Large deviations of a tracer in the symmetric exclusion process

T. Imamura, K. Mallick, T. Sasamoto

Large deviations of a tracer in the symmetric exclusion process

Takashi Imamura,[1] Kirone Mallick,[2] and Tomohiro Sasamoto[3]

[1] *Department of Mathematics and Informatics, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan*
[2] *Institut de Physique Théorique, CEA Saclay and URA 2306, CNRS, 91191 Gif-sur-Yvette cedex, France*
[3] *Department of Physics, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8551, Japan*

(Dated: February 21, 2017)

The one-dimensional symmetric exclusion process, the simplest interacting particle process, is a lattice-gas made of particles that hop symmetrically on a discrete line respecting hard-core exclusion. The system is prepared on the infinite lattice with a step initial profile with average densities ρ_+ and ρ_- on the right and on the left of the origin. When $\rho_+ = \rho_-$, the gas is at equilibrium and undergoes stationary fluctuations. When these densities are unequal, the gas is out of equilibrium and will remain so forever. A tracer, or a tagged particle, is initially located at the boundary between the two domains; its position X_t is a random observable in time, that carries information on the non-equilibrium dynamics of the whole system. We derive an exact formula for the cumulant generating function and the large deviation function of X_t, in the long time limit, and deduce the full statistical properties of the tracer's position. The equilibrium fluctuations of the tracer's position, when the density is uniform, are obtained as an important special case.

PACS numbers: 05.40.-a,05.60.-k

The collective dynamics of a complex system can be probed by attaching a neutral tag to a particle, that does not alter its interactions with the environment, and by monitoring the position of the tagged particle in time. This technique is a powerful tool to study flows in material sciences, biological systems and even social groups (see e.g., [1–4] and references therein). The averaged trajectory of a tracer carries information on the overall motion of the fluid whereas its fluctuations are sensitive to the statistical properties of the medium. The canonical example is the Brownian motion of a grain of pollen immersed in water at thermal equilibrium, and the simplest model for this diffusion is given by independent random walkers symmetrically hopping on a lattice; the position of any walker, as a function of time t, spreads as $\sqrt{t}$. In presence of weak-interactions, diffusive behavior generically prevails but the amplitude of the spreading, measured by the diffusion constant, is a function of the total density of particles [1, 5, 6].

If the interactions induce long-range correlations either in space or time direction, or if the environment is out of equilibrium (by carrying some internal currents), the motion of a tagged particle can exhibit unusual statistical properties such as anomalous diffusion and/or non-Gaussian fluctuations. For example, a tracer trapped in a linear array of convection rolls spreads only as $t^{1/3}$ with time [7, 8]. Correlations are usually enhanced in low dimensional systems such as narrow quasi-one-dimensional channels, in which the order amongst the particles is preserved because of steric hindrance. For such a single-file motion, the typical displacement X_t of a tracer at large times grows as $t^{1/4}$, which is much slower than the usual $\sqrt{t}$ law, regardless of the precise form of the interaction. However, collective diffusion of local density fluctuations remains normal and behaves as $\sqrt{t}$. Similarly, the time-integrated current at a given location of a single-file channel also displays $t^{1/4}$-fluctuations. This anomalous *single-file diffusion* has been demonstrated in various experiments involving different types of physical systems such as zeolites, capillary pores, carbon nanotubes or colloids [9–15]. Single-file diffusion is also discussed in numerous theoretical papers, at various levels of physical intuition [16] or mathematical rigor [5, 17, 18].

One of the simplest models in non-equilibrium statistical physics is the Symmetric Exclusion Process (SEP) [17], a lattice gas of particles performing symmetric random walks in continuous time and interacting by hard-core exclusion: each particle attempts to hop with rate unity from its location to an empty neighboring site; double occupancy of a site is forbidden. Thanks to the wealth of analytical knowledge accumulated during the last few decades, this process and its variants, are used as paradigms in non-equilibrium statistical mechanics [6, 19–22]. In a one dimensional lattice, the SEP is a pristine model of a single-file diffusion, amenable to quantitative analysis. In equilibrium case with uniform density ρ, the variance of the position X_t of a tagged particle initially located at the origin is given, in the long time limit, by [5, 17]

$$\langle X_t^2 \rangle = \frac{2(1-\rho)}{\rho} \sqrt{\frac{t}{\pi}}. \tag{1}$$

It has also been proved that the *rescaled* position $\frac{X_t}{t^{1/4}}$ satisfies a central limit theorem and converges to a fractional Brownian motion with Hurst index 1/4 [17, 23].

The full distribution of X_t and its higher cumulants are, however, not known. The tracer, being immersed in fluctuating environment, far from equilibrium, can display large and non-typical excursions. Such rare events are quantified by a large deviation function [24, 25]. Large deviation functions appear as appropriate candidates for macroscopic potentials under non-equilibrium conditions. Moreover, the fluctuation theorem, which is one of the few exact results valid far from thermodynamic equilibrium, is stated as a property of large deviation functions [26, 27]. It emphasizes the role of microscopic time-reversal symmetry for macroscopic fluctuations. In present day statistical physics, large deviations play an increasingly important role [21, 22, 28, 29].

Recently, the large deviation principle for the tracer position has been proved rigorously [30]: when $t \to \infty$ there exists a large-deviation function $\phi(\xi)$, such that

$$\mathrm{Prob}\left(\frac{X_t}{\sqrt{4t}} = -\xi\right) \sim \exp[-\sqrt{t}\phi(\xi)]. \tag{2}$$

Note the prefactor $\sqrt{t}$ in the exponent; for non-interacting particles, the prefactor would be t [31]. Alternatively, one studies the characteristic function of X_t, which behaves as

$$\langle e^{sX_t}\rangle \sim e^{-\sqrt{t}C(s)} \quad \text{when} \quad t \to \infty. \tag{3}$$

The Taylor expansion of the cumulant generating function $C(s)$ with respect to s generates all the cumulants of X_t. The functions $C(s)$ and $\phi(\xi)$ are related by Legendre transform [24, 25]:

$$C(s) = \min_{\xi}\left(2s\xi + \phi(\xi)\right). \tag{4}$$

Each of these functions carries information on the long time behavior of the process. Although the SEP has been studied for more than 40 years, analytic formulas for these functions are not yet known.

In this letter, we shall present an exact formula for the large-deviation function $\phi(\xi)$ in (2) of the tracer position in the SEP. As an initial condition, we prepare a step density profile with an average density ρ_+ on the right of the origin and ρ_- on the left. (See the right figure in Fig. 2.) In a parametric representation, $\phi(\xi)$ is given by

$$\begin{cases} \phi(\xi) & = \mu(\xi, \lambda^*), \\ \frac{\partial \mu(\xi,\lambda^*)}{\partial \lambda} & = 0, \end{cases} \tag{5}$$

where the second equation defines implicitly $\lambda^* = \lambda^*(\xi)$, and $\mu(\xi, \lambda)$ is

$$\mu(\xi,\lambda) = \sum_{n=1}^{\infty} \frac{(-\omega)^n}{n^{3/2}} A(\sqrt{n}\,\xi) + \xi \log \frac{1+\rho_+(e^{\lambda}-1)}{1+\rho_-(e^{-\lambda}-1)}. \tag{6}$$

Here $\omega = r_+(e^{\lambda}-1) + r_-(e^{-\lambda}-1)$ with $r_\pm = \rho_\pm(1-\rho_\mp)$ and

$$A(\xi) = \frac{e^{-\xi^2}}{\sqrt{\pi}} + \xi(1 - \mathrm{erfc}(\xi)), \tag{7}$$

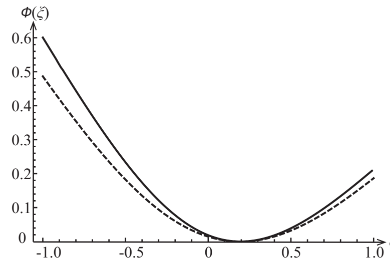

FIG. 1. The large deviation function $\phi(\xi)$ of the tracer position in the SEP (solid curve) for the case $\rho_+ = 0.3, \rho_- = 0.15$. The one for the reflective Brownian particles (14) with the same $\rho_\pm$ is also shown (dashed curve).

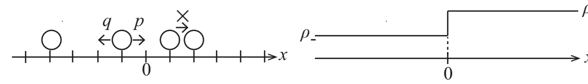

FIG. 2. ASEP. Left: Particles hop asymmetrically on the lattice under volume exclusion. Right: Step initial condition with densities ρ_+ (resp. ρ_-) to the right (resp. left).

where the complementary error function is defined by $\mathrm{erfc}(z) = \frac{2}{\sqrt{\pi}}\int_z^{\infty} e^{-u^2} du$. This is the central result in this letter. For $\xi = 0$, we have $A(0) = 1/\sqrt{\pi}$ and $\mu(0,\lambda)$ reduces to the expression found in [32] for the current fluctuations in the SEP at the origin. Our formula (6) generalizes it and leads us to a complete analytic description of the statistical properties of the tracer in the long time limit. The figure is also easily drawn, see Fig 1.

To explain the meaning of μ and the derivation of our formula, we first recall the set-up of the asymmetric simple exclusion process (ASEP), see Fig. 2. The position of a particle is labeled by an integer $x \in \mathbb{Z}$. Particles hop to the right and to the left with rates p and q, respectively. The asymmetry parameter is $\tau = p/q$ with $0 \le \tau \le 1$. The symmetric model, which is the main target of our study, corresponds to $p = q = \tau = 1$. We adopt the convention that a current flowing from right to left is counted positively [33, 34]. The initial condition is the step density profile with ρ_+ and ρ_-. Typically, we have $\rho_+ \ge \rho_-$. The stationary case corresponds to $\rho_+ = \rho_- = \rho$. We emphasize that the initial profile displays randomness: statistical averages will be taken both over the dynamics and the initial conditions. The tracer is defined to be the particle in the region $x > 0$, which is initially the closest to the origin. (For quantities in the long time, which we are interested in this article, this is equivalent to putting the tracer at the origin at $t = 0$.) Its position at time t is denoted by X_t.

In order to study the position of the continuously moving tracer, it is useful to relate X_t to a local observable. Let N_t denote the integrated current through the bond $(0, 1)$ for the duration $[0, t]$; i.e., N_t is equal to the total number of particles having hopped from 1 to 0 minus the total number of particles having hopped from 0 to 1 during the time interval $[0, t]$. We define the following quantity [35, 36]

$$N(x,t) = N_t + \begin{cases} +\sum_{y=1}^{x} \eta_y(t)\,, & x > 0, \\ 0, & x = 0, \\ -\sum_{y=x+1}^{0} \eta_y(t)\,, & x < 0, \end{cases} \tag{8}$$

where $\eta_x(t) = 1$(resp. 0) when the site x is occupied (resp. empty) at time t. We recall that the observable $h(x,t) = N(x,t) - x/2$ is the local height function appearing when the ASEP is mapped to a growth process [6, 37, 38].

Using particle number conservation, one can verify [31] that the tagged particle position X_t and $N(x,t)$ satisfy

$$\mathbb{P}[X_t \leq x] = \mathbb{P}[N(x,t) > 0]\,. \tag{9}$$

This identity will allow us to relate the statistical properties of the tracer X_t with those of the height $N(x,t)$ and, in particular, to express the cumulant generating function and the large deviation function of X_t in terms of the corresponding quantities for $N(x,t)$.

In the long time limit, the characteristic function of $N(x,t)$ behaves as

$$\langle e^{\lambda N(x,t)} \rangle \sim e^{-\sqrt{t}\mu(\xi,\lambda)} \tag{10}$$

with $\xi = -\frac{x}{\sqrt{4t}}$ and its cumulants are obtained by expanding $\mu(\xi,\lambda)$ with respect to λ. This $\mu(\xi,\lambda)$ is nothing but the one in (6). From the identity (9), we see [31] that the large deviation functions of X_t is given through the characteristic function of $N(x,t)$ as $\phi(\xi) = \max_\lambda \mu(\xi,\lambda)$, which is equivalent to (5).

We now investigate some properties of the above formulas and extract some concrete results from them. We also retrieve and generalize some results previously known in certain particular cases.

The tracer's large deviation function $\phi(\xi)$ satisfies a version of the Fluctuation Theorem [26, 27],

$$\phi(\xi) - \phi(-\xi) = 2\xi \log \frac{1-\rho_+}{1-\rho_-}. \tag{11}$$

The Fluctuation Theorem is a symmetry relation that originates from an underlying time-reversal invariance. It implies, in particular, that the Einstein relation is true for the SEP [39]. The proof of (11) is based on the fact that $\lambda^*(-\xi) = \log\frac{r_-}{r_+} - \lambda^*(\xi)$ [31]. We also note that, while the fluctuation theorems have been established mainly for a large system in the infinitely late time, ours is for a system on the infinite lattice and for a large time.

Explicit formulae for the first few cumulants of X_t can be obtained by substituting an expression of $\phi(\xi)$ in (5) into (4). For a stationary initial condition, $\rho_+ = \rho_- = \rho$, we have calculated the first few cumulants: the variance is given by (1) and at the fourth order, we find

$$\frac{\langle X_t^4 \rangle_c}{\sqrt{4t}} = \frac{1-\rho}{\sqrt{\pi}\rho^3}[1 - (4 - (8 - 3\sqrt{2})\rho)(1-\rho) + \frac{12}{\pi}(1-\rho)^2]$$

(the subscript c indicates a cumulant), in agreement with calculations based on the Macroscopic Fluctuation Theory (MFT) [40]. Considering the MFT is a description at the level of hydrodynamics, this coincidence provides a highly nontrivial check of the MFT. The procedure can be carried out to higher orders in s [31].

For non-equilibrium initial conditions, $\rho_+ > \rho_- > 0$, the tracer drifts away from the origin as

$$\frac{\langle X_t \rangle}{\sqrt{4t}} = -\xi_0, \tag{12}$$

where ξ_0 is the unique solution of

$$2\xi_0\rho_- = (\rho_+ - \rho_-)\int_{\xi_0}^{\infty} \mathrm{erfc}(u)\mathrm{d}u. \tag{13}$$

Solving (5) around ξ_0 leads to the variance of the tracer

$$\mathrm{Var}(X_t) = \frac{4K(\rho_+ - \rho_-)^2 A(\xi_0)\sqrt{t}}{(\rho_+\mathrm{erfc}(\xi_0) + \rho_-\mathrm{erfc}(-\xi_0))^2}$$

with

$$K = \frac{\rho_+^3 + \rho_-^3 - 3\rho_+^2\rho_- - 3\rho_+\rho_-^2 + 4\rho_+\rho_-}{(\rho_+ + \rho_-)(\rho_+ - \rho_-)^2} - \frac{A(\sqrt{2}\,\xi_0)}{\sqrt{2}A(\xi_0)}.$$

In the special case $\rho_- = 0$, the tracer is the leftmost particle of a SEP expanding in a half-empty space and finding the distribution of X_t becomes identical to a problem in extreme value statistics. From the above expressions, it can be shown that $\langle X_t \rangle \sim \sqrt{t\log t}$ and $\mathrm{Var}(X_t) \sim \frac{t}{\log t}$. The tracer follows a Gumbel law, which is well-known to appear for independent walkers, in spite of interaction effects in the SEP [17, 41].

In the low density limit $\rho_-, \rho_+ \ll 1$, the SEP becomes equivalent to an ensemble of reflecting Brownian particles [17]. This system can be viewed as independent Brownian motions that exchange their labels when they collide and has been solved exactly using various techniques. Retaining only the first order terms in $\rho_\pm$ in the formula (6), and using (5), we obtain the large deviation of a tracer in the reflecting Brownian limit:

$$\phi(\xi) = \left\{\sqrt{\rho_+\Xi(\xi)} - \sqrt{\rho_-\Xi(-\xi)}\right\}^2 \tag{14}$$

where $\Xi(\xi) = \int_\xi^\infty \mathrm{erfc}(u)\mathrm{d}u$. This generalizes the known result in the uniform case $\rho_+ = \rho_-$ [16, 40, 42, 43]. A figure of this large deviation function is also drawn in Fig. 1. By comparing to the one for the SEP, the effect of interaction among particles of the SEP is clearly seen.

In the last part of this work, we outline the derivation of the main formula (6). The strategy is to find exact expressions for all the moments of $N(x,t)$ and then construct the cumulant generating function $\mu(\xi,\lambda)$. The time evolution equations for the moments of $N(x,t)$ form a hierarchy of coupled differential equations that must be solved simultaneously. This seems to be a daunting task.

Our strategy is to make a detour though the ASEP, with $\tau<1$, for which the observable $N_\tau(x,t)$, defined in (8), satisfies a remarkable *self-duality* property [34–36]. For $x_1<x_2<\ldots<x_n$, n-point correlations of the type

$$\phi(x_1,\ldots,x_n;t)=\langle \tau^{N_\tau(x_1,t)}\ldots\tau^{N_\tau(x_n,t)}\rangle$$

follow the same dynamical equations as the ASEP with a finite number n of particles located at $x_1,\ldots,x_n$. Using the fact that the ASEP with n particles is solvable by Bethe Ansatz, these τ-correlations can be expressed as a multiple contour integral in the complex plane [33, 34, 44]. For the step initial condition with the densities $\rho_\pm$, we can write

$$\langle \tau^{nN_\tau(x,t)}\rangle=\tau^{-n\frac{x}{2}}\tau^{n(n-1)/2}\prod_{i=1}^{n}\left(1-\frac{r_-}{\tau^i r_+}\right)$$
$$\times\int\cdots\int\prod_{i<j}\frac{z_i-z_j}{z_i-\tau z_j}\prod_{i=1}^{n}\frac{F_{x,t}(z_i)}{(1-\frac{z_i}{\tau\theta_+})(z_i-\theta_-)}\mathrm{d}z_i \quad (15)$$

with $r_\pm$ defined below (6), $\theta_\pm=\rho_\pm/(1-\rho_\pm)$ and

$$F_{x,t}(z)=\left(\frac{1+z}{1+z/\tau}\right)^x e^{-\frac{q(1-\tau)^2z}{(1+z)(\tau+z)}t}.$$

The contour of z_i include $-1,\tau\theta_+$ and $\{\tau z_j\}_{j>i}$ but not $-\tau,\theta_-$; integrations are performed from z_n down to z_1, see Fig. 3. This contour formula is a generalization of the $\rho_-=0$ case studied in [34]. See also a recent related work [45]. The symmetric limit, $\epsilon=1-\tau\to 0$, is performed using the identity

$$\sum_{j=0}^{n}(-1)^j\binom{n}{j}\langle\tau^{(n-j)N_\tau}\rangle=\langle(1-\tau^{N_\tau})^n\rangle=\epsilon^n\langle N^n\rangle+o(\epsilon^n)$$

that relates the τ-moments of $N_\tau(x,t)$ in the ASEP to the n-th moment of the observable $N(x,t)$ in the SEP. Each term on the left-hand side is given by a complex contour integral, that has to be expanded with respect to ϵ. This is achieved first by evaluating the residues of the contour integrals at the poles in the vicinity of θ_+, leading to a formula in the form,

$$\langle(1-\tau^{N_\tau})^n\rangle=\sum_{k=0}^{n}\mu_{n,k}(\epsilon)J_k\epsilon^k$$

where the combinatorial coefficients $\mu_{n,k}(\epsilon)$ contain the contributions of the residues and the J_k's are k-fold integrals localized around the origin. Then, explicit recursive relations for the $\mu_{n,k}(\epsilon)$'s are found and large time

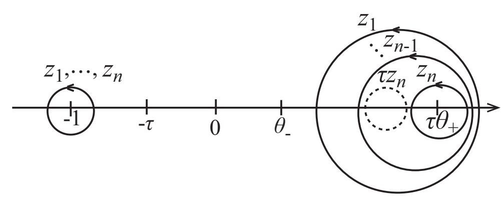

FIG. 3. The integration contours in (15).

asymptotics of the J_k's are extracted. This allows us, finally, to obtain a formula for the n-th moment of $N(x,t)$ and for its n-th cumulant. The expressions for the cumulants are given by

$$\frac{\langle N(x,t)^n\rangle_c}{\sqrt{t}}\sim\sum_{l=1}^{n}\frac{\alpha_{n,l}(r_+,r_-)}{\sqrt{l}}\Xi(-\sqrt{l}\xi)-2\alpha_{n,l}(1,0)\xi\rho_+^l \quad (16)$$

where $\Xi(\xi)$ was defined below (14) and

$$\frac{\alpha_{n,l}(a,b)}{(l-1)!}=(-1)^l\sum_{\substack{\sum jl_j=n\\ \sum l_j=l}}\frac{n!}{\prod_{j=1}^{n}l_j!}\prod_{j=1}^{n}\left(\frac{a+(-1)^jb}{j!}\right)^{l_j}. \quad (17)$$

Taking the generating function of the cumulants leads to (6). The full details of the derivation will be given in [46].

In this work, we obtain the exact formula for the large deviations of a tracer in the one dimensional symmetric simple exclusion process. This formula yields all the cumulants of the tracer position, in the long time limit. This answers a problem that has eluded solution for years [17, 30]. Our results are valid both when the system is at equilibrium with uniform density, and when the system is out of equilibrium, starting with a step density profile, the tracer being initially located at the boundary between the two domains of unequal density. Some of our formulas for the cumulants are prone to experimental tests, e.g. using colloidal particles [13]. They can also be used as benchmarks for numerical methods to evaluate large deviations, such as the one proposed in [47].

The derivation of the central formula (6) uses the powerful mathematical arsenal of integrable probabilities developed to solve the one-dimensional Kardar-Parisi-Zhang (KPZ) equation, the ASEP and related asymmetric models [34, 44, 48–51]. Generalizations of the ideas and techniques in this article will allow us to reveal various intricate properties of the SEP and related symmetric models, which would have been difficult with other means.

Infinite systems out of equilibrium keep in general the memory of the initial conditions [16]. For the models in the KPZ universality class, it has been well established that different initial conditions can lead to different statistical laws in the long time limit [38, 51–53]. This must also be true in the tagged particle problem in the SEP and one would like to study more general set-ups than the step profile. In particular, instead of taking averages over an ensemble of fluctuating initial step profiles and over the dynamics (*annealed case*), one could start with a deterministic initial configuration and average only over the history of the process (*quenched case*). For the latter case, even less is known [32, 43, 54] compared with the former, but new progress is expected to be achieved by extending our approach, combined with results for the ASEP, e.g. [55].

Finally, we would like to relate our derivation to the macroscopic fluctuation theory (MFT) [21, 29], one of the most promising approaches to study systems far from equilibrium. The MFT is based on a variational principle, that determines the optimal path that produces a given fluctuation, leading to two coupled nonlinear Euler-Lagrange equations. For reflecting Brownian particles, these equations can be linearized and solved, leading to the large deviations of the tracer [54]. However, for the symmetric exclusion process, the MFT equations are, for the moment, intractable. Our exact calculations may give some hint to solve the MFT equations for this nonlinear case.

The authors are grateful to S. Mallick for a careful reading of the manuscript and to P. L. Krapivsky for interesting discussions. We thank the JSPS core-to-core program "Non-equilibrium dynamics of soft matter and information" which initiated this work. Parts of this work were performed during stays at ICTS Bangalore and at KITP Santa Barbara. This research was supported in part by the National Science Foundation under Grant No. NSF PHY11-25915. The works of T.I and T.S. are also supported by JSPS KAKENHI Grant Numbers JP25800215, JP16K05192 and P25103004, JP14510499, JP15K05203, JP16H06338 respectively.

[1] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, 2000).
[2] J. Kärger and D. Ruthven, *Diffusion in zeolites and other microporous solids* (Wiley, 1992).
[3] R. Y. Tsien, Annu. Rev. Biochem. **67**, 509 (1998).
[4] A. Schadschneider, D. Chowdhury, and K. Nishinari, *Stochastic Transport in Complex Systems* (Elsevier, 2010).
[5] H. Spohn, *Large Scale Dynamics of Interacting Particles* (Springer-Verlag, New York, 1991).
[6] P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A Kinetic View of Statistical Physics* (Cambridge University Press, 2010).
[7] W. Young, A. Pumir, and Y. Pomeau, Phys. Fluids A **1**, 462 (1989).
[8] J.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).
[9] A. L. Hodgkin and R. D. Keynes, Journal of physiology **128**, 61 (1955).
[10] E. J. Lea, J. Theor. Biol. **5**, 102 (1963).
[11] V. Kukla, J. Kornatowski, D. Demuth, I. Girnus, H. Pfeifer, L. V. C. Rees, S. Schunk, K. K. Unger, and J. Karger, Science **272**, 702 (1996).
[12] T. Chou and D. Lohse, Phys. Rev. Lett. **82**, 3552 (1999).
[13] Q.-H. Wei, C. Bechinger, and P. Leiderer, Science **287**, 625 (2000).
[14] M. Kollmann, Phys. Rev. Lett. **90**, 180602 (2003).
[15] B. Lin, M. Meron, B. Cui, S. A. Rice, and H. Diamant, Phys. Rev. Lett. **94**, 216001 (2005).
[16] D. Levitt, Phys. Rev. A **8**, 3050 (1973); J. Percus, *ibid.* **9**, 557 (1974); P. M. Richards, Phys. Rev. B **16**, 1393 (1977); S. Alexander and P. Pincus, *ibid.* **18**, 2011 (1978); H. van Beijeren, K. Kehr, and R. Kutner, *ibid.* **28**, 5711 (1983); S. N. Majumdar and M. Barma, *ibid.* **44**, 5306 (1991); C. Rödenbeck, J. Kärger, and K. Hahn, Phys. Rev. E **57**, 4382 (1998); N. Leibovich and E. Barkai, *ibid.* **88**, 032107 (2013); P. Illien, O. Bénichou, C. Mejia-Monasterio, G. Oshanin, and R. Voituriez, Phys. Rev. Lett. **111**, 038102 (2013); L. Lizana, M. A. Lomholt, and T. Ambjörnsson, Physica A **395**, 148 (2014).
[17] T. E. Harris, J. Appl. Prob. **2**, 323 (1965); F. Spitzer, Adv. in Math. **5**, 246 (1970); R. Arratia, Ann. Prob. **11**, 362 (1983); H. Spohn, J. Stat. Phys. **59**, 1227 (1990).
[18] T. M. Liggett, *Interacting Particle Systems* (Springer, 2004).
[19] S. Katz, J. L. Lebowitz, and H. Spohn, J. Stat. Phys. **34**, 497 (1984).
[20] J. L. Lebowitz, E. Presutti, and H. Spohn, J. Stat. Phys. **51**, 841 (1988).
[21] B. Derrida, J. Stat. Mech. **P07023** (2007).
[22] B. Derrida, J. Stat. Mech. **P01030** (2011).
[23] M. Peligrad and S. Sethuraman, ALEA Lat. Am. J. Prob. Stat. **4**, 245 (2008).
[24] H. Touchette, Phys. Rep. **478**, 1 (2009).
[25] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications* (Springer, 2009).
[26] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995).
[27] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).
[28] G. Jona-Lasinio, J. Stat. Mech. **P02004** (2014).
[29] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, Rev. Mod. Phys. **87**, 593 (2015).
[30] S. Sethuraman and S. R. S. Varadhan, Ann. Prob. **41**, 1461 (2013).
[31] See Supplementary Material.
[32] B. Derrida and A. Gerschenfeld, J. Stat. Phys. **136**, 1 (2009); **137**, 978 (2009).
[33] C. A. Tracy and H. Widom, Com. Math. Phys. **279**, 815 (2008).
[34] A. Borodin, I. Corwin, and T. Sasamoto, Ann. Prob. **42**, 2314 (2014).
[35] G. M. Schütz, J. Stat. Phys. **86**, 1265 (1997).
[36] T. Imamura and T. Sasamoto, J. Stat. Phys. **142**, 919 (2011).
[37] T. Halpin-Healy and Y.-C. Zhang, Phys. Rep. **254**, 215 (1995).
[38] M. Prähofer and H. Spohn, in *In and out of equilibrium, vol. 51 of Progress in Probability*, edited by V. Sidoravicius (2002) pp. 185–204.

[39] P. A. Ferrari, S. Goldstein, and J. L. Lebowitz, *Diffusion, mobility and the Einstein relation, J. Fritz (Ed.)* (Birkhaüser, Boston, 1985) pp. 405–441.
[40] P. L. Krapivsky, K. Mallick, and T. Sadhu, Phys. Rev. Lett. **113**, 078101 (2014).
[41] S. Sabhapandit, J. Stat. Mech. **L05002** (2007).
[42] C. Hegde, S. Sabhapandit, and A. Dhar, Phys. Rev. Lett. **113**, 120601 (2014).
[43] B. Derrida and T. Sadhu, J. Stat. Mech. **P09008** (2015).
[44] A. Borodin and I. Corwin, Prob. Th. Rel. Fields **158**, 225 (2014).
[45] A. Aggarwal, arXiv:1608.04726.
[46] T. Imamura, K. Mallick and T. Sasamoto, in preparation (2017).
[47] C. Giardina, J. Kurchan and L. Peliti, Phys. Rev. Lett. **96**, 120603 (2006).
[48] T. Sasamoto and H. Spohn, Phys. Rev. Lett. **834**, 523 (2010).
[49] T. Imamura and T. Sasamoto, J. Stat. Phys. **128**, 799 (2007).
[50] T. Imamura and T. Sasamoto, Phys. Rev. Lett. **108**, 190603 (2012).
[51] I. Corwin, Random Matrices: Theory and Applications **1** (2012).
[52] T. Sasamoto, Eur. Phys. Jour. B. **64**, 373 (2008).
[53] K. A. Takeuchi, M. Sano, T. Sasamoto, and H. Spohn, Sci. Rep. **1**, 11 (2011).
[54] P. L. Krapivsky, K. Mallick, and T. Sadhu, J Stat. Phys. **160**, 885 (2015).
[55] J. Ortmann, J. Quastel, and D. Remenik, Ann. App. Prob. **26**, 507 (2016).

Supplemental note for "Large deviations of a tracer in the symmetric exclusion process"

This supplemental note contains some extra-information and the derivation of some results given in the main text.

LARGE DEVIATION FUNCTION FOR NON-INTERACTING PARTICLES

When particles do not interact the motion of a tracer is a single-body problem. We recall the calculation of the large deviation function of a single particle on a one-dimensional lattice, hopping to the right and to the left with rates p and q, respectively. We use the same notations as in the main text. Discretizing time in infinitesimal steps dt, we write the position X_t of the particle at time t as

$$X_t = \epsilon_1 + \dots \epsilon_{t/dt}$$

where the increments ϵ_i are independent and can take three values, +1, -1 and 0 with probabilities pdt, qdt and $1-(p+q)dt$, respectively. Because of independence, we have

$$\langle e^{sX_t}\rangle = \langle e^{s\epsilon_i}\rangle^{t/dt} \sim e^{-tC(s)}$$

with

$$C(s) = q(1-e^{-s}) - p(e^s - 1).$$

The Large Deviation Function, defined as

$$\text{Prob}\left(\frac{X_t}{t} = -\xi\right) \simeq \exp[-t\phi(\xi)]$$

is obtained from $C(s)$ by Legendre Transform and is given by

$$\phi(\xi) = p + q - \sqrt{\xi^2 + 4pq} + \xi \ln \frac{\sqrt{\xi^2+4pq}+\xi}{2q}\,. \tag{S1}$$

Note that the rate for large deviations is proportional to the time t as expected in a non-interacting system (whereas for the SEP we have a $\sqrt{t}$ prefactor to the large deviations).

PROOF OF THE RELATION (9) BETWEEN X_t AND $N(x,t)$

We first define $Q(x,t)$ the time-integrated current that has flown through the bond $(x, x+1)$ between time 0 and t. The total current $Q(x,t)$ is equal to the total number of particles that have jumped from $x+1$ to x *minus* the total number of particles that have jumped from x to $x+1$ during the time interval $(0,t)$ (because we take the convention that a current flowing from right to left is counted positively).

By particle conservation, we find that the relation between $Q(x,t)$ and $N(x,t)$ is given by:

- for $x = 0$, $Q(0,t) = N(0,t) = N_t$.
- For $x > 0$, $Q(x,t) = Q(0,t) + \sum_{y=1}^{x}\left(\eta_y(t) - \eta_y(0)\right)$, or equivalently $N(x,t) = Q(x,t) + \sum_{y=1}^{x}\eta_y(0)$.
- For $x < 0$, $N(x,t) = Q(x,t) - \sum_{y=x+1}^{0}\eta_y(0)$.

Consider a site $x > 0$, located to the right of X_0, the initial position of the tracer. For the tracer X_t to be to the right of x, it is necessary that all the particles that were initially between X_0 and x have crossed the bond $(x, x+1)$ from left to right (including the tracer itself). This means that the total current $Q(x,t)$ has to be less than $-\sum_{i=X_0}^{x}\eta_i(0) = -\sum_{i=1}^{x}\eta_i(0)$ [here we use the fact the tracer is defined to be the particle which at $t = 0$ is the closest to the origin from the right; therefore all sites between 1 and $X_0 - 1$ are empty at $t = 0$]. We conclude that for $x \geq X_0$,

$$\begin{aligned}\text{Prob}\left(X_t > x\right) &= \text{Prob}\left(Q(x,t) \leq -\sum_{i=1}^{x}\eta_i(0)\right)\\ &= \text{Prob}\left(N(x,t) \leq 0\right).\end{aligned} \tag{S2}$$

A similar reasoning allows us to show that for $x < X_0$,

$$\text{Prob}\left(X_t \le x\right) = \text{Prob}\left(N(x,t) > 0\right). \tag{S3}$$

The two identities (S2-S3) imply the relation (6) of the main text.

RELATION BETWEEN $\phi(\xi)$ AND $\mu(\xi,\lambda)$

Here we show that $\phi(\xi)$, the large deviation function of tracer's position can be written in terms of $\mu(\xi,\lambda)$, the characteristic function of $N(x,t)$, as

$$\phi(\xi) = \max_{\lambda} \mu(\xi,\lambda). \tag{S4}$$

For this purpose we introduce $\Phi(\xi,q)$, the large deviation function of $N(x,t)$, given by

$$\text{Prob}\left(\frac{N(x,t)}{\sqrt{t}} = q\right) \simeq \exp[-\sqrt{t}\Phi(\xi,q)] \tag{S5}$$

with $\xi = -\frac{x}{\sqrt{4t}}$. From (S3) and (S5), we find

$$\phi(\xi) = \Phi(\xi, q=0). \tag{S6}$$

Note that the two functions Φ and μ are Legendre transforms of each other

$$\Phi(\xi,q) = \max_{\lambda}\left(\mu(\xi,\lambda) + \lambda q\right). \tag{S7}$$

Thus from (S6) and (S7), we get (S4).

PROOF OF THE FLUCTUATION RELATION

We start with the parametric representation of the large deviation function (eq. (5) in the main text). The function $\phi(\xi)$, for a given value of ξ is given by $\phi(\xi) = \mu(\xi,\lambda^*)$ where $\lambda^*(\xi)$ is such that $\frac{\partial\mu(\xi,\lambda^*)}{\partial\lambda} = 0$, which, using the formula (6) for $\mu(\xi,\lambda)$, is equivalent to

$$\left(r_+ e^{\lambda^*} - r_- e^{-\lambda^*}\right)\sum_{n=1}^{\infty}\frac{(-1)^{n-1}\omega(\lambda^*)^{n-1}}{n^{1/2}}A(\sqrt{n}\,\xi) = \xi\frac{\omega(\lambda^*)+\rho_+ + \rho_-}{\omega(\lambda^*)+1}. \tag{S8}$$

First, we note under the change $\lambda \to \log\frac{r_-}{r_+} - \lambda$,

$$\omega\Big(\log\frac{r_-}{r_+} - \lambda\Big) = \omega(\lambda), \quad r_+ e^{\log\frac{r_-}{r_+}-\lambda} - r_- e^{-\log\frac{r_-}{r_+}+\lambda} = -\left(r_+e^{\lambda} - r_-e^{-\lambda}\right). \tag{S9}$$

Using these properties and $A(\xi) = A(-\xi)$ to (S8), we deduce that

$$\lambda^*(-\xi) = \log\frac{r_-}{r_+} - \lambda^*(\xi). \tag{S10}$$

Therefore we have

$$\phi(\xi) - \phi(-\xi) = \mu(\xi,\lambda^*(\xi)) - \mu(-\xi,\lambda^*(-\xi)) = \mu(\xi,\lambda^*(\xi)) - \mu\Big(-\xi, \log\frac{r_-}{r_+} - \lambda^*(\xi)\Big).$$

Using again the formula (6) and the fact that ω is invariant, we find that this expression is equal to

$$\xi\log\frac{\left(1+\rho_+(e^{\lambda}-1)\right)\left(1+\rho_+(e^{\lambda'}-1)\right)}{\left(1+\rho_-(e^{-\lambda}-1)\right)\left(1+\rho_-(e^{-\lambda'}-1)\right)},$$

where we have written λ for $\lambda^*(\xi)$ and λ' for $\lambda^*(-\xi)$. After simplification this expression reduces to

$$2\xi\log\frac{1-\rho_+}{1-\rho_-}$$

thus proving the Fluctuation Theorem.

CALCULATION OF THE HIGHER CUMULANTS OF X_t

To extract the cumulants of the X_t from $\mu(\xi,\lambda)$, it is useful to use a parametric representation for the cumulant generating function $C(s)$:

$$\frac{\partial \mu(\xi,\lambda)}{\partial \lambda} = 0, \tag{S11}$$

$$\frac{\partial \mu(\xi,\lambda)}{\partial \xi} = 2s, \tag{S12}$$

$$C(s) = 2s\xi + \mu(\xi,\lambda). \tag{S13}$$

The first two equations define implicitly two functions $\xi(s)$ and $\lambda(s)$, that, after substitution in the third equation, provide us with the cumulant generating function $C(s)$.

Equilibrium case with uniform density

For a stationary initial condition, $\rho_+ = \rho_- = \rho$, the functions $\xi(s)$ and $\lambda(s)$ vanish when $s \to 0$. The strategy is to write power-series expansions w.r.t. s for these two functions

$$\xi(s) = \sum_{i\geq 1} x_i s^i,$$

$$\lambda(s) = \sum_{i\geq 1} l_i s^i.$$

In order to calculate the unknown coefficients x_i's and l_i's, we substitute these expansions in (S11) and in (S12). At each order n, we obtain two inhomogeneous linear equations for x_n and l_n, the r.h.s of which involve higher powers of x_i's and l_i's with $i < n$. This system can be solved systematically to any desired order. This can be done by hand up to $n = 4$ and for higher orders by using a symbolic computation software (such as Mathematica). Substituting the series truncated at order n for $\xi(s)$ and $\lambda(s)$ in $C(s)$ using (S13) gives the values of the first n cumulants. All odd cumulants vanish and the first few even cumulants are given by the following expression.
At order 2:

$$\frac{\langle X_t^2\rangle_c}{\sqrt{4t}} = \frac{1-\rho}{\rho\sqrt{\pi}}.$$

At order 4:

$$\frac{\langle X_t^4\rangle_c}{\sqrt{4t}} = \frac{1-\rho}{\sqrt{\pi}\rho^3}[1-(4-(8-3\sqrt{2})\rho)(1-\rho)+\frac{12}{\pi}(1-\rho)^2].$$

At order 6:

$$\begin{aligned}\frac{\langle X_t^6\rangle_c}{\sqrt{4t}} = \frac{1-\rho}{\pi^{5/2}\rho^5}\Big[&\Big(1020-450\pi+45\pi^2\Big)\\ &-\rho\Big(4800-\pi(2700-540\sqrt{2})+\pi^2(270-45\sqrt{2})\Big)\\ &+\rho^2\Big(6120-\pi(5250-1620\sqrt{2})+\pi^2(570-225\sqrt{2}+40\sqrt{3})\Big)\\ &-\rho^3\Big(4080-\pi(4200-1620\sqrt{2})+\pi^2(480-300\sqrt{2}+80\sqrt{3})\Big)\\ &+\rho^4\Big(1020-\pi(1200-540\sqrt{2})+\pi^2(136-120\sqrt{2}+40\sqrt{3})\Big)\Big].\end{aligned}$$

Non-equilibrium case with $\rho_+ \neq \rho_-$

When the system starts with an initial step-profile, it remains out-of-equilibrium and the tracer drifts away from the origin as shown by (12) in the main text. In order to calculate the cumulants we must again solve the system (S11), (S12) and (S13). Here, when $s \to 0$, we still have $\lambda \to 0$ but $\xi \to \xi_0$ where ξ_0 satisfies (13) in the main text. The procedure explained above for the equilibrium case can still be applied but if one just wants to calculate the variance of the tracer's position it is simpler to use the parametric representation of the large deviation function: $\phi(\xi) = \mu(\xi, \lambda^*)$ where $\lambda^*(\xi)$ is such that $\frac{\partial \mu(\xi,\lambda^*)}{\partial \lambda} = 0$ (eq. (8) in the main text).

The large deviation function is strictly positive for $\xi \neq \xi_0$ and vanishes at ξ_0, i.e., $\mu(\xi_0, \lambda^*(\xi_0)) = 0$. It is elementary to check from the formula (6) in the main text that $\lambda^*(\xi_0) = 0$.

To calculate the variance of X_t, we perform a second order expansion of $\phi(\xi_0 + \epsilon)$ with respect to ϵ. Writing $\lambda^*(\xi_0 + \epsilon) \simeq 1 + a\epsilon$, we determine a such that $\partial_\lambda \mu(\xi_0 + \epsilon, \lambda^*(\xi_0 + \epsilon)))$ vanishes at first order in ϵ. We find that $a = -U/V$ with

$$U = \rho_+ \mathrm{erfc}(\xi_0) + \rho_- \mathrm{erfc}(-\xi_0)$$

and

$$V = \frac{\rho_+^3 + \rho_-^3 - 3\rho_+^2\rho_- - 3\rho_+\rho_-^2 + 4\rho_+\rho_-}{(\rho_+ + \rho_-)} A(\xi_0) - (\rho_+ - \rho_-)^2 \frac{A(\sqrt{2}\,\xi_0)}{\sqrt{2}}.$$

Substituting this result in $\phi(\xi_0 + \epsilon) \simeq \mu(\xi_0 + \epsilon, 1 + a\epsilon)$ gives the correct dominant term in the large deviation function at order ϵ^2:

$$\phi(\xi) \simeq \frac{U^2}{2V} (\xi - \xi_0)^2 \,.$$

From this Gaussian limiting form, we deduce that

$$\mathrm{Var}(\xi) = \frac{V}{U^2 \sqrt{t}}$$

which, after rearranging the terms, leads to the formula for the variance of the tracer, given in the main text:

$$\mathrm{Var}(X_t) = \frac{4K(\rho_+ - \rho_-)^2 A(\xi_0)\sqrt{t}}{(\rho_+ \mathrm{erfc}(\xi_0) + \rho_- \mathrm{erfc}(-\xi_0))^2}$$

with

$$K = \frac{\rho_+^3 + \rho_-^3 - 3\rho_+^2\rho_- - 3\rho_+\rho_-^2 + 4\rho_+\rho_-}{(\rho_+ + \rho_-)(\rho_+ - \rho_-)^2} - \frac{A(\sqrt{2}\,\xi_0)}{\sqrt{2} A(\xi_0)} \,.$$