Self-Averaging and Weak Ergodicity Breaking of Diffusion in Heterogeneous Media

Anna Russian

Dipartimento di Ingegneria Civile e Ambientale,

Politecnico di Milano, Milano, Italy[*]

Marco Dentz

Spanish National Research Council (IDAEA-CSIC), 08034 Barcelona, Spain[†]

Philippe Gouze

Géosciences, Université de Montpellier, CNRS, Montpellier, France

(Dated: October 16, 2017)

Abstract

Diffusion in natural and engineered media is quantified in terms of stochastic models for the heterogeneity-induced fluctuations of particle motion. However, fundamental properties such as ergodicity and self-averaging and their dependence on the disorder distribution are often not known. Here, we investigate these questions for diffusion in quenched disordered media characterized by spatially varying retardation properties, which account for particle retention due to physical or chemical interactions with the medium. We link self-averaging and ergodicity to the disorder sampling efficiency R_n, which quantifies the number of disorder realizations a noise ensemble may sample in a single disorder realization. Diffusion for disorder scenarios characterized by a finite mean transition time is ergodic and self-averaging for any dimension. The strength of the sample to sample fluctuations decreases with increasing spatial dimension. For an infinite mean transition time, particle motion is weakly ergodicity breaking in any dimension because single particles cannot sample the heterogeneity spectrum in finite time. However, even though the noise ensemble is not representative of the single particle time statistics, subdiffusive motion in $q \geq 2$ dimensions is self-averaging, which means that the noise ensemble in a single realization samples a representative part of the heterogeneity spectrum.

---

[*] anna.russian@yahoo.it

[†] marco.dentz@csic.es

I. INTRODUCTION

The diffusion of particles in a fluid or gas at rest is described through Brownian motion, characterized by a linear increase of the mean square displacement and Gaussian displacement statistics. Fick's laws form the basis for the quantification of tracer motion and chemical interactions between diffusing species. These mechanisms, however, are very different for diffusion in heterogeneous media, in which particle motion may be hindered due to spatial variability in the medium properties [1, 2]. In fact, anomalous diffusion is a ubiquitous phenomenon [1, 3–5] observed in natural and engineered media, which are intrinsically heterogeneous. Diffusion in heterogeneous media refers to processes as diverse as radionuclide migration in low permeability geological media [6–10], surface and solid state diffusion [11–13], natural gas production from tight reservoirs [14–16], contaminant and heat transfer in the subsurface [17–21], the motion of charge carriers in amorphous semi-conductors [22], and the transmission of light in optical media [23], as well as the motion of endogenous and artificial tracers in living cells and cell membranes [24–29]

Stochastic modeling of particle motion in quenched random systems is particularly challenging because particles sample the disorder only by diffusion, whose efficiency in turn is affected by the medium properties. The fundamental assumption underlying the use of stochastic models for the prediction of particle motion is that the ensemble behavior may be representative for the behavior in a single realization. This may refer to the representation of the temporal statistics of a single particle trajectory through a noise ensemble, or to the representation of particle statistics in a single medium realization through the disorder ensemble statistics. The former relates to the notion of ergodicity [30], the latter to the notion of self-averaging [1].

Self-averaging can only apply to infinite systems which are characterized by statistically stationary disorder distributions so that diffusing particles have the possibility to sample a representative part of the disorder spectrum in a single realization. The lack of self-averaging for finite size quenched trap models has been recently discussed by Akimoto *et al.* [31]. Infinity of space and statistical stationarity, however, are only necessary conditions for self-averaging. It critically depends on the sampling efficiency and thus on the diffusion process, which in turn is affected by the heterogeneity [32]. Despite its fundamental nature, only few studies address the problem of self-averaging in q–dimensional quenched random

2

media, specifically the evolution of the variance of the mean square displacement, for diffusive or subdiffusive systems.

Motivated by experimental results on single particle tracking, the ergodicity of particle motion in fluctuating environments has received considerable attention [27, 30, 33, 34] with specific focus on subdiffusive motion [34–38]. Weak ergodicity breaking in disordered media is related to long waiting times between particle transitions [33], which impedes significant disorder sampling in finite time. The latter has put the continous time random walk (CTRW) [19, 30, 39] in the spotlight as a model for subdiffusive particle motion. The CTRW renews particle waiting times at each random walk step independently and thus can be seen as an annealed disorder model. Subdiffusion in the CTRW is weakly ergodicity breaking as a consequence of aging, this means non-stationarity of particle displacement caused by diverging particle waiting times [35, 40]. For the same reason, subdiffusion in quenched random trap models, which show CTRW behavior in $q > 2$ dimensions, has been found to be weakly ergodicity breaking [38, 41]. Bertin and Bouchaud [42] analyze weak ergodicity breaking and localization of subdiffusion in the $q = 1$ dimensional quenched random trap model. Massignan *et al.* [26] study ergodicity breaking of subdiffusion in a $q = 1$ dimensional inhomogeneous environment. Only few works address the ergodicity in $q \geq 1$–dimensional quenched disorder under diffusive and subdiffusive conditions despite its central importance for the understanding of diffusion in heterogeneous media.

In this paper we study self-averaging and ergodicity for diffusion in q–dimensional quenched random media. Diffusion is hindered due to spatially varying retardation properties, which account for physical and chemical interactions with the medium. This disorder model is equivalent to the quenched random trap model [1, 32]. The nature of average particle motion in the quenched random trap model depends on the dimensionality q of space [1], which is strongly related to the notion of the average number S_n of distinct sites visited as a function of step number n of the random walk. The latter determines the efficiency of the diffusion process for disorder sampling in a single medium realization. For $q > 2$, average particle motion describes a CTRW because particles do in average not return to the same site that has already been visited. For $q < 2$ this is different. Particle motion describes a CTRW-like behavior, which however is characterized by dependence of subsequent waiting times because particles may return to the same site with finite probability. Clearly this property has an impact on the sampling efficiency and thus self-averaging property.

This paper extends previous work [32, 38], which focused on the questions of self-averaging and weak ergodicity breaking for disorder scenarios leading to subdiffusive average behaviors. Here, we derive a general framework for the analysis and quantification of self-averaging and ergodicity for arbitrary disorder distributions, characterized by diffusive and subdiffusive average behaviors. We study particle motion in single disorder realizations to illustrate the impact of medium heterogeneity on the segregation of particle distributions and their sample to sample fluctuations. Ergodicity and self-averaging are quantified systematically in terms of the noise variance of the time-averaged mean square displacement in single medium realizations (ergodicity), and the disorder variance of the noise mean squared displacement (self-averaging). We derive analytical expressions for the disorder variance of the mean square displacement and the noise variance of the time-averaged mean square displacement under both diffusive and subdiffusive conditions. Note that the concepts of ergodicity and self-averaging are different. Ergodicity refers to the equivalence between temporal average and noise average, this means average over all particles in a given disorder realization. Self-averaging on the other hand denotes the property that the mean square displacement, defined as a noise average in a single realization, may be represented by its disorder average. Under self-averaging conditions, the mean square displacement is independent of the specific disorder realization. Thus, it may be weakly ergodicity breaking, but self-averaging.

The paper is organized as follows. Section II describes the diffusion problem, presents the coarse-graining procedure, and derives explicit expressions for the noise average, a central part of the paper. Furthermore, it defines the disorder average in terms of the coarse grained particle trajectories. Section III discusses the disorder scenarios under consideration and illustrates the behavior of the mean square displacement and particle distributions in the average over all disorder realizations as well as in single disorder realizations. Section IV analyzes self-averaging for different heterogeneity scenarios and derives explicit expressions for the variance of the mean square displacement. Section V discusses the ergodicity of single particle trajectories in terms of the noise variance of the time-averaged mean square displacement.

II. DIFFUSION IN HETEROGENEOUS MEDIA

Particle transport in a q–dimensional quenched random medium, characterized by a spatially random diffusion coefficient $D(\mathbf{x})$, can be described by the Langevin equation

$$\frac{d\mathbf{x}(t)}{dt} = \sqrt{2D[\mathbf{x}(t)]}\boldsymbol{\xi}(t), \tag{1}$$

which is characterized by a multiplicative noise. We employ here and in the following the Ito interpretation. The Gaussian white noise has zero mean and covariance $\langle \xi_i(t)\xi_j(t')\rangle = \delta_{ij}\delta(t - t')$. The angular brackets denotes the noise average over all particles. The particle distribution $p(\mathbf{x}, t) = \langle \delta[\mathbf{x} - \mathbf{x}(t)]\rangle$ is then transported according to the Fokker-Planck equation [43]

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} - \nabla^2 D(\mathbf{x})p(\mathbf{x}, t) = 0. \tag{2}$$

The diffusion coefficient in (2) is given by $D(\mathbf{x}) = \kappa\theta(\mathbf{x})^{-1}$, where κ is the constant molecular diffusion coefficient. The quenched retardation coefficient $\theta(\mathbf{x})$ accounts for particle retention due to physical and chemical interactions with the medium and thus represents the medium heterogeneity. Equation (2) describes diffusion under spatially variable retardation properties. The retardation coefficient is modeled here as a stationary spatial random field which is characterized by the characteristic correlation scale ℓ and the distribution $p_\theta(\theta)$ of point values $\theta(\mathbf{x})$. Note that we consider here two stochastic processes. The temporal stochastic process $\boldsymbol{\xi}(t)$, which defines the noise ensemble, and the spatial stochastic process $\theta(\mathbf{x})$ which defines the disorder ensemble. We consider in the following deterministic initial conditions of instantaneous particle injection at $t = 0$ at $\mathbf{x} = \mathbf{0}$. As discussed in the Introduction and also below, particles sample the medium heterogeneity, which is represented by the disorder ensemble, due to diffusion, which is represented by the noise ensemble. Note that under certain experimental conditions, for example, the inital particle position in a single disorder relaization may be uncertain and an additional average over the initial particle locations may be required. In such a case, the noise ensemble samples part of the heterogeneity from the beginning. On the other hand, there may be situations, for example leachage of a contaminant in an underground repository, for which the initial positions are well confined and thus may be denominated determinstic.

The mean square displacement in a single medium realization is defined by $m(t) = \langle \mathbf{x}(t)^2\rangle$.

From (1) we obtain

$$m(t) = 2 \int\limits_0^t dt' \int\limits_0^t dt'' \left\langle \sqrt{D[\mathbf{x}(t')]D[\mathbf{x}(t'')]}\boldsymbol{\xi}(t') \cdot \boldsymbol{\xi}(t'') \right\rangle$$

Using the Ito interpretation of the stochastic integral we obtain the compact expression

$$m(t) = 2q \int\limits_0^t dt' \langle D[\mathbf{x}(t')]\rangle \tag{3}$$

because $\langle D[\mathbf{x}(t)]\boldsymbol{\xi}(t)\rangle = 0$. We set now $D(\mathbf{x}) = \kappa\theta(\mathbf{x})^{-1}$ and define

$$ds = dt\,\theta[\mathbf{x}(t)]^{-1}. \tag{4}$$

Thus, we can write $m(t)$ in the form

$$m(t) = 2\kappa q\langle s(t)\rangle. \tag{5}$$

In the following, we use the variable transform $t \to s$ as the basis for the coarse-graining of the Langevin equation (1).

A.   Coarse Graining

The Langevin equation (1) can be written in the Ito interpretation as

$$d\mathbf{x}(t) = \sqrt{2\kappa\theta[x(t)]^{-1}dt}\boldsymbol{\eta}(t), \tag{6}$$

where

$$\boldsymbol{\eta}(t) = \frac{1}{\sqrt{dt}} \int\limits_t^{t+dt} dt'\boldsymbol{\xi}(t') \tag{7}$$

is a Gaussian random vector with zero mean and unit variance such that $\langle \eta_i(t)\eta_j(t)\rangle = \delta_{ij}$. Using the transformation (4), we can separate (6) into the two equations

$$d\mathbf{x}(s) = \sqrt{2\kappa ds}\boldsymbol{\eta}(s), \qquad\qquad dt(s) = \theta[\mathbf{x}(s)]ds. \tag{8}$$

This means we have separated the non-linearity from the spatial particle movement and transferred it to the evolution of the particle time with s. The process $s(t) = \max\{s|t(s) \le t\}$ denotes the operational time that passes to reach the clock time t.

It is natural to choose the correlation scale ℓ of $\theta(\mathbf{x})$ as the spatial coarsening scale. This gives for the coarse grained equations of motion

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \ell\,\boldsymbol{\eta}_{n+1}, \qquad\qquad t_{n+1} = t_n + \theta(\mathbf{x}_n)\Delta s_n. \qquad (9)$$

The random vector $\boldsymbol{\eta}_n$ has unit length, is always aligned with one of the axes of the coordinate system, and its ensemble mean is zero. Thus, (9) describes a random walk on a q–dimensional lattice. The operational time increment Δs_n is the first passage time to the boundaries of a region of size ℓ by regular diffusion. The distribution of the first passage time Δs_n is cut-off exponentially for times larger than the the characteristic diffusion time $\tau_\kappa = \ell^2/(2q\kappa)$ over the length scale ℓ. Thus, we approximate the first passage time PDF by the exponential distribution [44]

$$\psi_0(\Delta s_n) = \tau_\kappa^{-1}\exp(-\Delta s_n/\tau_\kappa). \qquad (10)$$

The transition time over the length ℓ is given by

$$\tau(\mathbf{x}_n) = \theta(\mathbf{x}_n)\Delta s_n, \qquad (11)$$

which, through its dependence on $\theta(\mathbf{x}_n)$ is a quenched random variable. Equation (9) describes a time-domain random walk (TDRW) [9, 45, 46], which can be seen as a spatially inhomogeneous CTRW[44, 47] because the transition time depends on the particle position. The choice of the coarse-graining scale as the correlation scale of the heterogeneous retardation field $\theta(\mathbf{x})$ implies that all $\tau(\mathbf{x}_n)$ and $\tau(\mathbf{x}_{n'})$ are independent for $n \neq n'$. This means that the distribution of $\tau(\mathbf{x}_n)$ can be characterized by the one-point probability density function (PDF) $\psi(\tau)$ of transition times, which is given by

$$\psi(\tau) = \int\limits_0^\infty d\theta\,\theta^{-1}p_\theta(\theta)\psi_0(\tau/\theta). \qquad (12)$$

It is fully determined in terms of the distribution $p_\theta(\theta)$ of point values of $\theta(\mathbf{x})$. This is an important fact because it allows relating the statistics of the medium heterogeneity to average particle motion as outlined in Section III. With these definitions, (9) is also equivalent to the quenched random trap model [1].

The space-time trajectory of a particle diffusing in the heterogeneous medium are obtained from (9):

$$\mathbf{x}_n = \sum_{k=0}^{n-1}\ell\boldsymbol{\eta}_k, \qquad\qquad t_n = \sum_{k=0}^{n-1}\tau(\mathbf{x}_k). \qquad (13)$$

It is worth noticing that here the sequence of $\tau(\mathbf{x}_k)$ is in general not composed of independent time increments because particles may visit the same site more than once [1] depending on the dimensionality of space. This differentiates the TDRW (9) from a classical CTRW characterized by completely independent random time increments. This point is discussed in more detail in Section II C below. The particle position at time t is given by $\mathbf{x}(t) = \mathbf{x}_{n_t}$, where $n_t = \max(n|t_n \leq t)$ denotes the number of steps needed to reach time t by the time random walk (9). Likewise, the operational time $s(t)$ to arrive at time t is given by $s(t) = s_{n_t}$. The particle density in this coarse-grained framework is given by $p(\mathbf{x}, t) = \langle \delta(\mathbf{x} - \mathbf{x}_{n_t}) \rangle$. We can expand $p(\mathbf{x}, t)$ as

$$p(\mathbf{x}, t) = \sum_{n=0}^{\infty} \langle \delta(\mathbf{x} - \mathbf{x}_n) \mathbb{I}(t_n \leq t < t_{n+1}) \rangle, \tag{14}$$

where the indicator function $\mathbb{I}(t_n \leq t < t_{n+1})$ is 1 if the statement in the parenthesis is true and 0 else. Furthermore, the mean square displacement (5) is now given in terms of the coarse grained trajectory (9) as

$$m(t) = \ell^2 \langle n_t \rangle. \tag{15}$$

B. Noise Average

The average diffusion behavior and its self-averaging properties depend on the efficiency of the random walker to sample the disorder, which in turn depends on the sampling efficiency due to the random noise in single medium realizations. It is intuitive that a random walker explores a larger number of new sites as a function of random walk steps in $q = 3$ than in $q = 1$ spatial dimensions. In the following, we quantify this notion. First, we consider the coarse-grained trajectory (13) in a single disorder realization.

As outlined in the previous section, subsequent transition times $\tau(\mathbf{x}_k)$ are in general not independent because particles may return to the same position after a certain number of steps. The recurrence of the random walk may be described by the average number of independent sites visited S_n [1, 48]. The number S_n depends on the dimensionality of space in accord with Polya's theorem. For $q < 2$ spatial dimensions, $S_n \sim n^{q/2}$, for $q = 2$, $S_n = n/\ln(n)$ and for $q > 2$, $S_n \sim n$. Thus, as outlined in Bouchaud and Georges [1], the $\tau(\mathbf{x}_k)$ in (13) may be grouped into families of independent τ_i with n/S_n members each. If
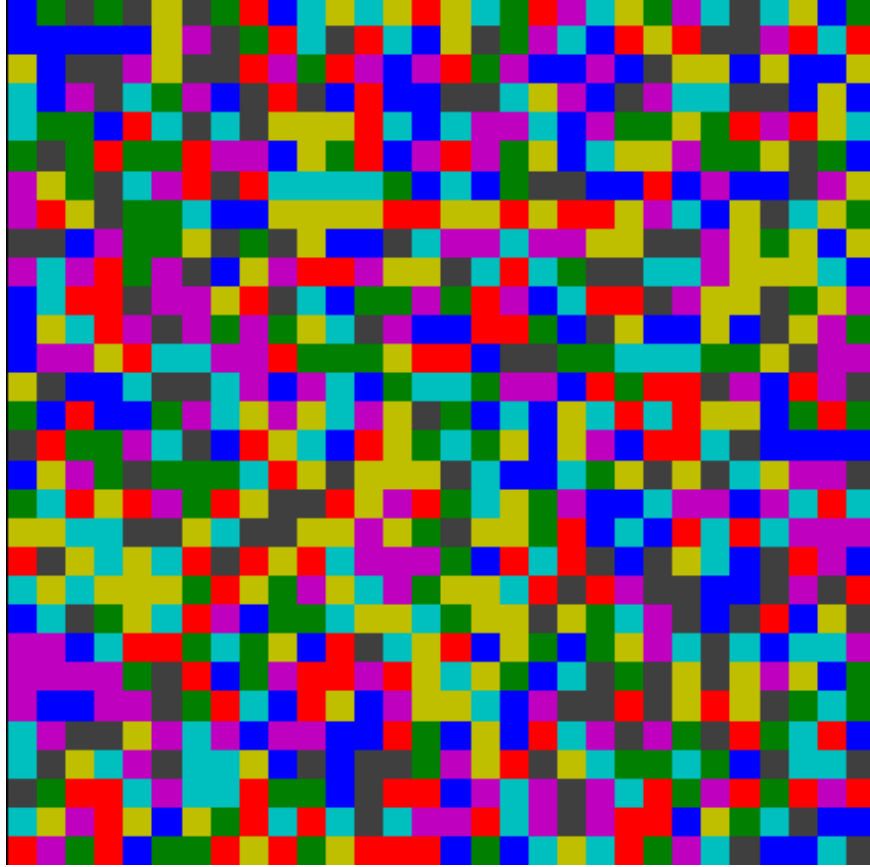
FIG. 1. Realization of a quenched random medium organized in equally sized pixels. The retardation coefficient $\theta(\mathbf{x})$ is distributed according to a heavy-tailed PDF for $\beta = 1/4$. Different colors denote different values of $\theta(\mathbf{x})$ on a logarithmic scale.

all members of a family are approximated by the same value τ_i, the trajectories for single disorder realizations can be re-written as

$$\mathbf{x}_n = \sum_{i=1}^{n} \ell \boldsymbol{\eta}_i, \qquad\qquad t_n = \sum_{i=1}^{S_n} \frac{n}{S_n} \tau_i, \qquad (16)$$

where the τ_i are distributed according to (12). In a single disorder realization, the set of available τ_i to form a replica of t_n depends on the volume sampled by the random walkers after n steps. This property and the independence of subsequent transition times are the basis for the formalisation of the noise average as described below, and of the disorder ensemble average as outlined in the next section.

In order to determine the noise average, we consider expression (15) for the mean square

9

displacement. It can be expanded according to

$$m(t) = \ell^2 \sum_{n=0}^{\infty} n \langle \mathbb{I}(t_n \leq t < t_{n+1}) \rangle, \tag{17}$$

where we used (14). The noise average can be written as the average over all particles released in a disorder realization such that

$$m(t) = \ell^2 \sum_{n=0}^{\infty} n \left\{ \lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} \mathbb{I} \left[t_n^{(j)} \leq t < t_{n+1}^{(j)} \right] \right\}, \tag{18}$$

where $t_n^{(j)}$ denotes the time of the jth particle after n steps, which is defined by (16). In order to simplify the sum over the particles j, we first note that the $t_n^{(j)}$ are given by the sum of waiting times τ_i, which in turn are determined by the disorder explored by the particles. Furthermore, we note that the average number of different sites sampled by the ensemble of random walkers after n steps is given by $V_n = (2n)^{q/2}$, which corresponds to the characteristic volume explored by the random walkers after n steps. This means that the $t_n^{(j)}$ are series of τ_i which are sampled from the same set of V_n statistically independent copies. Next we recall from (16) that a replica of t_n is formed of S_n independent τ_i, which are drawn from a set of V_n independent copies. Thus, the number of copies of t_n which can be formed from V_n members τ_i is equal to the binomial coefficient $B_n = \binom{V_n}{S_n}$. The number of independent replica of t_n that can be formed from V_n members on the other hand, is equal to $R_n = V_n/S_n$. Consequently, the number of families with members, who are statistically independent is given by B_n/R_n. Then, since each of the B_n/R_n families explore in average the same disorder, they are in average the same. This means that the number of independent replica of t_n in a sample of N is R_n. Thus, the sum in the curly brackets of (18) can be substituted after regrouping as

$$\sum_{j=1}^{N} \mathbb{I} \left[t_n^{(j)} \leq t < t_{n+1}^{(j)} \right] = \sum_{k=1}^{R_n} \frac{N}{R_n} \mathbb{I} \left[t_n^{(k)} \leq t < t_{n+1}^{(k)} \right]. \tag{19}$$

Expression (18) for $m(t)$ then simplifies to [41]

$$m(t) = \ell^2 \sum_{n=0}^{\infty} \frac{n}{R_n} \sum_{k=1}^{R_n} \mathbb{I} \left[t_n^{(k)} \leq t < t_{n+1}^{(k)} \right], \tag{20}$$

where the $t_n^{(k)}$ are independent replica of t_n. Each replica is formed according to (16), where the τ_i are independent identically distributed random variables characterized by the
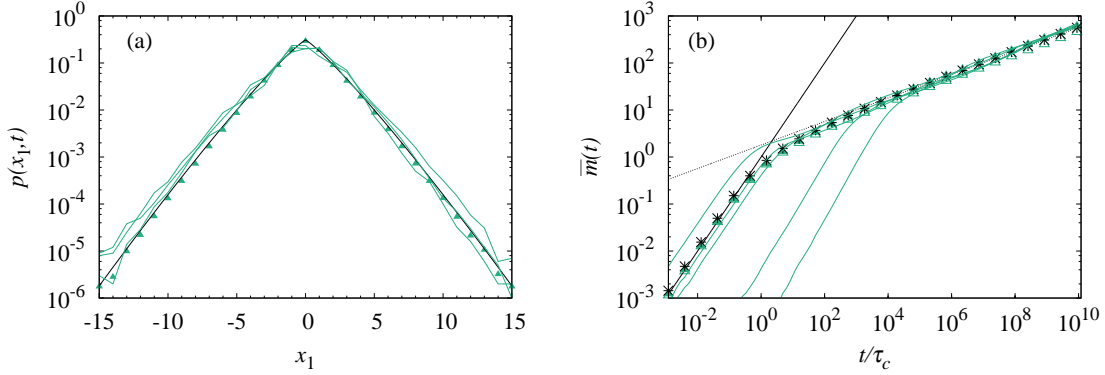
10

FIG. 2. (a) Concentration profile at $t = 20\tau_\kappa$ and (b) disorder average mean square displacement for a pulse injection into $q = 3$ dimensional random media for $\beta = 1/4$ and $\kappa = 1$. The full symbols represent the disorder ensemble average, lines the behaviors in single realizations. (Right) Open symbols denote the behavior of the corresponding annealed disorder model, stars the behavior of the equivalent CTRW discussed in Section II C. The solid black line denotes the early time behavior (31), the dashed line denotes the late time behavior (33).

PDF (12). Expression (20) captures the impact of noise on the disorder sampling in a single medium realization. In general, we obtain for the noise average of the kth power n_t^k of the number of steps needed to arrive at time t

$$\langle n_t^k \rangle \approx \sum_{n=0}^{\infty} \frac{n^k}{R_n} \sum_{k=1}^{R_n} \mathbb{I}\left[t_n^{(k)} \leq t < t_{n+1}^{(k)}\right]. \tag{21}$$

Similarly, the particle density (14) in a single realization may be approximated in terms of independent replica of t_n as

$$p(\mathbf{x}, t) = \sum_{n=0}^{\infty} \langle \delta(\mathbf{x} - \mathbf{x}_n) \rangle \frac{1}{R_n} \sum_{i=1}^{R_n} \mathbb{I}(t_n^{(i)} \leq t < t_{n+1}^{(i)}). \tag{22}$$

These expressions form the basis for the quantification of the disorder average mean squared displacements, its self-averaging properties and ergodicity in the following sections.

The key quantity for the determination of the impact of diffusion on disorder sampling in single disorder realizations and thus the average diffusion behavior is the number of independent replica of t_n that can be formed in a single realization,

$$R_n = \frac{(2n)^{q/2}}{S_n}. \tag{23}$$

11

The sampling efficiency R_n may be seen as the number of disorder realizations which can be sampled by the noise ensemble in a single disorder realization. For $q < 2$ spatial dimension $R_n \sim 1$ which implies that all particles in a single disorder realization see in average the same disorder. There is no diversification and disorder sampling is minimum. Differences between disorder realizations decrease as the ensemble of particles in single realizations experiences a disorder spectrum that is representative for the ensemble of media. For $q = 2$ we find that $R_n \sim \ln(n)$. This means that the number of disorder configuration that may be sampled within a single realization increases logarithmically with the number of steps; there is a weak sampling. For $q > 2$, we find that $R_n \sim n^{q/2-1}$. As expected, the efficiency of disorder sampling increases with the dimensionality of space. These behaviors have an impact on the self-averaging properties of the mean square displacement as discussed in the following.

C. Disorder Average

The average diffusion behavior can be described by the disorder average over the particle distribution (22),

$$\overline{p}(\mathbf{x}, t) = \sum_{n=0}^{\infty} \langle \delta(\mathbf{x} - \mathbf{x}_n) \rangle \frac{1}{R_n} \sum_{i=1}^{R_n} \overline{\mathbb{I}(t_n^{(i)} \leq t < t_{n+1}^{(i)})}. \tag{24}$$

Here and in the following the overbar indicates the average over the ensemble of disorder realizations. The disorder ensemble average removes here the dependence of the $t_n^{(i)}$ on the single disorder realizations such that we can write

$$\overline{p}(\mathbf{x}, t) = \sum_{n=0}^{\infty} \langle \delta(\mathbf{x} - \mathbf{x}_n) \rangle p_n(t), \tag{25}$$

where $p_n(t)$ is the distribution of numbers of steps needed to reach time t,

$$p_n(t) = \overline{\mathbb{I}(t_n \leq t < t_{n+1})}. \tag{26}$$

The particle time t_n after n steps is given by (16).

In $q > 2$ spatial dimensions the number S_n of distinct sites visited increases linearly with step number, $S_n = bn$, where b depends on the type of lattice on which the medium is discretized, and thus on the geometry of the medium. Here, we consider a simple cubic lattice, for which $b \approx 0.66$ [48]. Accordingly, for $q > 2$, the renormalised particle trajectory (16) can
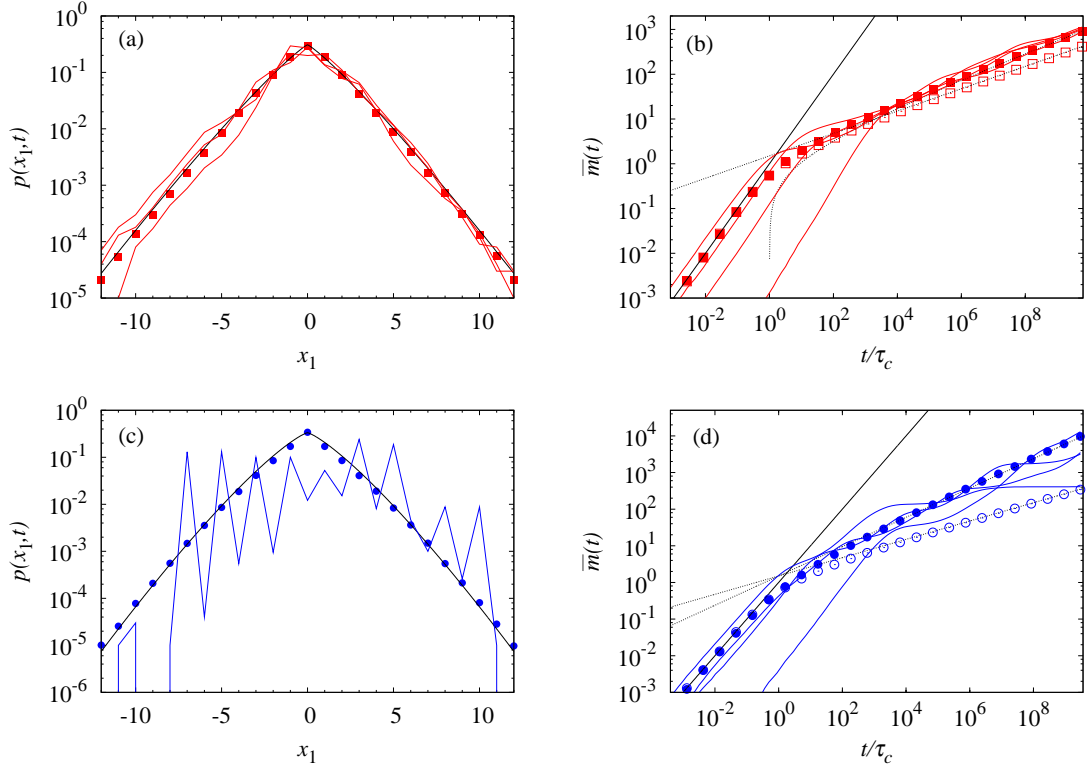
12

FIG. 3. (a, c) Concentration profiles at $t = 20\tau_\kappa$ for a pulse injection into (a) $q = 2$ and (c) $q = 1$ dimensional random media for $\beta = 1/4$ and $\kappa = 1$. The full symbols represent the disorder ensemble average, lines the behaviors in single realizations. (b, d) Disorder average mean square displacement for a pulse injection into (b) $q = 2$ and (d) $q = 1$ (b) dimensional random media for $\beta = 1/4$ and $\kappa = 1$. The full symbols represent the disorder average, lines the behaviors in single realizations. Open symbols denote the behavior of the corresponding annealed disorder model. The solid black line denotes the early time behavior (31), the dashed lines denotes the late time behaviors (35).

then be written as

$$\mathbf{x}_N = \sum_{i=1}^{N} \frac{\ell}{\sqrt{b}} \boldsymbol{\eta}_i, \qquad\qquad t_N = \sum_{i=1}^{N} \frac{\tau_i}{b}, \qquad (27)$$

see Appendix A. These implies that the average particle motion describes a CTRW [1]. For $q \leq 2$ it is not possible to formulate the renormalised space-time particle motion (16) in the form of a CTRW because of the step number dependence of the ratio S_n/n, which is $S_n/n \propto n^{q/2-1}$ in $q < 2$ and $S_n/n = 1/\ln(n)$ in $q = 1$ spatial dimension.

13

III. DIFFUSION BEHAVIOR

In this Section we investigate the diffusion behavior in $q = 1, 2$ and 3 spatial dimensions both numerically and analytically. We study the disorder ensemble average diffusion behavior as well as diffusion in single disorder realizations. The media are characterized by a spatially random distribution of retardation coefficients $\theta(\mathbf{x})$ and organized in equisized hypercubes of length ℓ as illustrated in Figure 1. The retardation coefficients $\theta(\mathbf{x})$ are assigned to the hypercubes randomly. This type of medium has a linear tent-shaped correlation function. As discussed in Section II A, such media can be considered caricatures of more general random media characterized by a constant correlation distance ℓ. The spatial random fields $\theta(\mathbf{x})$ generated in this way are stationary and isotropic. The distribution of point values of $\theta(\mathbf{x})$ is denoted by $p_\theta(\theta)$. We consider in the following disorder scenarios characterized by

$$p_\theta(\theta) \propto \theta^{-1-\beta} \tag{28}$$

with $\beta > 0$ and $\theta \geq 1$. According to (12), this implies a transition time PDF that behaves as

$$\psi(t) \propto \left(\frac{t}{\tau_\kappa} \right)^{-1-\beta}. \tag{29}$$

Moments of order higher than $\lfloor \beta \rfloor$ are not defined, where $\lfloor \cdot \rfloor$ denotes the floor function. The setup of the numerical random walk simulations and the specific choice for $p_\theta(\theta)$ are described in Appendix F. In the following, we first discuss the disorder average behavior, which we then contrast with the diffusion behavior in single medium realizations.

We study the diffusion behavior in terms of the particle distributions and mean square displacements. For illustration, we consider the integrated particle distributions $p(x_1, t) = \langle \delta[x_1 - x_1(t)] \rangle$ for $q = 2$ and 3As initial condition we consider an instantaneous particle injection at $\mathbf{x} = \mathbf{0}$. We first consider the disorder average diffusion behavior and briefly summarize the well known long time scaling of the disorder average mean square displacement $\overline{m}(t)$. We then study the diffusion behavior in single realizations and the sample to sample fluctuations as a function of the spatial dimensions and the disorder distribution.

A. Disorder average behavior

The disorder averaged mean square displacement $\overline{m}(t) = \overline{\langle \mathbf{x}(t)^2 \rangle}$ is obtained by averaging (20) as

$$\overline{m}(t) = \ell^2 \overline{n_t}, \tag{30}$$

where the overbar indicates the average over the ensemble of realizations. During the first step, n_t can be approximated by a Poisson variable whose mean is given by $\overline{n}_t = t/\tau_\kappa \overline{\theta(\mathbf{0})^{-1}}$. Thus, the early time behavior of $\overline{m}(t)$ is

$$\overline{m}(t) = \frac{2q\kappa}{\theta_H} t, \tag{31}$$

where $\theta_H = 1/\overline{\theta^{-1}}$ is the harmonic average of θ. This behavior is valid for times shorter than the characteristic diffusion time τ_c across the injection region of size ℓ,

$$\tau_c = \theta_H \tau_\kappa, \tag{32}$$

In the following, we briefly discuss the diffusion behavior for disorder scenarios characterized by infinite and finite mean transition times.

1. Infinite mean transition time ($0 < \beta < 1$)

For infinite mean transition time the disorder average long time diffusion behavior is subdiffusive. The scalings of the mean squared displacements for diffusion in quenched disorder depend on the dimensionality of space and are reported in Ref. [1] and derived in the current framework in Appendix D 1.

As discussed in Section II C, for $q > 2$ dimensions, for $q > 2$ dimensions, the disorder average diffusion behavior describes a CTRW and $\overline{m}(t)$ scales as

$$\overline{m}(t) \propto t^\beta. \tag{33}$$

Figure 2a shows the areally integrated profile, which behaves as the stretched exponential [39]

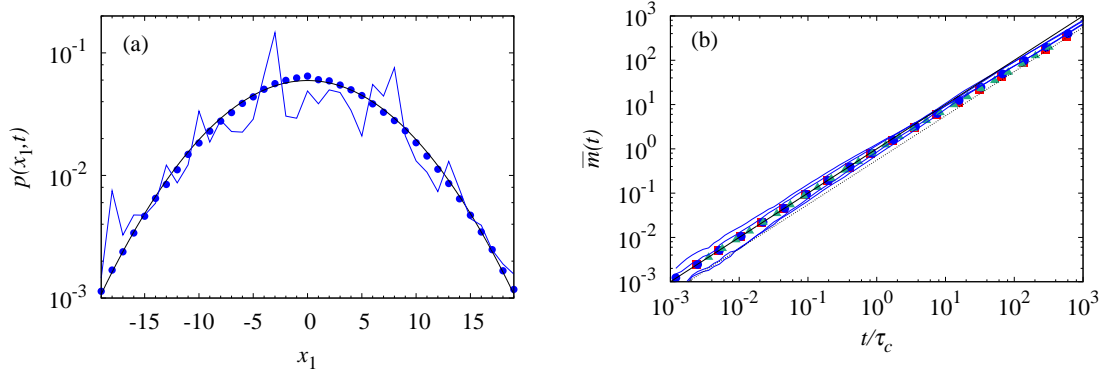$$\overline{p}(x_1, t) \sim \exp\left[-a|x_1/\overline{m}(t)|^{2/(2-\beta)}\right] \tag{34}$$

FIG. 4. (a) Concentration profile at $t = 20\tau_c$ for a pulse injection into a $q = 1$ dimensional random media for $\beta = 3/2$ and $\kappa = 1$. The full symbols represent the disorder average, the line the behavior in a single medium realization. (b) Disorder average mean square displacement for a point injection into (triangles) $q = 3$, (squares) $q = 2$ and (circles) $q = 1$ dimensional random media. The full symbols represent the ensemble average, the lines the behavior in single $q = 1$ dimensional medium realizations. The solid black line denotes the early time behavior (31), the dashed line the late time behavior (37).

with a a positive constant and $\overline{m}(t)$ given by (33). Figure 2b shows the numerically obtained $\overline{m}(t)$, the exact early time behavior (31) and the long time scaling (33). The behavior of the equivalent renormalised CTRW discussed in Section II C and the diffusion behavior in an annealed disorder scenario characterized by $\psi(t)$ are also displayed in Figure 2b. Note that the annealed scenario does not account for the renormalisation of the space and time increments. The mean square displacement in the annealed scenario is equal to the quenched case at short times. It underestimates $\overline{m}(t)$ at times $t > \tau_c$ because it does not account for the renormalised time and space increments. Conversely, the equivalent CTRW coincides with the quenched behavior for $t \gg \tau_c$. The disorder average behavior behaves as in an equivalent annealed disorder. The quenched nature of the spatial disorder is only revealed in the sample to sample fluctuations as discussed in the next section.

In $q = 2$ and $q < 2$ dimensions, $\overline{m}(t)$ behaves as

$$\overline{m}(t) \propto t^{\beta} \ln(t)^{1-\beta}, \qquad\qquad \overline{m}(t) \propto t^{\frac{2\beta}{2\beta-q\beta+q}}, \qquad (35)$$

respectively. The mean square behavior is sub-diffusive, but increases faster than in $q > 2$. In the following, we focus on $q = 1$ and $q = 2$ as well as $q > 2$ spatial dimensions.

16

Figures 3a and c show the disorder average particle distributions in $q = 2$ and $q = 1$ dimensions and illustrates that $\overline{m}(t)$ increases faster as dimension decreases. The profile for $q = 2$ dimensions adjusts to the same stretched exponential (34) as in $q = 3$ with $\overline{m}(t)$ given by (35). For $q = 1$ the profile is well described by the stretched exponential [42]

$$p(x_1, t) \sim \exp(-b|x/\overline{m}(t)|^{1+\beta}) \tag{36}$$

with b a positive constant and $\overline{m}(t)$ the disorder average mean square displacement (35) for $q = 1$.

Figures 3b and d display the disorder average mean square displacement for $q = 2$ and $q = 1$ dimensions. For comparison, we plot the diffusion behavior for an annealed disorder scenario characterized by the same $\psi(t)$. The short time behavior is given exactly by (31); the long time behavior for $t \gg \tau_c$ is given by (35). At short times $t < \tau_c$ the annealed and quenched scenarios behave in the same way because the behavior is given by the harmonic average of the residence times in the injection region. At larger time, when particles start sampling the disorder, the space dependency of transition times implies correlation of successive time increments. Thus, in quenched disorder particles visit in average a lower number of different sites than in annealed disorder. Consequently particles encounter a lower number of high retardation sites that determines long time behaviour. As a consequence, mean square displacement increases slower in annealed than in quenched disorder.

2. *Finite mean transition time ($\beta > 1$)*

Here, we consider media characterized by a power law distribution of retardation coefficients with $\beta > 1$. In this case, the mean transition time $\overline{\tau}$ exists. The long time diffusion behavior is normal. The disorder average mean square displacement behaves in leading order as (see Appendix D 2)

$$\overline{m}(t) = \frac{2q\kappa}{\theta_A}t, \tag{37}$$

where $\theta_A = \overline{\theta(\mathbf{x})}$ is the arithmetic average of $\theta(\mathbf{x})$. This behavior holds in any dimension. The quenched nature of the disorder manifests in the self-averaging behavior of the mean square displacement as discussed below.

Figure 4a illustrates the particle distribution for $q = 1$ and $\beta = 3/2$. It is well described by a Gaussian distribution. Figure 4b shows the disorder average mean square displacement

17

obtained for $q = 1, 2$ and $q = 3$ dimensional disorder distribution with $\beta = 3/2$ and the particle distribution for $q = 1$. We observe that $\overline{m}(t)$ scale linearly with time for any dimension. For $t < \tau_c$ it behaves as predicted by (31) and for $t > \tau_c$ as (37).

B. Single medium realizations

In this Section we study diffusion in single disorder realizations. We discuss the sample to sample fluctuations of the mean square displacement and the particle distribution between realizations. The early time mean square displacement in a single realization is given by

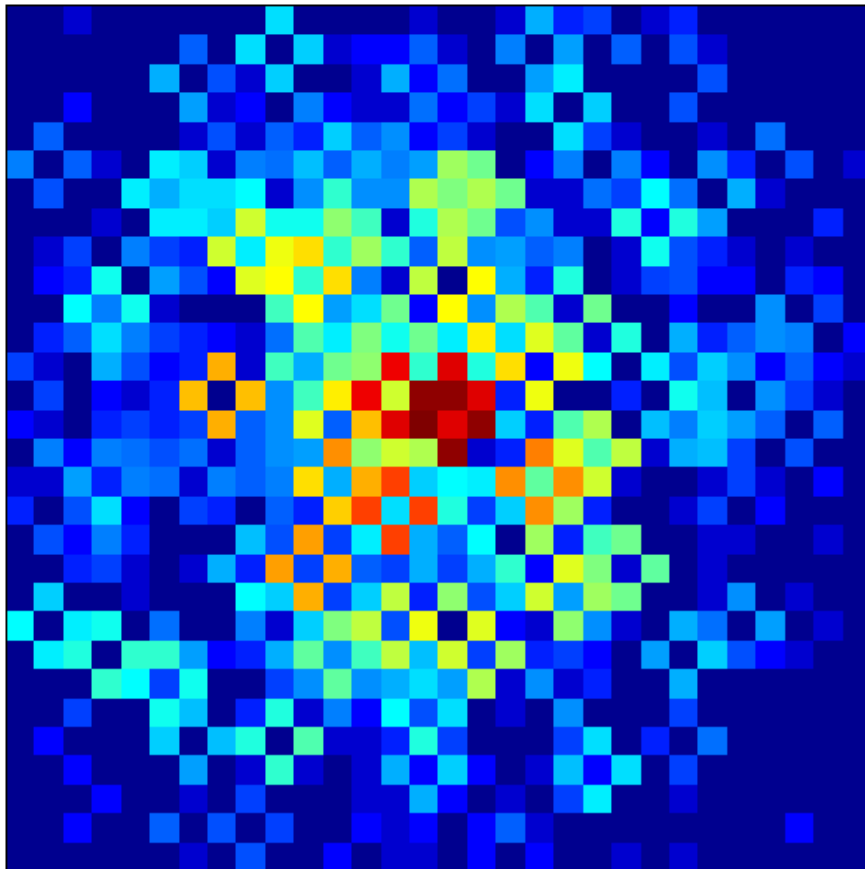$$m(t) = \frac{2q\kappa}{\theta(\mathbf{0})}t. \tag{38}$$

FIG. 5. Particle distribution for a pulse injection into a $q = 2$ dimensional heterogeneous medium for $\beta = 1/4$ and $\kappa = 1$. Different colors denote different values of concentration with a logarithmic color scheme.
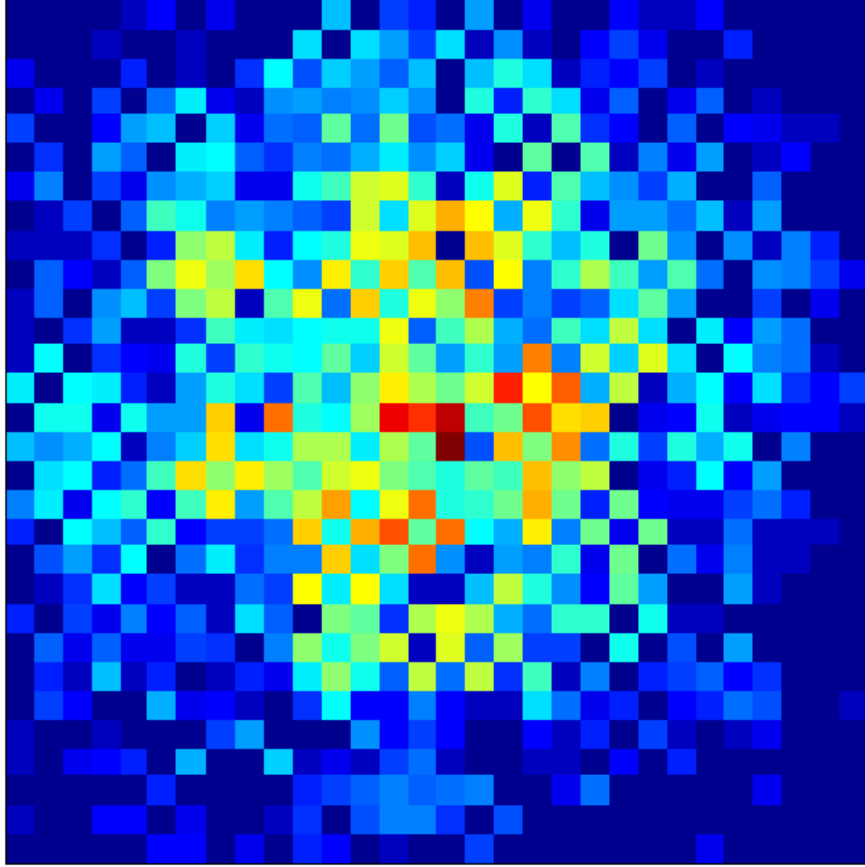
FIG. 6. Particle distribution across a section for a pulse injection into a $q = 3$ dimensional hetero-geneous medium for $\beta = 1/4$ and $\kappa = 1$. Different colors denote different values of concentration with a logarithmic color scheme.

This behavior persists as long as the particle displacement is smaller than the characteristic length scale ℓ, this means for times $t < \theta(\mathbf{0})\tau_\kappa$. This time scale is determined by the disorder at the injection point, which varies from realization to realization.

1. Infinite mean transition time ($0 < \beta < 1$)

Figure 3c compares the particle distributions at $t = 20\tau_\kappa$ after a pulse injection in a single $q = 1$–dimensional disorder realizations with the disorder ensemble average. The spatial particle distribution is characterized by multiple peaks which are due to the retention of particles in regions of high θ. Figure 3d shows the mean square displacement for different $q = 1$ dimensional disorder realizations. As predicted, at short times the mean square displacement increases in each realization linearly in time until $\theta(\mathbf{0})\tau_\kappa$ with $\theta(\mathbf{0})$ the retention

of the first pixel. At times larger than $\theta(\mathbf{0})\tau_\kappa$ the injected particles start sampling the medium heterogeneity of high and low θ values. Thus, diffusion at long times is slowed down due to particle retention in regions of high θ and the mean square displacement fluctuates around the ensemble mean behavior given by (35).

Figure 5 shows the particle distribution in a single realization of a $q = 2$ dimensional random medium. We observe similar localization features as in $q = 1$. The corresponding integrated particle distributions are compared in Figure 3a to the average distribution. The integrated profiles are naturally much smoother than for $q = 1$. The mean square displacement displayed in Figure 3b shows the exact behavior (38) for $t < \theta(\mathbf{0})\tau_\kappa$. As in $q = 1$, $m(t)$ slows down and fluctuates around the ensemble mean (35). The fluctuations are smaller than in $q = 1$ due to the more efficient disorder sampling. This feature is even more pronounced in $q = 3$.

Figure 6 shows the particle distribution in a section of a $q = 3$ dimensional medium.Again, we observe certain localization features, but the particle distribution is smoother than in $q = 2$ due to more efficient disorder sampling. The more efficient disorder sampling is also manifest in the behavior of the mean square displacement shown in 2. While the short time behavior depends strongly on the single medium properties, the long time behavior fluctuates much less about the mean behavior (33) than in $q < 3$.

In summary, fluctuations of the particle distribution and the mean square displacement are decreasing with increasing for spatial dimension. This is due to the increase in the sampling efficiency with q. As discussed in Section II B, the number R_n of possibles sites that may be explored evolves with step number as $R_n \propto n^{q/2-1}$ and the ability S_n of sampling new sites increases with as a function of q.

2. Finite mean transition time ($\beta > 1$)

Here we consider media characterized by the disorder distribution (28) with $\beta > 1$, which implies that the ensemble mean transition time $\overline{\tau}$ is finite. Figure 4b illustrates the mean square displacements in single $q = 1$ dimensional disorder realizations with $\beta = 3/2$. For compactness, we omit the behaviors for $q = 2$ and $q = 3$. At short times, for $t < \theta(0)\tau_\kappa$, $m(t)$ increases linearly with time as (38). At long times, the behaviors are approximately linear in all realizations and fluctuate around the disorder average (37). For $q = 2$ and

20

$q = 3$, the behaviors are similar, but the fluctuations around the mean behavior decrease with increasing spatial dimension (not shown).

IV. SELF-AVERAGING

The differences in the diffusion behaviors between disorder realizations depend on the efficiency by which the diffusing particles sample the heterogeneity in single disorder realizations. This efficiency is expressed by S_n, the average number of distinct sites visited after n random walk steps, and by the number of different disorder configurations particles are able to see in a single realization, which is measured by R_n defined by (23). As discussed in Section II B, in $q = 1$ dimension the number of different disorder configurations accessible to the diffusing particles within a single realization is of the order of 1. Thus, we expect that sample to sample fluctuations persist, while in $q \geq 2$, R_n increases with step number. In the following, we discuss how these attributes impact on the sample-to-sample fluctuations of the mean square displacement for anomalous as well as normal average transport.

The fluctuations of the mean square displacement $m(t)$ are quantified by its disorder variance

$$\sigma_m^2(t) = \overline{m(t)^2} - \overline{m}(t)^2 \tag{39}$$

where $m(t)$ for a single realization is given by (20). Again, we obtain an exact result for early times $t < \tau_\kappa / \theta_H$. Using expression (38) for $m(t)$ and (31) for its disorder average in (39), we obtain

$$\sigma_m^2(t) = 4q\kappa^2 t^2 \overline{\left(\theta^{-1} - \theta_H^{-1}\right)^2}. \tag{40}$$

For times $t \gg \tau_c$, we derive in Appendix B the following compact expression:

$$\sigma_m^2(t) = \ell^4 \overline{n_t(n_t - \overline{n}_t)/R_{n_t}}. \tag{41}$$

The mean square displacement is quantified in terms of the disorder averages of the renewal process $n_t = \max(n|t_n \leq t)$, which describes the number of steps needed to reach time t in the process t_n defined in (16). The self-averaging property is quantified by the relative variance [1, 49, 50]

$$\Sigma(t) = \frac{\sigma_m^2(t)}{\overline{m}(t)^2}. \tag{42}$$
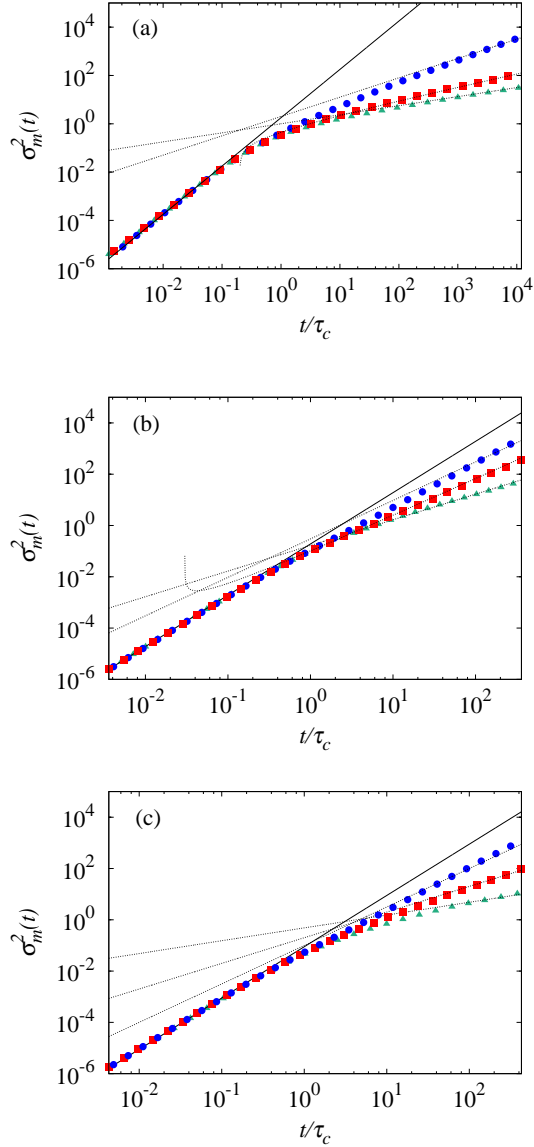
21

FIG. 7. Evolution of the variance $\sigma_m^2(t)$ for a pulse injection of 10^5 particles into 10^3 realizations of (triangles) $q = 3$, (squares) $q = 2$ and (circles) $q = 1$ dimensional random media with $\kappa = 1$ and (a) $\beta = 1/4$, (b) $\beta = 3/2$ and (c) $\beta = 5/2$. The solid black line shows the analytical early time behavior (40), the dashed black lines the late time behaviors (top) (43)–(45), (middle) (46)–(47) and (bottom) (49).

If $\lim_{t\to\infty} \Sigma(t) = 0$, diffusion is referred to as self-averaging, if $\lim_{t\to\infty} \Sigma(t) > 0$ as non self-averaging. Note that the fact that $\Sigma(t)$ goes to 0 with increasing time does not imply that the variance disappears. It may in fact be large. However, the ratio of fluctuation to mean value goes to 0, which means that the relative error of the disorder ensemble averaged quantity

22

disappears in the limit of infinite times. This property is defined here as self-averaging in agreement with Aslangul *et al.* [49] and Bouchaud and Georges [1], who consider the evolution of the relative variance over time, and Wiseman and Domani [50], who consider the evolution of the relative variance as a function of the spatial size of a disorder sample.

In the following, we discuss the variance of the mean square displacement $m(t)$ and its self-averaging behavior for disorder scenarios that are characterized by infinite and finite mean transition times. The detailed calculations leading to the results discussed in the following can be found in Appendix E and, for the case of infinite mean transition time in Ref. [41], which we summarize here for completeness.

A. Infinite mean transition time ($0 < \beta < 1$)

The disorder average mean square displacement for $0 < \beta < 1$ is sub-diffusive in any spatial dimension. The variance of $m(t)$ in $q = 1$ spatial dimensions behaves for $t \gg \tau_c$ as

$$\sigma_m^2(t) \propto t^{\frac{4\beta}{2\beta - q\beta + q}}. \tag{43}$$

Thus, using (43) and (35) for the relative variance we have $\Sigma(t) \sim 1$. This implies that the mean square displacement is not self-averaging for $q = 1$ as expected from the constant number of possible realities ($R_n \sim 1$) that may be sampled in a single realization. Nevertheless, (43) implies that $m(t) \propto \overline{m}(t) \propto t^{\frac{2\beta}{2\beta - q\beta + q}}$, as discussed in the previous section.

For $q = 2$, one finds the following long time behavior scaling:

$$\sigma_m^2(t) \propto t^{2\beta} \ln(t)^{1-2\beta}. \tag{44}$$

Here the relative variance $\Sigma(t) \propto 1/\ln(t)$, see (35). The process is self-averaging, but the relative variance decays very slowly. This behavior is in line with the slow increase of $S_n \sim \ln(n)$ of new sites that may be sampled.

For $q > 2$, the disorder variance of $m(t)$ scale as

$$\sigma_m^2(t) \propto t^{\beta(3-q/2)}. \tag{45}$$

This implies, together with (33), that the relative variance scales as $\Sigma(t) \propto t^{\beta(1-q/2)}$. The diffusion behavior is self-averaging and the decay of the relative variance decreases with the spatial dimension. Again, this behavior is reflected by the increasing sampling efficiency

23

expressed both by the ability of sampling new sites, which scales as $S_n \sim n$ and the number of different realities that may be sampled by the random walk, which scales as $R_n \sim n^{q/2-1}$.

Figure 7a shows the temporal evolution of $\sigma_m^2(t)$ sampled from numerical Monte-Carlo simulations using 10^4 realizations of the random medium for $\beta = 1/4$ and $q = 1$, 2 and 3. The analytical results (40) for early times and the late time scalings (43)–(44) are consistent with the numerical simulation data. At short times $t \ll \tau_c$ there is a quadratic increase of uncertainty of the mean square displacement, which is due to the poor disorder sampling; particles have not yet diffused over lengths larger than the characteristics disorder scale ℓ. For times $t > \tau_c$, particles start seeing the disorder and the variance increase slows down.

B.   Finite mean transition time ($\beta > 1$)

For finite mean transition time, the disorder average mean square displacement evolves linearly with time. This means that average particle motion is diffusive, i.e. $m(t) \propto t$ for any dimension q. The behavior of the variance of $m(t)$, however, depends on the spatial dimension and on the value of β. We distinguish the cases $1 < \beta < 2$ for which the mean transition time is finite while $\overline{\tau^2}$ is infinite, and the case $\beta > 2$ for which both the mean and the mean square transition time are finite. The results presented in the following are derived in Appendix E.

1.   Self-averaging property for $1 < \beta < 2$

For $1 < \beta < 2$ the variance of the transition time is infinite. The scaling of $\sigma_m^2(t)$ for $q = 1$ is:

$$\sigma_m^2(t) \propto (t/\overline{\tau})^{\frac{4-q(\beta-1)}{2}}. \tag{46}$$

This indicates that the variance increases with time faster than in the case $0 < \beta < 1$. The relative variance, however, decays algebraically with time as $\Sigma(t) \propto (t/\overline{\tau})^{-\frac{q(\beta-1)}{2}}$. The mean square displacement is self-averaging despite the poor disorder sampling in $q = 1$ for which the number of possible disorder realizations that may be sampled is $R_n \sim 1$. Disorder sampling is still poor as expressed by the slow decay of the relative variance. For $q = 2$, we

find:

$$\sigma_m^2(t) \propto (t/\overline{\tau})^{3-\beta} \ln(t/\overline{\tau})^{\beta-2}. \tag{47}$$

The relative variance here decreases as $\Sigma(t) \propto (t/\overline{\tau})^{1-\beta} \ln(t/\overline{\tau})^{\beta-2}$, i.e. slower than in the case $0 < \beta < 1$.

For $q > 2$ we obtain for $\sigma^2(t)$ the long time behavior

$$\sigma_m^2(t) \propto (t/\overline{\tau})^{4-\beta-q/2}. \tag{48}$$

Similarly to the case $q \leq 2$ dimensions, the disorder variance increases faster than in the case of an infinite mean transition time. The relative variance, however, scales as $\Sigma(t) \propto (t/\overline{\tau})^{2-\beta-q/2}$, which indicates that diffusion is self-averaging. These behaviors are illustrated in Figure 7 for $\beta = 3/2$ for $q = 1$, 2 and 3. At early times, we observe the characteristic quadratic increase, which slows down when the particles start experiencing the disorder.

2. *Self averaging property for $\beta > 2$*

For $\beta > 2$ both the mean and mean squared transition times are finite. In this case, one obtains for the disorder variance $\sigma_m^2(t)$:

$$\sigma_m^2(t) \propto (t/\overline{\tau})^{2-q/2}. \tag{49}$$

This implies that the relative variance behaves as $\Sigma(t) \propto (t/\overline{\tau})^{-q/2}$. As expected the behavior is always self-averaging, but the relative variance decreases slower with decreasing spatial dimension. Figure 7c illustrates the behavior of $\sigma_m^2(t)$ for $\beta = 5/2$ and $q = 1$, 2 and 3 spatial dimensions.

V. ERGODICITY

In this section we study the issue of ergodicity. While self-averaging refers to the sample to sample fluctuations of the mean square displacement $m(t)$ with respect to its disorder average $\overline{m}(t)$, the ergodicity property considers the fluctuations of the time averaged mean square displacement with respect to its noise average. In other words, self-averaging refers to the variability of noise averages between disorder realizations, while ergodicity refers to

variability of time averages between noise realizations (particles) in single disorder realizations.

The mean square displacement of $\mathbf{x}(s)$ in (8) is ergodic in operational time s because it describes an ordinary random walk characterized by a stationary increment. This is different for the process $\mathbf{x}(t) \equiv \mathbf{x}[s(t)]$ that in clock time samples the heterogeneity slower. Ergodicity is related to the ability of a single particle to sample the available fluctuation spectrum in a finite time. The time process here is directly related to the disorder process as given in (8). Thus, disorder sampling in time is limited by the disorder itself. In order to probe the ergodicity of the random walk, we consider the time average particle increments [35]

$$m_\Delta(t) = \frac{1}{t - \Delta} \int_0^{t-\Delta} dt' \left[\mathbf{x}(t' + \Delta) - \mathbf{x}(t') \right]^2 . \tag{50}$$

It is interesting to note that this quantity is related to the empirical variogram [51], which is used in geostatistics to characterize the variability of a spatial random field. Eq. (50) is typically referred to as time average mean square displacement in the literature. Here, we refer to it as the time average of the increment process $\mathbf{x}_\Delta(t) \equiv \mathbf{x}(t + \Delta) - \mathbf{x}(t)$. Using (8), we can write

$$\mathbf{x}_\Delta(t) = \sqrt{2\kappa}\, \boldsymbol{w}_\Delta(t), \qquad\qquad \boldsymbol{w}_\Delta(t) = \int_{s(t)}^{s(t+\Delta)} ds'\, \boldsymbol{\eta}(s') \tag{51}$$

with $\langle \boldsymbol{w}_\Delta(t) \rangle = \mathbf{0}$ and $\langle \boldsymbol{w}_\Delta(t)^2 \rangle = q[s(t + \Delta) - s(t)]$. Thus, we may write (50) in a weak sense as

$$m_\Delta(t) = \frac{2q\kappa}{t - \Delta} \int_0^{t-\Delta} dt' \left[s(t' + \Delta) - s(t') \right] . \tag{52}$$

We furthermore expand this expression for $t \gg \Delta$ in order to obtain

$$m_\Delta(t) = \frac{2q\Delta\kappa}{t} s(t) + \dots , \tag{53}$$

where the dots denote contributions of order Δ^2. We define now the quantity

$$m_a(t) \equiv \frac{m_\Delta(t)}{\Delta} t = 2q\kappa s(t), \tag{54}$$

which can be seen as the time averaged mean square displacement while $m_\Delta(t)/2\Delta$ denotes an apparent diffusion coefficient. Note that Lubelski $et~al.$ [36] suggest to consider $m_\Delta(t)/2t$

as an effective diffusion coefficient. These authors focus on the displacement increment as a function of the time increment Δ, while here, we study the long time behavior of the time-averaged mean square displacement. According to (5), the noise mean of $m_a(t)$ is equal to the mean square displacement $\langle m_a(t) \rangle = m(t)$. The process $s(t)$ is identified in the coarse grained random walk (9) by $s(t) \equiv \tau_\kappa n_t$ such that

$$m_a(t) = \ell^2 n_t. \tag{55}$$

The disorder average of $m_a(t)$ is equal to $\overline{m}(t)$. Note that Miyaguchi and Akimoto [38] defined a diffusion coefficient as $D_t = n_t/t$, which is related to $m_a(t)$ as $D_t = m_a(t)/\ell^2 t$. These authors study the full distribution of D_t over the disorder ensemble as well as its disorder mean and variance. The disorder mean of D_t is related to the disorder mean square displacement $\overline{m}(t)$ as $\overline{D}_t = \overline{m}(t)/\ell^2 t$.

Ergodicity of the mean square displacement is measured in terms of the variance of $m_a(t)$ with respect to its noise average $m(t)$,

$$\sigma_a^2(t) = \left\langle [m_a(t) - m(t) \rangle]^2 \right\rangle = \ell^4 \left(\langle n_t^2 \rangle - \langle n_t \rangle^2 \right). \tag{56}$$

Let us first consider the short time behavior of $\sigma_a^2(t)$. The variance of $m_a(t)$ in a single realization is given by

$$\sigma_a^2(t) = \frac{\ell^4 t}{\tau_\kappa \theta(\mathbf{0})} = \ell^2 m(t), \tag{57}$$

because at the first step in a single medium realization, the transition time is exponentially distributed with mean $\theta(\mathbf{0})\tau_\kappa$. This renders n_t approximately a Poisson variable as long as $\langle n_t \rangle \leq 1$. Thus, in the diffusive short time regime, $\sigma_a^2(t)$ behaves as the mean square displacement. Its disorder average is given at short times $t < \tau_\kappa$ by

$$\overline{\sigma}_a^2(t) = \ell^2 \frac{2q\kappa t}{\theta_H}. \tag{58}$$

In order to quantify this variance in a systematic way at long times $t \gg \tau_c$, we consider the disorder average of (56), which can be tautologically rewritten as

$$\overline{\sigma}_a^2(t) = \ell^4 \left(\overline{\langle n_t^2 \rangle} - \overline{\langle n_t \rangle^2} \right) - \ell^4 \left(\overline{\langle n_t \rangle^2} - \overline{\langle n_t \rangle}^2 \right). \tag{59}$$
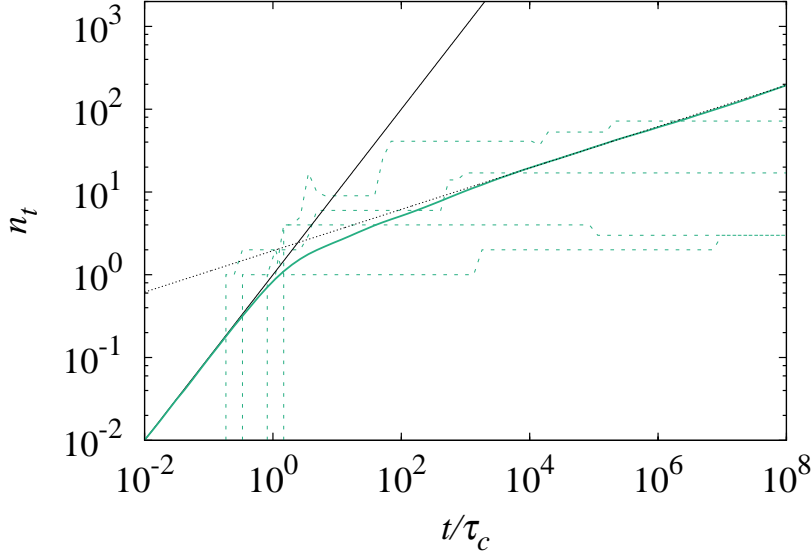
FIG. 8. Evolution of the time-averaged mean square displacement $m_a(t)$ defined by (55) (dashed) for 4 particle trajectories, and the noise average $m(t)$ for (solid) 10^5 particles injected at the origin of a $q = 3$ heterogeneous medium for $\beta = 1/4$. The solid black line shows the analytical early time behavior (38), the dashed black lines the late time behavior (33).

The second term on the right is identified with the variance $\sigma_m^2(t)$ defined by (39). The first term on the right is the disorder variance of the step number process n_t, which is independent of the noise in a single disorder realization,

$$\sigma_n^2(t) = \overline{n_t^2} - \overline{n_t}^2. \tag{60}$$

Thus, $\overline{\sigma}_a^2(t)$ is given by

$$\overline{\sigma}_a^2(t) = \ell^4 \sigma_n^2(t) - \sigma_m^2(t). \tag{61}$$

This implies that $\sigma_n^2(t) \geq \sigma_m^2(t)/\ell^4$ due to the positiveness of the variance. Thus, the long time behavior of $\overline{\sigma}_a^2(t)$ is dominated by $\sigma_n^2(t)$. The ratio of $\sigma_n^2(t)$ and $\overline{n_t}^2$ has been termed ergodicity breaking parameter EB in the literature [52],

$$\text{EB} = \frac{\sigma_n^2(t)}{\overline{n_t}^2}. \tag{62}$$

If the value of EB is nonzero in the limit of infinite times, one speaks of weak ergodicity breaking. Note that $\sigma_n^2(t) = [\overline{m_a(t)^2} - \overline{m}(t)^2]/\ell^4$ is proportional to the disorder variance of the

28

time average mean square displacement. Note also, that EB is equal to the relative disorder variance of the diffusion coefficient D_t studied by Miyaguchi and Akimoto [38]. Specifically, the variance $\sigma_n^2(t)$ is related to the disorder variance of D_t as $\sigma_n^2(t) = (\overline{D_t^2} - \overline{D}_t^2)t^2$.

Appendix C derives the asymptotic long time behaviors for the moments and the PDF of n_t with respect to the disorder average for q dimensions and for disorder scenarios that are characterized by infinite and finite mean transition times. In the following, we discuss $m_a(t)$ and its disorder variance in terms of $\sigma_n^2(t)$ for different values of q and β.

A. Infinite mean transition time ($0 < \beta < 1$)

As outlined above, for infinite mean transition time, $m(t)$ evolves subdiffusively with time in any spatial dimension. The time average mean square displacement $m_a(t)$ show subdiffusive scaling as illustrated in Figure 8 with significant fluctuations around the noise mean $m(t)$. As outlined above, these fluctuations can be quantified in terms of the disorder variance $\sigma_n^2(t)$. Miyaguchi and Akimoto [38] report results for the variance of D_t, which is equivalent to $\sigma_n^2(t)$. Also, these authors report explicit expressions for EB. In the following,
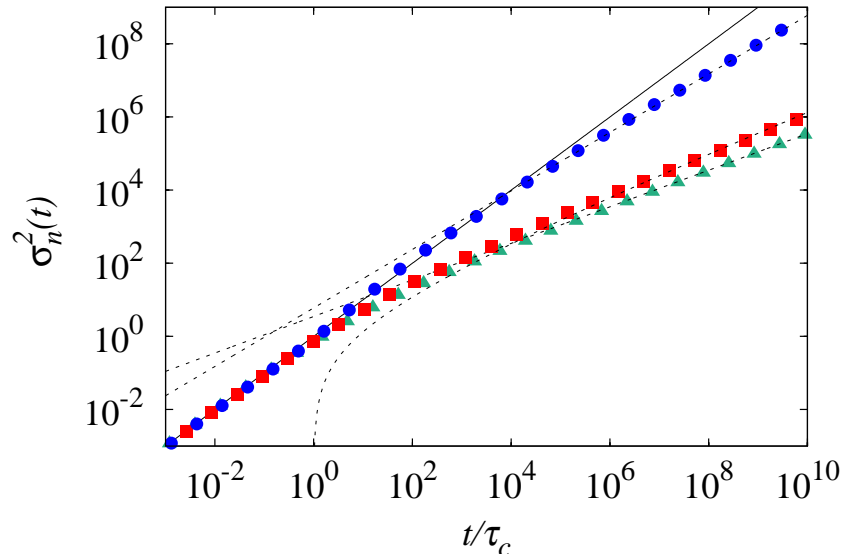

FIG. 9. Evolution of the variance $\sigma_n^2(t)$ for a pulse injection into 10^3 realizations of a (circles) $q = 1$, (squares) $q = 2$ and (triangles) $q = 3$ random medium for $\beta = 1/4$. The solid black line shows the analytical early time behavior (58), the dashed black lines the late time behaviors (63)–(65).
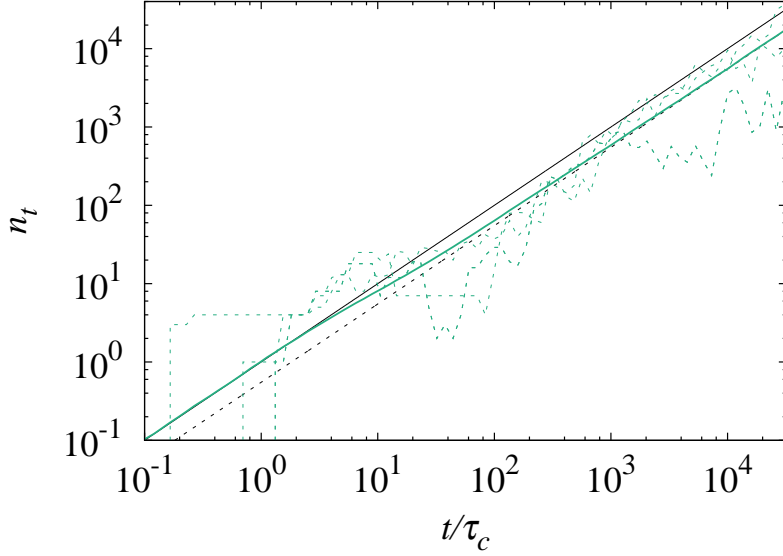
FIG. 10. Evolution of the time-averaged mean square displacement $m_a(t)$ defined by (55) (dashed) for 4 particle trajectories, and the noise average $m(t)$ for (solid) 10^5 particles injected at the origin of a $q = 3$ heterogeneous medium for $\beta = 3/2$. The solid black line shows the analytical early time behavior (38), the dashed black lines the late time behavior (37).

we summarize these results for completeness before we discuss the behavior of $\sigma_n^2(t)$ and thus $\overline{\sigma}_a^2(t)$ for disorder distributions characterized by finite mean transition times.

For $q = 1$, the scaling of the disorder variance $\sigma_n^2(t)$ is given by

$$\sigma_n^2(t) \propto t^{\frac{4\beta}{2\beta - q\beta + q}}, \tag{63}$$

see also (D6). It scales in the same way as the disorder variance of $m(t)$ given in (43). As $\overline{n}_t \propto t^{\frac{2\beta}{2\beta - q\beta + q}}$, the ergodicity breaking parameter is EB ~ 1. This means that diffusion is here weakly ergodicity breaking.

For $q = 2$ dimensions, $\sigma_n^2(t)$ behaves as

$$\sigma_n^2(t) \propto t^{2\beta} \ln(t)^{2-2\beta}, \tag{64}$$

see also (D9). This implies that the disorder variance of $\overline{\sigma}_a^2(t)$ increases faster than $\sigma_m^2(t)$ given by (44). In fact, $\sigma_n^2(t)$ scales as $\overline{n}_t^2$. Thus the ergodicity breaking parameter EB ~ 1 and again, diffusion is weakly ergodicity breaking.

For $q > 2$, the scaling for the disorder variance of n_t is

$$\sigma_n^2(t) \propto t^{2\beta}, \tag{65}$$

see also (D11). The variance $\sigma_n^2(t)$ scales in the same way as $\bar{n}_t^2$ and thus, again EB ~ 1. Also in $q = 3$, diffusion is weakly ergodicity breaking. These behaviors are illustrated in Figure (9) and compared with data from numerical random particle tracking simulations. Explicit expressions for EB can be found in Miyaguchi and Akimoto [38].

B. Finite mean transition time ($\beta > 1$)

For disorder scenarios that are characterized by finite mean transition times $\bar{\tau} < \infty$, the average diffusion behavior is normal and the diffusion is asymptotically self-averaging. The time average mean square displacement $m_a(t)$ evolves approximately linear but fluctuates about the noise mean $m(t)$ as illustrated in Figure 10.

For $q = 1$, we obtain from (D14) for $\sigma_n^2(t)$ the following scaling behavior:

$$\sigma_n^2(t) \propto (t/\langle\tau\rangle)^{2 + \frac{q(1-\beta)}{2}}. \tag{66}$$
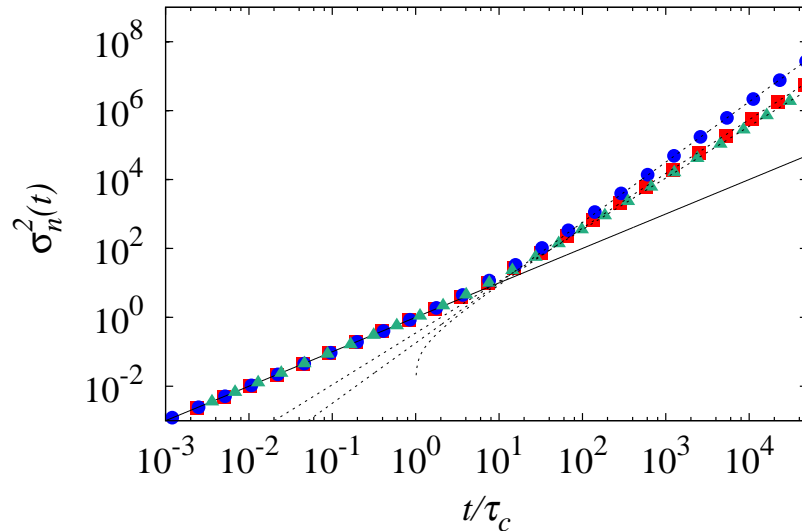

FIG. 11. Evolution of the variance $\sigma_n^2(t)$ for a pulse injection into 10^3 realizations of a (circles) $q = 1$, (squares) $q = 2$ and (triangles) $q = 3$ random medium for $\beta = 3/2$. The solid black line shows the analytical early time behavior (58), the dashed black lines the late time behaviors (66)–(68).

31

The behavior for $\beta > 2$ is obtained by setting $\beta = 2$. Since the disorder mean $\overline{n}_t$ scales as t, the ergodicity breaking parameter behaves as EB $\propto (t/\langle\tau\rangle)^{\frac{q(1-\beta)}{2}}$. This means it decreases relatively slowly with time as $t^{-\gamma}$, where $0 < \gamma \leq q/2$.

For $q = 2$, expression (D17) gives:

$$\sigma_n^2(t) \propto (t/\overline{\tau})^{3-\beta} \ln(t/\overline{\tau})^{\beta-1} \tag{67}$$

Again, the behavior for $\beta > 2$ is obtained by setting $\beta = 2$. The ergodicity breaking parameter here scales as EB $\propto [(t/\overline{\tau})/\ln(t/\overline{\tau})]^{1-\beta}$. Thus, it decays as $[t/\ln(t)]^{-\gamma}$ with $0 < \gamma \leq 1$.

For $q > 2$, we obtain from (D20)

$$\sigma_n^2(t) \propto (t/\langle\tau\rangle)^{3-\beta}. \tag{68}$$

The ergodicity breaking parameter scales as EB $\propto (t/\overline{\tau})^{1-\beta}$. It decays as $t^{-\gamma}$ with $0 < \gamma \leq 1$. Thus, while the time average mean square displacement is asymptotically ergodic, its variance decays algebraically with time. Figure 11 shows the evolution of the disorder variance of n_t in $q = 1$, 2 and 3 dimensions from numerical random walk simulations.

For disorder scenarios characterized by a finite mean transition time, diffusion is asymptotically ergodic and self-averaging, as discussed in Section IV. Thus, the ensemble of particles is representative of individual particle trajectories (ergodicity), and the disorder ensemble is representative of the diffusion behavior in single realization (self-averaging).

VI. SUMMARY AND CONCLUSIONS

We study self-averaging and ergodicity of diffusion in heterogeneous media. Ergodicity refers to the question whether the particle ensemble statistics in a single realization are representative of the temporal statistics sampled along a single particle trajectory. It is a fundamental property for the inference of stochastic particle rules from single particle tracking observations. The notion of ergodicity is different from the the notion of self-averaging. The latter refers to the representation of the particle statistics in a single medium realization though the disorder ensemble statistics. It is a fundamental condition for the use of stochastic modeling for the prediction of diffusion behavior in heterogeneous media. If these properties apply, they are approached asymptotically.

We analyze these behaviors for diffusion in heterogeneous media that are characterized by a spatially variable retardation coefficients, which quantifies particle retention due to physical and chemical interactions with the host medium. The retardation coefficient is modeled as a stationary spatial random field. The diffusion problem is coarse-grained on the characteristic heterogeneity length scale and numerically implemented using a time-domain random walk (TDRW). We derive an explicit expression for the noise average in a single realization which depends on the number R_n of different disorder realizations that the noise ensemble may sample after n diffusion steps in a single disorder realization. As is well known, for $q > 2$, the average particle motion in this model describes a continous time random walk (CTRW). We derive the renormalised space and time increments to match the long time diffusion behavior obtained from the TDRW simulations for $q = 3$.

We then focus on the mean-square particle displacement and particle distributions averaged over the disorder ensemble and in single disorder realizations. We derive explicit expressions for the early time behavior of the mean square displacement and confirm through the numerical simulations the known analytical results for the asymptotic scaling of the mean-square displacement and disorder average particle distribution for $q \geq 1$ spatial dimensions. The disorder ensemble average behavior is contrasted to the evolution of the mean-square displacement and particle distributions in single realizations. We observe significant sample to sample fluctuations and segregation of the particle distribution due to localized strong retention, which are more significant for decreasing spatial dimension. These observations are discussed in the light of the sampling efficiency R_n, which depends on the spatial dimension q; R_n increases faster with n for increasing q. For $q = 1$, $R_n \sim 1$, this means, particles sample in a average only one reality per realization. For $q = 2$, R_n increase logarithmically with step number and in $q = 3$ as $\sqrt{n}$.

The sample to sample fluctuations are quantified in terms of the disorder variance of the mean-square displacement between disorder realizations. The self-averaging behavior is probed by the relative variance $\Sigma(t)$. If $\Sigma(t)$ goes asymptotically to zero, $m(t)$ is self-averaging. This does not mean, however, that the fluctuations of $m(t)$ are necessarily small at finite times. We derive explicit expressions for the variance in q dimensions at early and late times, which are compared to TDRW simulations of the full heterogeneous diffusion problem for subdiffusive and diffusive disorder configurations. Under subdiffusive conditions, the mean square displacement is non self-averaging for $q = 1$ and self-averaging for $q \geq 2$.

Under diffusive conditions, i.e., when $\overline{m}(t) \propto t$, we distinguish scenarios for which the variance of the transition time is infinite and finite. Under both conditions, $m(t)$ is self-averaging in q dimension. The rate of convergence of $\Sigma(t)$ depends on the dimension of space and the heterogeneity distribution.

We study ergodicity in terms of the noise variance of the time-averaged mean square displacement $m_a(t)$, which is based on the definition of the time-averaged particle displacement during a time increment Δ. In a single realization $m_a(t) \propto n_t$, the number of steps a particle performs to reach time t. Due to the sample to sample fluctuations of the diffusion behavior between disorder realization, also the noise variance of $m_a(t)$ is a fluctuating quantity. Thus, we probe the ergodicity of $m_a(t)$ in terms of the disorder average of its variance, which is dominated by the disorder variance of the step number n_t. The ergodicity breaking parameter EB is defined as the ratio between the disorder variance of n_t and $\overline{n}_t^2$ consistent with the definition by Bel and Barkai [35] in the framework of CTRW. Ergodicity is studied both for diffusive and subdiffusive disorder configurations. We derive explicit analytical expressions for EB. Under subdiffusive conditions, the mean square displacement is weakly ergodicity breaking in all dimensions. Under diffusive conditions, for which the mean transition time is finite, diffusion is ergodic. The convergence toward ergodicity depends also here on the spatial dimension and disorder distribution.

In conclusion, under diffusive conditions, this means when the mean transition time is finite, the mean square displacement is both ergodic and self-averaging. Under subdiffusive conditions, on the other hand, the mean square displacement is weakly ergodicity breaking in all dimensions, while at the same time it is self-averaging for $d > 2$. This means, the noise ensemble is not representative of the single trajectory time statistics, but the disorder ensemble is representative for the noise ensemble in a single disorder realization.

Appendix A: Equivalent CTRW for $q > 2$ spatial dimensions

One difference between the TDRW and the CTRW arises from the correlation of successive time increments. Here we re-sample the trajectories in (13) such that the trajectories are constituted by series of S_n independent rescaled space and time increments as

$$\mathbf{x}_{S_n} = \sum_{i=1}^{S_n} \hat{\ell}_n \boldsymbol{\eta}_i, \qquad\qquad t_{S_n} = \sum_{i=1}^{S_n} \frac{n}{S_n} \tau_i. \qquad (A1)$$

We require that the variance of $\mathbf{x}_{S_n}$ be equal to the variance of $\mathbf{x}_n$, which implies

$$\hat{\ell}_n^2 S_n \equiv \ell^2. \qquad (A2)$$

Thus, we obtain directly $\hat{\ell}_n = \sqrt{n/S_n}\,\ell$. This gives for (A1)

$$\mathbf{x}_{S_n} = \sum_{i=1}^{S_n} \sqrt{\frac{n}{S_n}}\,\ell \boldsymbol{\eta}_i, \qquad\qquad t_{S_n} = \sum_{i=1}^{S_n} \frac{n}{S_n} \tau_i. \qquad (A3)$$

In order to obtain (27) for $q > 2$, we substitute $S_n = bn$ in (A3) and set $N = \lfloor bn \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function.

Appendix B: Disorder variance of mean square displacement

The disorder variance of the mean square displacement is defined by (39). In the following, we derive expression (41). To this end, we consider the second disorder moment of $m(t)$. We obtain by using (20)

$$\overline{m(t)^2} = \sum_{n,k=0}^{\infty} nk\ell^4 \frac{1}{R_n} \frac{1}{R_k} \sum_{i=1}^{R_n} \sum_{j=1}^{R_k} \overline{\mathbb{I}\left[t_n^{(i)} \le t < t_{n+1}^{(i)}\right] \mathbb{I}\left[t_k^{(j)} \le t < t_{k+1}^{(j)}\right]}. \qquad (B1)$$

By definition, the $t_n^{(i)}$ and $t_k^{(j)}$ are independent for $i \ne j$. For $i = j$ we have by virtue of the impulse functions that $n = k$. Thus, we can write $\overline{m(t)^2}$ as

$$\overline{m(t)^2} = \ell^4 \sum_{n=0}^{\infty} \frac{n^2}{R_n} p_n(t) + \Lambda(t), \qquad (B2)$$

where we defined

$$\Lambda(t) = \sum_{n,k=0}^{\infty} nk\ell^4 \frac{1}{R_n} \frac{1}{R_k} \sum_{i=1}^{R_n} \sum_{j=1,j\ne i}^{R_k} p_n(t) p_k(t). \qquad (B3)$$

We used here definition (26) of the number of steps needed to arrive at time t. $\Lambda(t)$ can be written as

$$\Lambda(t) = \sum_{n \geq k}^{\infty} nk\ell^4 \left(1 - \frac{1}{R_n}\right) p_n(t)p_k(t)$$

$$+ \sum_{n < k}^{\infty} nk\ell^4 \left(1 - \frac{1}{R_k}\right) p_n(t)p_k(t) \tag{B4}$$

and due to symmetry, we obtain more compactly

$$\Lambda(t) = \sum_{n,k=1}^{\infty} nk\ell^4 \left(1 - \frac{1}{R_n}\right) p_n(t)p_k(t). \tag{B5}$$

By comparison with (30), we identify

$$\Lambda(t) = \overline{m}(t)^2 - \overline{m}(t)\ell^2 \sum_{n=1}^{\infty} \frac{n}{R_n} p_n(t) \equiv \overline{m}(t)^2 - \overline{m}(t)\ell^2 \overline{n_t/R_{n_t}}. \tag{B6}$$

Inserting the latter into (B2) gives

$$\overline{m(t)^2} = \overline{m(t)}^2 + \ell^4 \left(\overline{n_t^2/R_{n_t}} - \overline{n_t}\,\overline{n_t/R_{n_t}}\right), \tag{B7}$$

where we note that

$$\overline{n_t^2/R_{n_t}} = \sum_{n=0}^{\infty} \frac{n^2}{R_n} p_n(t). \tag{B8}$$

and that $\overline{m}(t) = \ell^2 \overline{n_t}$, see (30). Using (B7) in (39) for the variance of the mean square displacement gives (41).

Appendix C: Distribution of number n_t of random walk steps

In order to determine the long-time behaviors of the disorder average mean square displacement, its variance and the variance of the time-average mean square displacement, we need to determine averages of function of the number n_t of random walk steps. Its distribution is given by

$$p_n(t) = \overline{\delta_{n,n_t}} = \overline{\mathbb{I}(t_n \leq t < t_{n+1})}. \tag{C1}$$

Its Laplace transform is given by

$$p_n^*(\lambda) = \lambda^{-1}\overline{[\exp(-\lambda t_n) - \exp(-\lambda t_{n+1})]}, \tag{C2}$$

where the times t_n and t_{n+1} here are

$$t_n = \sum_{i=1}^{S_n} \gamma_n \tau_i, \qquad\qquad t_{n+1} = \sum_{i=1}^{S_{n+1}} \gamma_{n+1} \tau_i. \qquad (C3)$$

We defined for convenience $\gamma_n \equiv n/S_n$. Using these expressions in (C2) and performing the disorder average we obtain

$$p_n^*(\lambda) = \lambda^{-1} \left[f_n^*(\lambda) - f_{n+1}^*(\lambda) \right] \approx -\lambda^{-1} \frac{d}{dn} f_n^*(\lambda), \qquad (C4)$$

where we defined

$$f_n^*(\lambda) \equiv \psi(\lambda \gamma_n)^{S_n}. \qquad (C5)$$

We consider here transition time PDFs $\psi(t)$ characterized by the power-law behavior (29) as well as distributions for which all moments exist. We distinguish the cases $0 < \beta < 1$, $1 < \beta < 2$ and $\beta > 2$ in (29). The latter covers the leading behavior expected if all moments of τ exist.

1. Infinite mean transition time: $0 < \beta < 1$

The Laplace transform of $\psi(t)$ can be expanded for small λ as

$$\psi^*(\lambda) = 1 - a_\beta \lambda^\beta + \dots. \qquad (C6)$$

where a_β a constant that depends on the specific form of $\psi(t)$. Inserting (C6) into (C4) gives for $\lambda \tau_\kappa \ll 1$

$$f_n^*(\lambda) \approx \exp\left(-a_\beta \alpha_n \lambda^\beta \right), \qquad (C7)$$

where we defined for convenience $\alpha_n = \gamma_n^\beta S_n$.

2. Finite mean transition time: $\beta > 1$

For $1 < \beta < 2$, this means, if the mean transition time is finite $\langle \tau \rangle < \infty$ and $\langle \tau^2 \rangle = \infty$, the Laplace transform of $\psi(t)$ can be expanded as

$$\psi^*(\lambda) = 1 - \langle \tau \rangle \lambda + b_\beta \lambda^\beta + \dots. \qquad (C8)$$

where b_β depends on the specific shape of $\psi(t)$. Inserting (C8) into (C4) gives

$$f_n^*(\lambda) = \psi^*(\lambda \gamma_n)^{S_n} \approx \exp\left(-n\lambda\langle\tau\rangle\right)\left(1 + b_\beta \alpha_n \lambda^\beta + \dots\right), \tag{C9}$$

where the dots denote sub-leading contributions.

For $\beta > 2$, this mean, if both the mean and mean squared transition time are finite, the expansion of $\psi^*(\lambda)$ is

$$\psi^*(\lambda) = 1 - \langle\tau\rangle\lambda + \frac{\langle\tau^2\rangle}{2}\lambda^2 + \dots, \tag{C10}$$

where the dots denote sub-leading contributions. Inserting (C10) into (C4) gives

$$f_n(\lambda) \equiv \psi(\lambda\gamma_n)^{S_n} \approx \exp\left(-n\lambda\langle\tau\rangle\right)\left(1 + \frac{n^2}{S_n}\langle\tau^2\rangle\lambda^2 + \dots\right), \tag{C11}$$

where the dots denote sub-leading contributions. This expansion is valid for all transition time PDFs, for which the first and second moments exist.

Appendix D: Disorder moments of number n_t of random walk steps

The kth disorder moment of the step number n_t is defined by by

$$h_k(t) = \sum_{n=0}^{\infty} n^k p_n(t). \tag{D1}$$

The Laplace transform of $h_k(t)$ can with the above definitions be approximated as

$$h_k^*(\lambda) \approx -\lambda^{-1} \int_0^\infty dn n^k \frac{d}{dn} f_n^*(\lambda) \tag{D2}$$

Integration by parts gives

$$h_k^*(\lambda) \approx k\lambda^{-1} \int_0^\infty dn n^{k-1} f_n^*(\lambda). \tag{D3}$$

1. Infinite mean transition time: $0 < \beta < 1$

We write (D3) by using (C7) as

$$h_k^*(\lambda) \approx k\lambda^{-1} \int_0^\infty dn n^{k-1} \exp\left(-a_\beta \alpha_n \lambda^\beta\right). \tag{D4}$$

38

a. Spatial dimension $q = 1$

In $q = 1$, $S_n \sim n^{q/2}$ and $\gamma_n = n^{1-q/2}$, therefore $\alpha_n = n^{\frac{2\beta - d\beta + d}{2}}$. We obtain by rescaling n

$$h_k^*(\lambda) \approx \lambda^{-1-\frac{2k\beta}{2\beta-d\beta+d}} k \int_0^\infty dn\, n^{k-1} \exp\left(-a_\beta n^{\frac{2\beta-d\beta+d}{2}}\right). \tag{D5}$$

Inverse Laplace transform gives

$$h_k(t) \propto t^{\frac{2k\beta}{2\beta-d\beta+d}}. \tag{D6}$$

b. Spatial dimension $q = 2$

In $q = 2$, $S_n \sim n/\ln(n)$ and $\gamma_n = \ln(n)$, therefore $\alpha_n = n\ln(n)^{\beta-1}$ such that

$$h_k^*(\lambda) \approx \lambda^{-1} k \int_0^\infty dn\, n^{k-1} \exp\left[-a_\beta \lambda^\beta n \ln(n)^{\beta-1}\right]. \tag{D7}$$

We rescale now $n \to n\lambda^\beta \ln(1/\lambda)^{\beta-1}$. Thus, we obtain in leading order in the limit $\lambda \to 0$

$$h_k^*(\lambda) \approx \lambda^{-1-k\beta} \ln(1/\lambda)^{k-k\beta} k \int_0^\infty dn\, n^{k-1} \exp\left(-a_\beta \beta^\beta n\right). \tag{D8}$$

Inverse Laplace transform gives

$$h_k(t) \propto t^{k\beta} \ln(t)^{k-k\beta}. \tag{D9}$$

c. Spatial dimension $q > 2$

In $q > 2$, $S_n \sim n$ and $\gamma_n = 1$, therefore $\alpha_n = n$ such that

$$h_k^*(\lambda) \approx \lambda^{-1-\beta} k \int_0^\infty dn\, n^{k-1} \exp\left(-a_\beta n\right). \tag{D10}$$

Inverse Laplace transform gives

$$h_k(t) \propto t^{k\beta}. \tag{D11}$$

2. Finite mean transition time $\beta > 1$

For $1 < \beta < 1$, we use (C9) in (C4) in the Laplace transform of (30) in order to obtain

$$h_k^*(\lambda) \approx \lambda^{-1}k \int_0^\infty dn\, n^{k-1} \exp(-\lambda n\bar{\tau})$$
$$+ \lambda^{-1}k \int_0^\infty dn\, n^{k-1} \exp(-\lambda n\bar{\tau})b_\beta \alpha_n \lambda^\beta + \dots, \tag{D12}$$

The case $\beta > 2$ is obtained according to (C11) by setting $\beta = 2$ and $b_\beta = b_2 = \langle \tau^2 \rangle / 2$.

a. $q = 1$

Using $\alpha_n = n^{\frac{2\beta - d\beta + d}{2}}$ and rescaling of n, we obtain

$$h_k^*(\lambda) \approx \lambda^{-1-k}k \int_0^\infty dn\, n^{k-1} \exp(-n\bar{\tau})$$
$$+ \lambda^{-1-k-\frac{d(1-\beta)}{2}}k \int_0^\infty dn\, n^{k-1+\frac{2\beta-d\beta+d}{2}} \exp(-n\bar{\tau})b_\beta. \tag{D13}$$

Inverse Laplace transform gives

$$h_k(t) = (t/\bar{\tau})^k + A_1(t/\langle \tau \rangle)^{k+\frac{d(1-\beta)}{2}}, \tag{D14}$$

where we defined

$$A_1 = \frac{b_\beta}{\bar{\tau}^\beta} \frac{\Gamma(k + \frac{2\beta - d\beta + d}{2})}{\Gamma(k + 1 + \frac{d(1-\beta)}{2})} \tag{D15}$$

b. $q = 2$

Using $\alpha_n = n \ln(n)^{\beta-1}$ and rescaling of n, we obtain in leading order

$$h_k^*(\lambda) \approx \lambda^{-1-k}k \int_0^\infty dn\, n^{k-1} \exp(-n\bar{\tau})$$
$$+ \lambda^{-2-k+\beta} \ln(1/\lambda)^{\beta-1}k \int_0^\infty dn\, n^k \exp(-n\bar{\tau})b_\beta. \tag{D16}$$

40

Inverse Laplace transform gives

$$h_k(t) = (t/\bar{\tau})^k + A_2(t/\bar{\tau})^{1+k-\beta} \ln(t/\bar{\tau})^{\beta-1}. \tag{D17}$$

where we defined

$$A_2 = \frac{b_\beta}{\bar{\tau}^\beta} \frac{k\Gamma(k+1)}{\Gamma(k+2-\beta)}. \tag{D18}$$

c. $q > 2$

Using $\alpha_n = n$ and rescaling of n, we obtain

$$h_k^*(\lambda) \approx \lambda^{-1-k} k \int_0^\infty dn\, n^{k-1} \exp(-n\bar{\tau})$$
$$+ \lambda^{-2-k+\beta} k \int_0^\infty dn\, n^k \exp(-n\bar{\tau}) b_\beta. \tag{D19}$$

Inverse Laplace transform gives

$$h_k(t) = (t/\bar{\tau})^k + A_2(t/\langle\tau\rangle)^{1+k-\beta}. \tag{D20}$$

Appendix E: Variance of the mean square displacement

In order to determine the behavior of the variance $\sigma_m^2(t)$ of the mean square displacement $m(t)$, we need to evaluate (41), which involves the disorder moments $\overline{n_t^2/R_{n_t}}$ and $\overline{n_t/R_{n_t}}$ of the random step number n_t.

1. Spatial dimension $q = 1$

For $q = 1$, $R_n = 2^{q/2}$. Thus, we obtain for $\sigma_m^2(t)$

$$\sigma_m^2(t) = \frac{\ell^4}{2^{q/2}} \left(\overline{n_t^2} - \overline{n_t}^2 \right) = \frac{\ell^4}{2^{q/2}} \left[h_2(t) - h_1(t)^2 \right]. \tag{E1}$$

a. Mean transition time infinite: $0 < \beta < 1$

We obtain from the previous section that

$$\sigma_m^2(t) \propto t^{\frac{4\beta}{2\beta-d\beta+d}} \tag{E2}$$

41

b.   Mean transition time finite: $\beta > 1$

For $1 < \beta < 2$, we obtain from the previous section

$$\sigma_m^2(t) \propto (t/\overline{\tau})^{2 + \frac{q(1-\beta)}{2}} \tag{E3}$$

This means, for $\beta > 2$ that

$$\sigma_m^2(t) \propto (t/\overline{\tau})^{2 - \frac{q}{2}}. \tag{E4}$$

2.   Spatial dimension $q = 2$

For $q = 2$, $R_n = 2\ln(n)$. Thus, we obtain for $\sigma_m^2(t)$

$$\sigma_m^2(t) = \ell^4 \left(\overline{n_t^2/\ln(n_t)} - \overline{n_t/\ln(n_t)\overline{n_t}} \right). \tag{E5}$$

This means, we need to determine

$$J_k(t) = \sum_{n=0}^{\infty} \frac{n^k}{\ln(n)} p_n(t) \tag{E6}$$

for $k = 1, 2$. We obtain from (C4) for its Laplace transform

$$J_k^*(\lambda) \approx -\lambda^{-1} \int_0^{\infty} dn \frac{n^k}{\ln(n)} \frac{d}{dn} f_n^*(\lambda). \tag{E7}$$

Integration by parts gives

$$J_k^*(\lambda) \approx \lambda^{-1} \int_0^{\infty} dn\, n^{k-1} \left[\frac{k}{\ln(n)} - \frac{1}{\ln(n)^2} \right] f_n^*(\lambda). \tag{E8}$$

a.   Infinite mean transition time: $0 < \beta < 1$

We obtain by using (C7) for $\alpha_n = n\ln(n)^{\beta-1}$

$$J_k^*(\lambda) \approx \lambda^{-1} \int_0^{\infty} dn \frac{kn^{k-1}}{\ln(n)} \left[1 - \frac{1}{k\ln(n)} \right] \exp\left[-a_\beta \lambda^\beta n \ln(n)^{\beta-1} \right]. \tag{E9}$$

Rescaling $n \rightarrow n\lambda^\beta \ln(1/\lambda)^{\beta-1}$, we obtain in leading order in the limit $\lambda \rightarrow 0$

$$J_k^*(\lambda) \approx \lambda^{-1-k\beta} \ln(1/\lambda)^{k-1-k\beta} \int_0^{\infty} dn\, kn^{k-1} \exp\left(-a_\beta \beta^\beta n \right). \tag{E10}$$

Inverse Laplace transform gives

$$J_k(t) \propto t^{k\beta} \ln(t)^{k-1-k\beta}. \tag{E11}$$

Thus, we obtain for $\sigma_m^2(t)$ the scaling

$$\sigma_m^2(t) \propto t^{2\beta} \ln(t)^{1-2\beta}. \tag{E12}$$

b. *Finite mean transition time: $\beta > 1$*

We obtain by using (C9) for $\alpha_n = n \ln(n)^{\beta-1}$

$$J_k^*(\lambda) \approx \lambda^{-1} \int_0^\infty dn \frac{k n^{k-1}}{\ln(n)} \left[1 - \frac{1}{k \ln(n)} \right] \exp(-n\lambda\bar{\tau})$$

$$+ \lambda^{-1+\beta} \int_0^\infty dn k n^k \ln(n)^{\beta-2} \left[1 - \frac{1}{k \ln(n)} \right] \exp(-n\lambda\bar{\tau}) b_\beta. \tag{E13}$$

Rescaling of $n \to n\lambda$ gives in leading order

$$J_k^*(\lambda) \approx \frac{\lambda^{-1-k}}{\ln(1/\lambda)} \int_0^\infty dn k n^{k-1} \exp(-n\bar{\tau})$$

$$+ \lambda^{-2-k+\beta} \ln(1/\lambda)^{\beta-2} \int_0^\infty dn k n^k \exp(-n\bar{\tau}) b_\beta. \tag{E14}$$

Inverse Laplace transform gives

$$J_k(t) \approx \frac{(t/\bar{\tau})^k}{\ln(t/\langle\tau\rangle)} + A_2 (t/\bar{\tau})^{1+k-\beta} \ln(t/\bar{\tau})^{\beta-2}. \tag{E15}$$

Thus, we obtain for $\sigma_m^2(t)$ the behavior

$$\sigma_m^2(t) \propto (t/\bar{\tau})^{3-\beta} \ln(t/\bar{\tau})^{\beta-2} \tag{E16}$$

For $\beta > 2$, we obtain

$$J_k(t) \approx \frac{(t/\bar{\tau})^k}{\ln(t/\langle\tau\rangle)} + A_2 (t/\bar{\tau})^{k-1} \tag{E17}$$

and

$$\sigma_m^2(t) \propto t/\bar{\tau}. \tag{E18}$$

43

3. Spatial dimension $q > 2$

For $q > 2$, $R_n \sim 2^{q/2} n^{q/2-1}$. Thus, we obtain for $\sigma_m^2(t)$ from (41)

$$\sigma_m^2(t) = \frac{\ell^4}{2^{q/2}} \left(\overline{n_t^{3-q/2}} - \overline{n_t^{2-q/2} \overline{n_t}} \right) = \ell^4 \left[h_{3-q/2}(t) - h_{2-q/2}(t) h_1(t) \right]. \tag{E19}$$

a. Infinite mean transition time: $0 < \beta < 1$

Using the results for the $h_k(t)$ from the previous section, we obtain

$$\sigma_m^2(t) \propto t^{\beta(3-q/2)}. \tag{E20}$$

b. Finite mean transition time: $\beta > 1$

Along the same lines, we obtain for $1 < \beta < 2$

$$\sigma_m^2(t) \propto t^{4-\beta-q/2}. \tag{E21}$$

For $\beta > 2$, we obtain accordingly

$$\sigma_m^2(t) \propto t^{2-q/2}. \tag{E22}$$

Appendix F: Numerical simulations

The numerical simulations use a time-domain random walk (TDRW) algorithm [44]. The TDRW approach is based on constant length displacement of particles described by the recursive relations:

$$\mathbf{x}_i(n+1) = \mathbf{x}_j(n) + \boldsymbol{\xi}_{ij}, \qquad\qquad t(n+1) = t(n) + \tau_j. \tag{F1}$$

The transition probability w_{ij} to move from voxel j to voxel i for a transition of length $|\boldsymbol{\xi}_{ij}|$ is given by $w_{ij} = 1/(2q)$ and transition time associated to voxel j is $\tau_j = \theta_j \ell^2/(2q\kappa)$. In the case of the annealed disorder model, the random variable θ_j is not linked to the voxel j, but on the n-th step, thus $\theta_j = \theta_n$ and consequently $\tau_j = \tau_n$.

For the numerical simulations, we consider the Pareto distribution

$$p_\theta(\theta) = \beta \theta^{-1-\beta} H(\theta - 1), \tag{F2}$$

44

with $H(\cdot)$ the Heaviside step function. We obtain according to (12) the following distribution of transition times

$$\psi(\tau) = \frac{\beta}{\tau_\kappa} \left(\frac{\tau}{\tau_\kappa} \right)^{-1-\beta} \gamma(1+\beta, \tau/\tau_\kappa), \tag{F3}$$

where $\gamma(a, z)$ is the lower incomplete Gamma function [53]. The numerical calculations of the disorder average particle distribution use $10^3 - 10^4$ particles in each of the $10^3 - 10^4$ disorder realizations. The calculations of the particle distribution in single realizations use $10^5 - 10^7$ particles. The calculations for average mean square displacement use $10^3 - 10^4$ particles in each of the 10^4 disorder realizations. The calculation for the disorder variance of the mean square displacement use 10^5 particles in each of the 10^3 realizations. The calculations for variance of the time average mean square displacement use 10^4 particles in each of the 10^3 disorder realizations.

[1] J. P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).

[2] S. Condamin, O. Bénichou, V. Tejedor, R. Voituriez, and J. Klafter, Nature **450**, 7 (2007).

[3] S. Havlin and D. Ben-Avraham, Adv. Phys. **51**, 187 (2002).

[4] J. Klafter and I. Sokolov, Phys. World **18**, 29 (2005).

[5] I. M. Sokolov, Soft Matter **8**, 9043 (2012).

[6] G. de Marsily, E. Ledoux, A. Barbreau, and J. Margat, Nature **197**, 519 (1977).

[7] B. Noetinger and T. Estebenet, Transp. Porous Media **39**, 315 (2000).

[8] R. Benke and S. Painter, Water Resour. Res. , 1324 (2003).

[9] F. Delay, G. Porel, and P. Sardini, C. R. Geoscience **334**, 967 (2002).

[10] C. Poinssot and H. Geckeis, eds., *Radionuclide behavior in the natural environment* (Woodhead Publishing, 2012).

[11] J. Nelson, Phys. Rev. B **59**, 15374 (1999).

[12] D. Wang, C. He, M. P. Stoykovich, and D. K. Schwartz, ACS Nano **9**, 1656 (2015).

[13] J. Dong, Y. Cheng, Q. Liu, H. Zhang, K. Zhang, and B. Hu, Energy Fuels **31**, 26432651 (2017).

[14] T. W. Patzek, F. Male, and M. Marder, Proc. Natl. Acad. Sci. U. S. A. **110**, 19731 (2013).

[15] K. Falk, B. Coasne, R. Pellenq, F.-J. Ulm, and L. Bocquet, Nat. Commun. **6**, 6949 (2014).

[16] A. Russian, P. Gouze, M. Dentz, and A. Gringarten, Transp. Porous Media, doi:10.1007/s11242-015-0504-y **109**, 109130 (2015).

[17] M. Quintard and S. Whitaker, Adv. Heat Transfer **23**, 369 (1993).

[18] B. Berkowitz and H. Scher, Phys. Rev. Lett. **79**, 4038 (1997).

[19] B. Berkowitz, A. Cortis, M. Dentz, and H. Scher, Rev. Geophys. **44**, RG2003 (2006).

[20] T. Le Borgne, M. Dentz, and J. Carrera, Phys. Rev. E **78**, 041110 (2008).

[21] M. Klepikova, T. Le Borgne, O. Bour, M. Dentz, R. Hochreutner, and N. Lavenant, Water Resour. Res. **52**, 54425457 (2017).

[22] H. Scher and M. Lax, Phys. Rev. B **7**, 4491 (1973).

[23] P. Barthelemy, J. Bertolotti, and D. S. Wiersma, Nature **453**, 495 (2008).

[24] F. Höfling and T. Franosch, Rep. Prog. Phys. **76**, 046602 (2013).

[25] B. M. Regner, D. Vucinic, C. Domnisoru, T. M. Bartol, M. W. Hetzer, D. M. Tartakovsky, and T. J. Sejnowski, Biophys. J. **104**, 1652 (2013).

[26] P. Massignan, C. Manzo, J. A. Torreno-Pina, M. F. García-Parajo, M. Lewenstein, and J. Lapeyre, G. J., Phys. Rev. Lett. **112**, 150603 (2014).

[27] E. Barkai, Y. Garini, and R. Metzler, Phys. Today, No. 8, 65, 29 (2012). **8**, 29 (2012).

[28] C. Manzo, J. A. Torreno-Pina, P. Massignan, G. J. Lapeyre, M. Lewenstein, and M. F. García-Parajo, Phys. Rev. X **5**, 011021 (2015).

[29] G. Seisenberger, M. U. Ried, T. Endreß, H. Buning, M. Hallek, and C. Brauchle, Science **294**, 1929 (2001).

[30] R. Metzler, J. Jae-Hyung, A. Cherstvy, and E. Barkai, Phys. Chem. Chem. Phys. **16**, 24128 (2014).

[31] T. Akimoto, E. Barkai, , and K. Saito, Phys. Rev. Letters **117**, 180602 (2016).

[32] M. Dentz, A. Russian, and P. Gouze, Phys. Rev. E **93**, 010101(R) (2016).

[33] J. P. Bouchaud, J. Phys. I **2**, 1705 (1992).

[34] X. Brokmann, J.-P. Hermier, P. Messin, J.-P. Bouchaud, and M. Dahan, Phys. Rev. Lett. **90**, 120601 (2003).

[35] G. Bel and E. Barkai, Phys. Rev. Lett. **94**, 240602 (2005).

[36] A. Lubelski, I. M. Sokolov, and J. Klafter, Phys. Rev. Lett. **100**, 250602 (2008).

[37] Y. Meroz, I. M. Sokolov, and J. Klafter, Phys. Rev. E **81**, 010101(R) (2010).

[38] T. Miyaguchi and T. Akimoto, Phys. Rev. E **83**, 031926 (2011).

[39] R. Metzler and J. Klafter, Phys. Rep. **339**, 1 (2000).

[40] I. M. Sokolov, A. Blumen, and J. Klafter, Physica A **302**, 268 (2001).

[41] M. Dentz, A. Russian, and P. Gouze, Phys. Rev. E **94**, 019902(E) (2016).

[42] E. M. Bertin and J.-P. Bouchaud, Phys. Rev. E. **67**, 026128 (2003).

[43] H. Risken, *The Fokker-Planck Equation* (Springer Heidelberg New York, 1996).

[44] M. Dentz, P. Gouze, A. Russian, J. Dweik, and F. Delay, Adv. Water Resour. **49**, 13 (2012).

[45] J. F. McCarthy, J. Phys. A: Math. Gen. **26**, 2495 (1993).

[46] B. Noetinger, D. Roubinet, A. Russian, T. Le Borgne, F. Delay, M. Dentz, J.-R. de Dreuzy, and P. Gouze, Transp. Porous Media **115**, 345 (2016).

[47] A. Russian, M. Dentz, and P. Gouze, Water Resour. Res. **52**, 33093323 (2016).

[48] G. H. Vineyard, J. Math. Phys. **4**, 1191 (1963).

[49] C. Aslangul, J.-P. Bouchaud, A. Georges, N. Pottier, and D. Saint-James, J. Stat. Phys. **55**, 461 (1989).

[50] S. Wiseman and E. Domani, Phys. Rev. E **52**, 3469 (1995).

[51] G. Matheron, Economic Geol. **58**, 1246 (1963).

[52] Y. He, S. Burov, R. Metzler, and E. Barkai, Phys. Rev. Lett. **101**, 058101 (2008).

[53] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover Publications, New York, 1972).