bar{\rho} - \Delta\rho}{1 - \bar{\rho} + \Delta\rho} \right). \tag{S50}
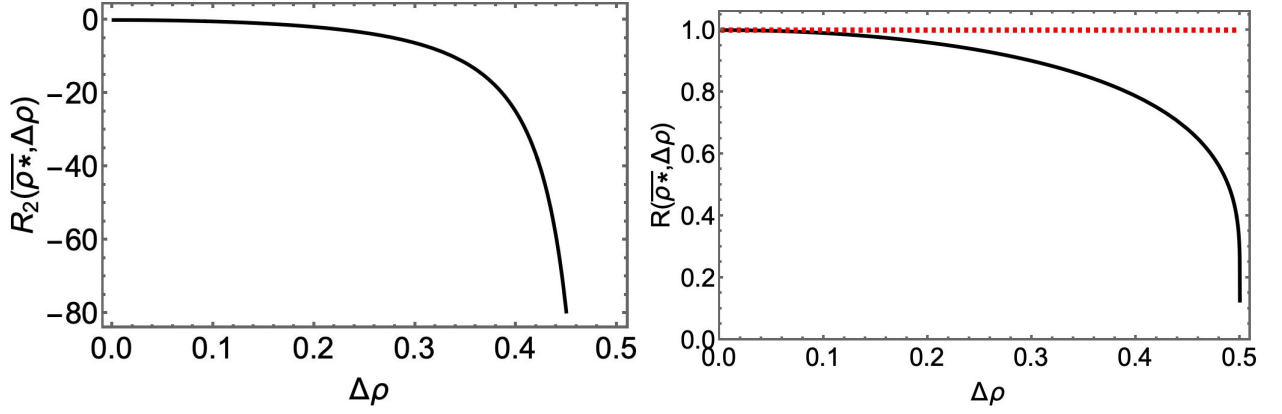$$

FIG. S2. *Left panel:* Plot of the second derivative $R_2(\bar{\rho}^*, \Delta\rho) = \left.\frac{\partial^2 R(\bar{\rho}, \Delta\rho)}{\partial \bar{\rho}^2}\right|_{\bar{\rho}=\bar{\rho}^*}$ where $\bar{\rho}^* = 1/2$. For all $\Delta\rho$, the second derivative takes only negative values. *Right panel:* Illustration of $R(\bar{\rho}^*, \Delta\rho)$ in Eq. (S53) as a function of $\Delta\rho$. The solid black line shows the plot while red dashed line indicates that $R(\bar{\rho}^*, \Delta\rho)$ is always bounded by the value one.

To derive the bound (S28), it is useful to define a ratio $R(\bar{\rho}, \Delta\rho)$ as

$$R(\bar{\rho}, \Delta\rho) = \frac{12 \ln 2 \; \mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)}{\Delta\mu \; \Delta\rho}. \tag{S51}$$

We are interested in finding the maximum value of this ratio. Below we show that this maximum value turns out to be one thereby proving the upper bound. Proceeding ahead, we take the first derivative of $R(\bar{\rho}, \Delta\rho)$ with respect to $\bar{\rho}$ for fixed $\Delta\rho$

$$\frac{\partial R(\bar{\rho}, \Delta\rho)}{\partial \bar{\rho}} = R(\bar{\rho}, \Delta\rho) \left[\frac{1}{\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)} \frac{\partial \; \mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)}{\partial \bar{\rho}} - \frac{1}{\Delta\mu} \frac{\partial \Delta\mu}{\partial \bar{\rho}} \right], \tag{S52}$$

and setting it to zero, we find the condition for optimality as $\bar{\rho}^* = 1/2$. Moreover, the second derivative of $R(\bar{\rho}, \Delta\rho)$ at $\bar{\rho}^*$ is always negative indicating that the extremum is a maxima. This is illustrated in the left panel of Figure S2. Therefore, we obtain

$$R(\bar{\rho}, \Delta\rho) \leq R\left(\bar{\rho}^* = \frac{1}{2}, \Delta\rho\right),$$
$$= \frac{3}{2\Delta\rho^2} \frac{(1 + 4 \; \Delta\rho^2) \; \text{Arctanh}(2\Delta\rho) - 2\Delta\rho + 2\Delta\rho \; \ln(1 - 4\Delta\rho^2)}{\ln(1 + 2\Delta\rho) - \ln(1 - 2\Delta\rho)}. \tag{S53}$$

Plotting this expression in Figure S2 (right panel), we find that the value of $R(\bar{\rho}^*, \Delta\rho)$ is always upper bounded by one. From our analysis above, this translates to $R(\bar{\rho}, \Delta\rho) \leq 1$ and we finally get

$$\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho) \leq \frac{\Delta\mu \; \Delta\rho}{12 \ln 2}. \tag{S54}$$

Hence for SSEP, we have rigorously derived that the bound (S28) is valid. We have also illustrated this in Figure S3 where we have plotted the two sides of (S54) along the two axes. The red dashed line indicates our upper bound. We clearly see that the positional information remains below this red line across all parameter values.

IV.   POSITIONAL INFORMATION FOR THE BOUNDARY DRIVEN ZRP

We now demonstrate positional information for another model, namely the open zero-range process (ZRP) and test our upper bound. The model consists of N lattice sites and each site can accommodate an arbitrary number of particles. This is different than the SSEP where each site can have at most a single particle. From any bulk site i, a particle can jump to either of its neighbouring sites with a rate pu_{n_i}, where u_{n_i} is a non-negative function of the number of particles n_i. At the boundaries ($i = 1$ and $i = N$), the dynamics are modified to allow for the addition (removal) of a particle with rate α_L ($\beta_L u_{n_1}$) for the left boundary and with rate α_R ($\beta_R u_{n_N}$) for the right one (see
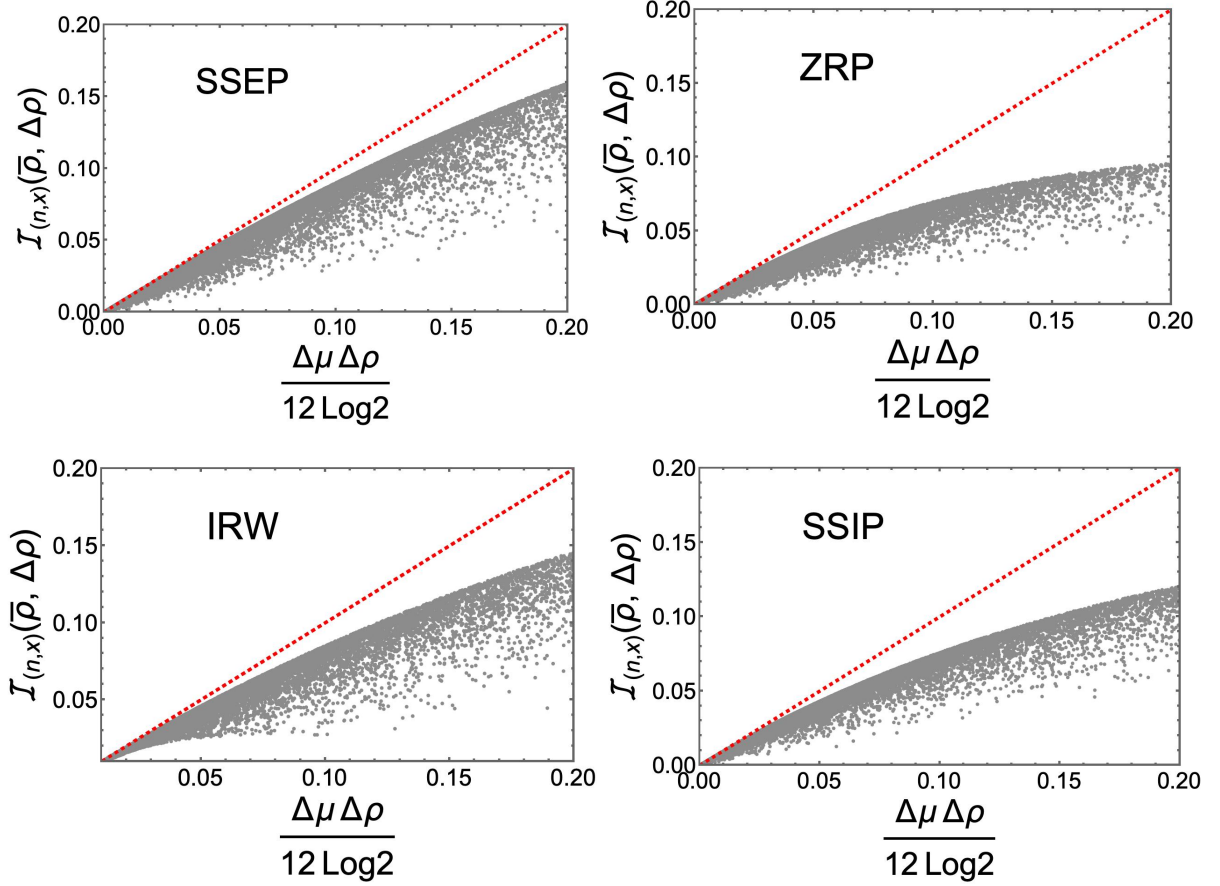
FIG. S3. Upper bound on the positional information in terms of the differences $\Delta\mu$ and $\Delta\rho$ for four different boundary-driven processes. For all cases, the gray symbols are the analytical results for different values of $\bar{\rho}$ and $\Delta\rho$, while the red line corresponds to the upper bound in (S28). Clearly, the bound is saturated in the near-equilibrium limit. We have set $k_B T = 1$.

Figure S4). Following [5], whenever steady state exists, the probability to find $n_i = n$ number of particles in the i-th site is given by

$$P(n\big|i) = \mathcal{N}_i(z_i)^n \prod_{n_k=1}^{n} \frac{1}{u_{n_k}}, \quad \text{with } z_i = \frac{\alpha_L}{\beta_L} - \left(\frac{\alpha_L}{\beta_L} - \frac{\alpha_R}{\beta_R}\right)\left(\frac{i-1}{N-1}\right), \tag{S55}$$

where $\mathcal{N}_k$ is the normalisation factor. For the case of $u_n = 1$, the probability $P(n\big|i)$ takes the form

$$P(n\big|i) = (z_i)^n(1-z_i), \quad \text{with } z_i = \frac{\alpha_L}{\beta_L} - \left(\frac{\alpha_L}{\beta_L} - \frac{\alpha_R}{\beta_R}\right)\left(\frac{i-1}{N-1}\right). \tag{S56}$$

As done for the SSEP, here again, we introduce the rescaled coordinate $x = i/N$ for large N and rewrite Eq. (S56) as

$$P(n\big|x) \simeq [z(x)]^n [1 - z(x)], \quad \text{with } z(x) = \frac{\alpha_L}{\beta_L} - \left(\frac{\alpha_L}{\beta_L} - \frac{\alpha_R}{\beta_R}\right)x. \tag{S57}$$

Since, we are interested in calculating the positional information in terms of the variables $\bar{\rho} = (\rho_L + \rho_R)/2$ and $\Delta\rho = (\rho_L - \rho_R)/2$, where $\rho_L = \langle n(0) \rangle$ and $\rho_R = \langle n(1) \rangle$ are average densities at the two ends

$$\rho_L = \frac{\alpha_L}{\beta_L - \alpha_L}, \quad \rho_R = \frac{\alpha_R}{\beta_R - \alpha_R}, \tag{S58}$$

we rewrite Eq. (S57) in the following manner

$$P(n\big|x) \simeq [z(x)]^n [1 - z(x)], \quad \text{with } z(x) = c_+ - x(c_+ - c_-). \tag{S59}$$

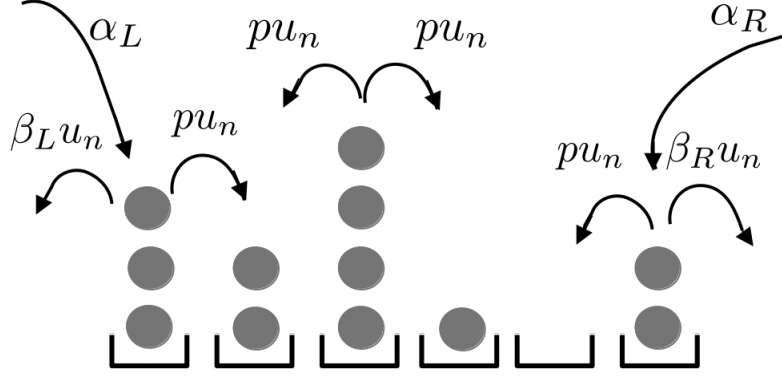FIG. S4. Zero-range process with open boundaries

where $c_\pm = (\bar\rho \pm \Delta\rho)/(1 + \bar\rho \pm \Delta\rho)$. For this model, the average density in the bulk takes a non-linear form

$$\rho(x) = \frac{\bar\rho(1+\bar\rho) - \Delta\rho\ (2x - 1 + \Delta\rho)}{(1+\bar\rho) + \Delta\rho\ (2x-1)}. \tag{S60}$$

In the remaining part of this section, we will demonstrate that even with this non-linear profile, the bound (S28) remains still valid. To see this, we use the probability in Eq. (S59) and calculate the entropy

$$S[P(n|x)] = -\sum_{n=0}^{\infty} P(n|x)\ \log_2 P(n|x) = -\frac{z(x)\log_2 z(x) + (1 - z(x))\log_2(1 - z(x))}{(1 - z(x))}. \tag{S61}$$

Taking average with respect to the prior $P_x(x) = 1$ gives

$$\begin{aligned}
&\langle S[P(n|x)]\rangle_x \\
&= -\int_0^1 dx\ \frac{z(x)\log_2 z(x) + (1 - z(x))\log_2(1 - z(x))}{(1 - z(x))}, \\
&= -\frac{1}{\ln 2\ (c_+ - c_-)}\Big[\mathrm{Li}_2(1 - c_+) - \mathrm{Li}_2(1 - c_-) - c_+\ln c_+ + c_-\ln c_- - (1 - c_+)\ln(1 - c_+) + (1 - c_-)\ln(1 - c_-)\Big],
\end{aligned} \tag{S62}$$

where $\mathrm{Li}_2(y)$ denotes the poly-logarithmic function. Next we calculate the other entropy term $S[P_n(n)]$ in Eq. (S6) for which we need the following probability

$$P_n(n) = \int_0^1 dx\ P(n|x)P_x(x) = \frac{f(n)}{(c_+ - c_-)}, \quad \text{with } f(n) = \frac{c_+^{n+1} - c_-^{n+1}}{n+1} - \frac{c_+^{n+2} - c_-^{n+2}}{n+2}. \tag{S63}$$

This yields

$$S[P_n(n)] = -\sum_{n=0}^{\infty} P_n(n)\log_2 P_n(n) = \log_2(c_+ - c_-) - \frac{1}{(c_+ - c_-)}\sum_{n=0}^{\infty} f(n)\log_2 f(n). \tag{S64}$$

Using Eqs. (S62) and (S64), the expression of the positional information turns out to be

$$\begin{aligned}
\mathcal{I}_{(n,x)}\left(\bar\rho, \Delta\rho\right) = \frac{1}{\ln 2\ (c_+ - c_-)}\Big[&\mathrm{Li}_2(1 - c_+) - \mathrm{Li}_2(1 - c_-) - c_+\ln c_+ + c_-\ln c_- - (1 - c_+)\ln(1 - c_+) \\
&+ (1 - c_-)\ln(1 - c_-) - \sum_{n=0}^{\infty} f(n)\ln f(n)\Big] + \log_2(c_+ - c_-).
\end{aligned} \tag{S65}$$

We have compared this result with the numerical simulations in Figure S1, and found a good agreement between them. Unlike in SSEP, the positional information for ZRP decreases monotonically with $\bar{\rho}$ and vanishes for large $\bar{\rho}$. For larger values of $\bar{\rho}$ but with fixed $\Delta\rho$, both boundary sites have a large number of particles available for hopping in the bulk, as there is no exclusion. Hence, in the steady-state, we expect the same number of particles in the bulk as well as in the boundaries. This can also be seen from Eq. (S60) where the density becomes independent of x for large enough $\bar{\rho}$. Therefore, the probability $P(x|n)$ is broad with respect to x, and consequently the entropy $S[P(x|n)]$ takes a large value. From Eq. (S4), this would mean that the positional information is small. In fact our study shows that, in processes with no exclusion, $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ exhibits a monotonic decay.

Our expression in Eq. (S65) is also consistent with the one derived with perturbative approach in Eq. (S24). This can be verified by expanding Eq. (S65) in $\Delta\rho$

$$\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho) \simeq \frac{\Delta\rho^2}{6\ln 2\,\bar{\rho}(1+\bar{\rho})} + \frac{[3 + \bar{\rho}(11 + 7\bar{\rho})]\,\Delta\rho^4}{180\ln 2\,\bar{\rho}^3(1+\bar{\rho})^3}. \tag{S66}$$

One essentially gets the same expression also from Eq. (S24) by plugging the central moments from Eq. (S59).

Now that we have obtained an exact expression for the positional information, we will employ it to test the bound in Eq. (S28). First, we compute the chemical potential difference using Eq. (S25) for which we need the first derivative of the free energy. In this regard, we follow a procedure similar to that of the SSEP and employ the fluctuation-response relation in Eq. (S23).

$$a_2(\bar{\rho}) = \frac{1}{\sigma_2(\bar{\rho})} = \frac{1}{\bar{\rho}(1+\bar{\rho})}, \implies a_1(\bar{\rho}) = \ln\left(\frac{\bar{\rho}}{1+\bar{\rho}}\right). \tag{S67}$$

The resulting expression for $\Delta\mu$ is now found to be

$$\Delta\mu = \ln\left(\frac{\bar{\rho} + \Delta\rho}{1 + \bar{\rho} + \Delta\rho}\right) - \ln\left(\frac{\bar{\rho} - \Delta\rho}{1 + \bar{\rho} - \Delta\rho}\right). \tag{S68}$$

We now utilize these exact expressions to plot $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ and $(\Delta\mu\,\Delta\rho)/(12\ln 2)$ for all possible values of $\bar{\rho}$ and $\Delta\rho$ in Figure S3. Across all these parameter values, we see that the bound (S28) is satisfied, and it is saturated in the near-equilibrium limit. To sum up, this section showcases an example of a model with non-linear density profile, where (S28) remains valid across all parameter values.

A. Independent random walkers

In our preceding analysis, we considered a special case of zero-range process where the rate $u_n = 1$ was chosen to be constant. We now provide another example of $u_n = n$ where an exact expression for the positional information can be obtained and the bound can be assessed. This example corresponds to the case of independent random walkers (IRW) [5]. We use $u_n = n$ in Eq. (S55) and obtain the conditional probability as

$$P(n|x) = \frac{[\rho(x)]^n}{n!}\exp\left[-\rho(x)\right], \quad \text{where } \rho(x) = \bar{\rho} - (2x - 1)\Delta\rho, \tag{S69}$$

where we have written the expression in terms of the x variable. The marginal probability $P_n(n)$ follows to be

$$P_n(n) = \frac{g(n)}{n!}, \quad \text{with } g(n) = \frac{\Gamma(n + 1, \bar{\rho} - \Delta\rho) - \Gamma(n + 1, \bar{\rho} + \Delta\rho)}{2\,\Delta\rho}. \tag{S70}$$

We now have both probabilities required in the definition of positional information in Eq. (S6). One can now proceed in the same manner as before and calculate the two entropies as

$$S[P_n(n)] = \sum_{n=0}^{\infty} P_n(n)\log_2 n! - \sum_{n=0}^{\infty} \frac{g(n)}{n!}\log_2 g(n), \tag{S71}$$

$$\langle S[P(n|x)]\rangle_x = \frac{1}{\ln 16\,\Delta\rho}\left[(\bar{\rho} - \Delta\rho)^2\ln(\bar{\rho} - \Delta\rho) - (\bar{\rho} + \Delta\rho)^2\ln(\bar{\rho} + \Delta\rho) + 6\bar{\rho}\,\Delta\rho\right] + \sum_{n=0}^{\infty} P_n(n)\log_2 n! \tag{S72}$$
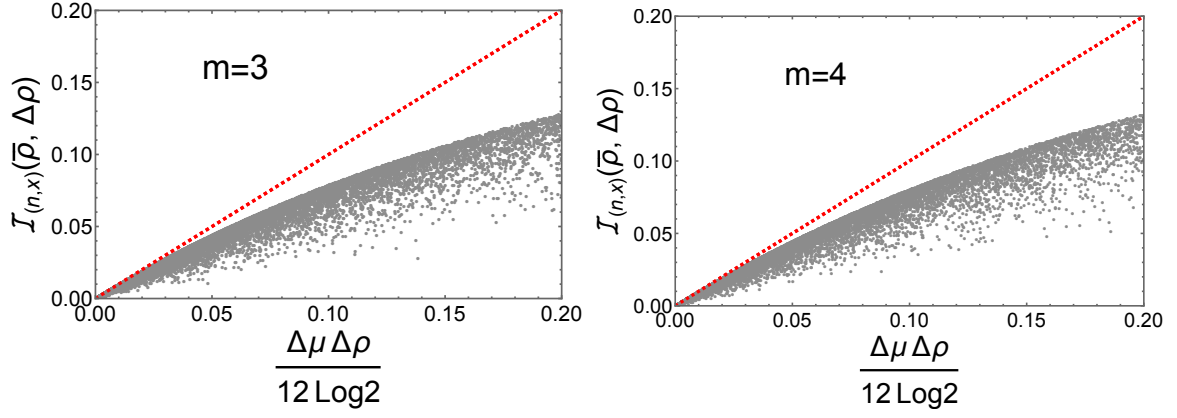
FIG. S5. Illustration of the upper bound on the positional information for SSIP model for two different values of the parameter m. The corresponding expressions of $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ and $\Delta\mu$ are given in Eqs. (S80) and (S81) respectively.

from which the expression of positional information can be calculated to be

$$\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho) = \frac{1}{\ln 16 \; \Delta\rho}\Big[\mathcal{Y}(\bar{\rho}, \Delta\rho) - 6\bar{\rho} \; \Delta\rho\Big] - \sum_{n=0}^{\infty} \frac{g(n)}{n!} \log_2 g(n), \tag{S73}$$

with functions $\mathcal{Y}(\bar{\rho}, \Delta\rho)$ and $g(n)$ defined respectively in Eqs. (S48) and (S70). Similarly, the chemical potential difference driving this system is

$$\Delta\mu = \ln(\bar{\rho} + \Delta\rho) - \ln(\bar{\rho} - \Delta\rho) \tag{S74}$$

We now have at our disposal all the quantities needed to test the upper bound (S28) on the positional information. Plotting $\mathcal{I}_{(n,x)}(\bar{\rho}, \Delta\rho)$ and $(\Delta\mu \; \Delta\rho)/(12\ln 2)$ in Figure S3, we have illustrated that the bound is satisfied even for this model. This model presents a third solvable example where the upper bound is valid.

V. POSITIONAL INFORMATION FOR THE BOUNDARY-DRIVEN SSIP

In all models examined thus far, the jump rate was either independent of the particle numbers or solely dependent on the occupation number of the source site from which the jump takes place. In this section, we will consider the SSIP model where the jump rate depends on the occupation numbers of both the source and the target sites [6, 7]. Like in ZRP, each lattice site is capable of accommodating an arbitrary number of particles. In the bulk site, a particle can jump from $i \to j$ [with $j = (i \pm 1)$] with a rate $pn_i(n_j + m)$ where m (> 0) is a parameter in the model. At the boundaries ($i = 1$ and $i = N$), a particle can be added with rate $\alpha_L(n_1 + m)$ for the left boundary and α_R ($n_N + m$) for the right one. Similarly, a particle, if present, can be removed from these sites with rates $\beta_L n_1$ and $\beta_R n_N$ respectively. The conditional probability to observe n particles at location x is [6]

$$P(n|x) = \left[\frac{m^m\Gamma(m+n)}{n! \; \Gamma(m)}\right] \frac{[\rho(x)]^n}{[m + \rho(x)]^{m+n}}, \quad \text{where } \rho(x) = \bar{\rho} - (2x - 1)\Delta\rho, \tag{S75}$$

from which the marginal probability $P_n(n)$ follows to be

$$P_n(n) = \frac{m \; \Gamma(m + n) \; \mathcal{S}(n)}{n! \; \Gamma(m)}, \quad \text{where} \tag{S76}$$

$$\mathcal{S}(n) = \frac{1}{2(n + 1)\Delta\rho}\left[\left(\frac{\bar{\rho} + \Delta\rho}{m}\right)^{n+1} {}_2F_1\left(1 + n, m + n; 2 + n; -\frac{\bar{\rho} + \Delta\rho}{m}\right) - \left(\frac{\bar{\rho} - \Delta\rho}{m}\right)^{n+1}\right.$$

$$\left. \times {}_2F_1\left(1 + n, m + n; 2 + n; -\frac{\bar{\rho} + \Delta\rho}{m}\right)\right]. \tag{S77}$$

Here $_2F_1\left(1+n, m+n; 2+n; z\right)$ stands for the hypergeometric function. Now that we possess both probabilities, we can compute the two entropies to be

$$S[P_n(n)] = -\left[\log_2 m + \sum_{n=0}^{\infty} P_n(n) \log_2\left(\frac{\Gamma(m+n)\ \mathcal{S}(n)}{n!\ \Gamma(m)}\right)\right], \tag{S78}$$

$$\langle S[P(n|x)]\rangle_x = \frac{1}{\ln 16\ \Delta\rho}\left[\mathcal{Y}(m+\bar{\rho}, \Delta\rho) - \mathcal{Y}(\bar{\rho}, \Delta\rho) - 2m\Delta\rho\right] - \sum_{n=0}^{\infty} P_n(n) \log_2\left(\frac{\Gamma(m+n)\ m^m}{n!\ \Gamma(m)}\right). \tag{S79}$$

Using these two expressions, the positional information can be calculated to be

$$\mathcal{I}_{(n,x)}\left(\bar{\rho}, \Delta\rho\right) = \frac{1}{\ln 16\ \Delta\rho}\left[\mathcal{Y}(\bar{\rho}, \Delta\rho) - \mathcal{Y}(m+\bar{\rho}, \Delta\rho) + 2m\ \Delta\rho\right] + (m-1)\log_2 m - \sum_{n=0}^{\infty} \frac{m\ \Gamma(m+n)\ \mathcal{S}(n)}{\Gamma(n+1)\Gamma(m)} \log_2 \mathcal{S}(n), \tag{S80}$$

with functions $\mathcal{Y}(\bar{\rho}, \Delta\rho)$ and $\mathcal{S}(n)$ given respectively in Eq. (S48) and (S77). Similarly, the chemical potential difference driving the system is

$$\Delta\mu = \ln\left(\frac{\bar{\rho} + \Delta\rho}{m + \bar{\rho} + \Delta\rho}\right) - \ln\left(\frac{\bar{\rho} - \Delta\rho}{m + \bar{\rho} - \Delta\rho}\right). \tag{S81}$$

We are now equipped with all the necessary elements to assess the upper bound (S28) on positional information. In Figure S3, we have again plotted $\mathcal{I}_{(n,x)}\left(\bar{\rho}, \Delta\rho\right)$ and $\Delta\mu\ \Delta\rho/12\ln 2$ for $m = 2$ for different values of $\Delta\rho$ and $\bar{\rho}$. Across all these values, we once again find that our bound is satisfied. We have also verified this for other values of m in Figure S5. This consistency further reinforces the conjecture that our bound holds for general boundary-driven systems.

VI. PROOF OF THE RELATIONS IN EQ. (S23)

For a system in equilibrium, we stated some relations in Eq. (S23) that give central moments of the density in terms of the derivatives of the associated free energy. Here, we present a mathematical proof of these relations. Note that the probability distribution to observe a density n at equilibrium is given by

$$P_{\mathrm{eq}}(n) = \frac{e^{-[a(n)-a(\bar{\rho})-\bar{\mu}(n-\bar{\rho})]}}{\int dn\ e^{-[a(n)-a(\bar{\rho})-\bar{\mu}(n-\bar{\rho})]}}, \quad \text{with } k_B T = 1, \tag{S82}$$

where $\bar{\rho} = \langle n \rangle_{\mathrm{eq}}$ is the average density, and $\bar{\mu}$ is the chemical potential related to $\bar{\rho}$ by

$$\bar{\mu} = \frac{da(\bar{\rho})}{d\bar{\rho}}. \tag{S83}$$

Eq. (S82) can be further simplified as

$$P_{\mathrm{eq}}(n) = \frac{\exp\left[-a(n) + \bar{\mu}n\right]}{Z_{\mathrm{P}}}, \quad \text{with } Z_{\mathrm{P}} = \int dn\ \exp\left[-a(n) + \bar{\mu}n\right]. \tag{S84}$$

Any moment of the density can be computed by taking the derivative of the partition function Z_{P}

$$\langle n^k \rangle_{\mathrm{eq}} = \frac{1}{Z_{\mathrm{P}}} \frac{\partial^k Z_{\mathrm{P}}}{\partial\bar{\mu}^k}. \tag{S85}$$

A. Variance of n

Putting $k = 2$ in Eq. (S85), we get the second moment

$$\langle n^2 \rangle_{\mathrm{eq}} = \frac{1}{Z_{\mathrm{P}}} \frac{\partial^2 Z_{\mathrm{P}}}{\partial\bar{\mu}^2} = \frac{\partial}{\partial\bar{\mu}}\left[\frac{1}{Z_{\mathrm{P}}} \frac{\partial Z_{\mathrm{P}}}{\partial\bar{\mu}}\right] + \left[\frac{1}{Z_{\mathrm{P}}} \frac{\partial Z_{\mathrm{P}}}{\partial\bar{\mu}}\right]^2. \tag{S86}$$

Now identifying $\left[\frac{1}{Z_P}\frac{\partial Z_P}{\partial \mu_L}\right] = \bar{\rho}$, one can rewrite the above expression as

$$\sigma_2(\bar{\rho}) \equiv \langle n^2 \rangle_{\text{eq}} - \bar{\rho}^2 = \frac{d\bar{\rho}}{d\bar{\mu}}. \tag{S87}$$

To simplify further, we take $\bar{\mu}$ from Eq. (S83) and take its derivative with $\bar{\rho}$ to get $a_2(\bar{\rho}) = \frac{d\bar{\mu}}{d\bar{\rho}}$. Plugging this in Eq. (S87) yields

$$\sigma_2(\bar{\rho}) = \frac{1}{a_2(\bar{\rho})}. \tag{S88}$$

This relation gives us the equilibrium variance of the density in terms of the underlying free energy of the system.

B. Third central moment of n

We next look at the third moment for which we put $k = 3$ in Eq. (S85). This gives

$$\langle n^3 \rangle_{\text{eq}} = \frac{1}{Z_P}\frac{\partial^3 Z_P}{\partial \bar{\mu}^3} = \frac{1}{Z_P}\frac{\partial}{\partial \bar{\mu}}\left[Z_P \langle n^2 \rangle_{\text{eq}}\right],$$

$$= \bar{\rho}\langle n^2 \rangle_{\text{eq}} + \frac{\partial \langle n^2 \rangle_{\text{eq}}}{\partial \bar{\mu}}. \tag{S89}$$

To evaluate the derivative with $\bar{\mu}$, we use Eqs. (S83) and (S87) as $\frac{\partial \langle n^2 \rangle_{\text{eq}}}{\partial \bar{\mu}} = \frac{\partial \langle n^2 \rangle_{\text{eq}}}{\partial \bar{\rho}}\frac{d\bar{\rho}}{d\bar{\mu}} = \sigma_2(\bar{\rho})\frac{\partial \langle n^2 \rangle_{\text{eq}}}{\partial \bar{\rho}}$. Furthermore, we also write $\langle n^2 \rangle_{\text{eq}} = \sigma_2(\bar{\rho}) + \bar{\rho}^2$. The expression of $\langle n^3 \rangle_{\text{eq}}$ then becomes

$$\langle n^3 \rangle_{\text{eq}} = 3\bar{\rho}\,\sigma_2(\bar{\rho}) + \bar{\rho}^3 + \sigma_2(\bar{\rho})\frac{\partial \sigma_2(\bar{\rho})}{\partial \bar{\rho}}. \tag{S90}$$

With this expression, the third central moment $\sigma_3(\bar{\rho}) = \langle (n - \bar{\rho})^3 \rangle_{\text{eq}}$ turns out to be

$$\sigma_3(\bar{\rho}) = \sigma_2(\bar{\rho})\,\frac{\partial \sigma_2(\bar{\rho})}{\partial \bar{\rho}} = -\sigma_2(\bar{\rho})^3\,a_3(\bar{\rho}), \tag{S91}$$

where for the second equality, we have used Eq. (S88). Proceeding in the same, one can show that the fourth central moment is

$$\sigma_4(\bar{\rho}) = -\sigma_2(\bar{\rho})^4\,a_4(\bar{\rho}) + \frac{3}{\sigma_2(\bar{\rho})}\left[\sigma_3(\bar{\rho})^2 + \sigma_2(\bar{\rho})^3\right]. \tag{S92}$$

[1] J. O. Dubuis, G. Tkačik, E. F. Wieschaus, T. Gregor, and W. Bialek, Positional information, in bits, Proceedings of the National Academy of Sciences **110**, 16301 (2013).

[2] L. Wolpert, Positional information and the spatial pattern of cellular differentiation, Journal of Theoretical Biology **25**, 1 (1969).

[3] B. Derrida, J. L. Lebowitz, and E. R. Speer, Free energy functional for nonequilibrium systems: An exactly solvable case, Physical Review Letters **87**, 150601 (2001).

[4] B. Derrida, J. L. Lebowitz, and E. R. Speer, Large deviation of the density profile in the steady state of the open symmetric simple exclusion process, Journal of Statistical Physics **107**, 599 (2002).

[5] E. Levine, D. Mukamel, and G. M. Schütz, Zero-range process with open boundaries, Journal of Statistical Physics **120**, 759 (2005).

[6] K. Vafayi and M. H. Duong, Weakly nonequilibrium properties of a symmetric inclusion process with open boundaries, Physical Review E **90**, 052143 (2014).

[7] C. Franceschini, P. Gonçalves, and F. Sau, Symmetric inclusion process with slow boundary: Hydrodynamics and hydrostatics, Bernoulli **28**, 1340 (2022).