SCIENTIFIC REPORTS

OPEN

Cellular automaton models for time-correlated random walks: derivation and analysis

J. M. Nava-Sedeño[1], **H. Hatzikirou**[1,2], **R. Klages**[3] **& A. Deutsch**[1]

Many diffusion processes in nature and society were found to be anomalous, in the sense of being fundamentally different from conventional Brownian motion. An important example is the migration of biological cells, which exhibits non-trivial temporal decay of velocity autocorrelation functions. This means that the corresponding dynamics is characterized by memory effects that slowly decay in time. Motivated by this we construct non-Markovian lattice-gas cellular automata models for moving agents with memory. For this purpose the reorientation probabilities are derived from velocity autocorrelation functions that are given a priori; in that respect our approach is "data-driven". Particular examples we consider are velocity correlations that decay exponentially or as power laws, where the latter functions generate anomalous diffusion. The computational efficiency of cellular automata combined with our analytical results paves the way to explore the relevance of memory and anomalous diffusion for the dynamics of interacting cell populations, like confluent cell monolayers and cell clustering.

Within the past two decades transport processes in many branches of the sciences were observed to be *anomalous*, in the sense that they do not obey the laws of conventional statistical physics and thermodynamics[1–8]. Important cases are diffusion processes where the long-time mean square displacement (MSD) does not grow linearly in time. That is, $\langle r^2 \rangle \propto t^\phi$, where the angular brackets denote an ensemble average, does not increase with $\phi = 1$ as expected for Brownian motion but either *subdiffusively* with $\phi < 1$ or *superdiffusively* with $\phi > 1$[9–11]. After pioneering work on amorphous semiconductors[12], more recently anomalous diffusion has been detected in many other complex systems[3,4,6,8]; here well-known examples of physcial systems are nanopores[13], plasmas[14] and glassy material[15].

Biological systems frequently exhibit anomalous properties as well: Prominent examples are the foraging of organisms[16], epidemic spreading[17] and the diffusion of macromolecules in biological cells[6]. Especially, it was found that many types of cells migrate anomalously: *Hydra* cells[18], mammary gland epithelial cells[19], *MDCKF* cells[20], amoeboid *Dictyostelium* cells[21,22], T cells[23], breast carcinoma cells[24] and stem cells[25] were all experimentally observed to move superdiffusively, typically with non-Gaussian position and/or velocity distribution functions[18–24] accompanied by either exponential or non-exponential position[23], and exponential[21,22,26] or power law[18–20] velocity autocorrelation function (VACF) decay. For T cells it was argued that superdiffusion optimizes their search to kill intruding pathogens[23,27]. While all these results are on single cell migration, currently collective cell migration is moving into the center of interest[28], where cells interact with each other, e.g., by chemical signalling[29]. Interesting phase transitions inside dense tissues of epithelial cell monolayers were reported[30] and partially traced back to particular features of single-cell migration[31]. It was also observed experimentally that superdiffusion appears to foster the formation of clusters of stem cells leading to tissue formation[25]. Other works investigate the role of interacting agents for phase transitions in active matter[32], and collective anomalous dynamics emerging from the interaction of single agents[33,34].

On the theoretical side there are many different ways to model anomalous diffusion in terms of stochastic processes, such as continuous time random walks (CTRW)[1,5], generalized Langevin equations[2], Lévy flights and walks[8], fractional diffusion equations[1,4,5], scaled Brownian motion and heterogeneous diffusion processes[7].

[1]Technische Universität Dresden, Center for Information Services and High Performance Computing, Nöthnitzer Straße 46, 01062, Dresden, Germany. [2]Department of Systems Immunology and Braunschweig Integrated Centre of Systems Biology, Helmholtz Center for Infection Research, Inhoffenstraße 7, 38124, Braunschweig, Germany. [3]School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, United Kingdom. Correspondence and requests for materials should be addressed to J.M.N.-S. (email: nava@mail.zih.tu-dresden.de)
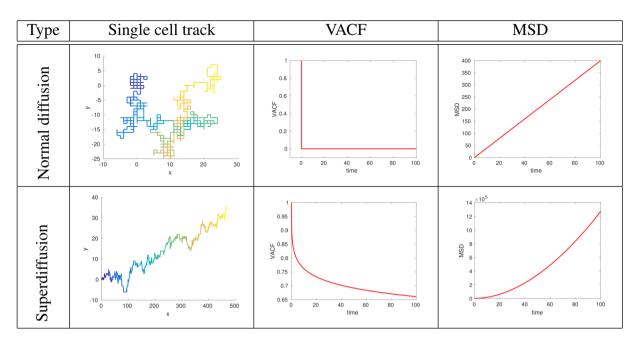
Type	Single cell track	VACF	MSD
Normal diffusion			
Superdiffusion			

Figure 1. Basic types of diffusive movement in two dimensions characterized by two key quantities (VACF and MSD). Shown in the first column are the tracks of a single particle starting at $(x, y) = (0, 0)$ that exhibits either long-time normal diffusion (top row) or superdiffusion (bottom row). The color gradient changes from blue to yellow with elapsed time. The second column displays the particle's velocity autocorrelation function (VACF) Eq. (3), the third column its mean square displacement (MSD) Eq. (4).

A subset of these models, most notably generalized Fokker-Planck equations[18-20] generalized Langevin equations[21,22], and generalized random walks[23,24], has been used to model anomalous movement of single biological cells. However, solving equations for anomalously moving single particles analytically or numerically is typically difficult already[1,2,4,5,7,8]. To our knowledge there exists no systematic attempt to generalize this theory to model interacting many-particle systems; the only exception we are aware of is a line of work in plasma physics[14].

On the other hand, several models have been introduced to study the collective movement of particles and cells[35-37]. Cellular automata (CA) in particular have the advantage of being less computationally demanding than continuous models when performing simulations. A specific type of CA is the so-called lattice-gas cellular automata (LGCA)[38]. In LGCA, each lattice node can contain several particles, which at each time step are rearranged within the lattice node according to the interaction rule, and subsequently moved to a neighboring node. In a biological context particles can be regarded as cells, while the LGCA rules mimic cell migration and interaction. Furthermore, LGCA have proved to be amenable to mathematical analysis[39]. For this reason, LGCA have been introduced as mesoscopic models for single and collective cell migration[40-43]. So far, none of the mentioned models has considered anomalous migration of single cells.

It thus arises the need to design simple fundamental schemes by which the collective properties of interacting agents can be studied whose individual dynamics is anomalous. Using our methods will enable to explore the relevance of microscopic single-particle dynamics for emerging collective phenomena. We thus devise a scheme by which anomalous dynamics of many interacting agents can be simulated efficiently, which is based on capturing the non-trivial decay of VACFs. This approach generates superdiffusion if the correlation decay is of power law-type[18-20]. We emphasize that our data-driven approach can be applied to any moving entity that exhibits dynamics with non-trivial correlation decay, a feature that may be expected to hold more generally for the movement of biological organisms[16].

We use the LGCA modeling framework and construct various time-correlated random walk models. After briefly introducing the LGCA concept, we define an LGCA model for unbiased random walk. Next, motivated by the biophysical mechanism of single cell crawling we construct a *persistent random walk* LGCA model wherein angular (orientation) correlations give rise to temporal correlations. Subsequently, we construct a first LGCA model for *time-correlated random walk* which is data-driven, as the model's reorientation probabilities are derived by assuming that the exact temporal dependence of the VACF is known a priori. Finally, we develop a *generalized time-correlated random walk* LGCA model for cell movement at short and medium time regimes by curing a deficiency of our first time-correlated random walk model for short times. Figure 1 shows single cell tracks with the corresponding VACFs and MSDs for our main two classes of LGCA models we are dealing with, which are Markovian and non-Markovian random walks, exemplified by showing their basic features.

Lattice-gas cellular automata

Cellular automata are mathematical models where the states of discrete lattice nodes are updated at discrete time steps. If the states of the lattice sites are Boolean, such states can be interpreted as presence/absence of a particle at a particular node. The lattice-gas cellular automaton is a specific CA type, which has two important
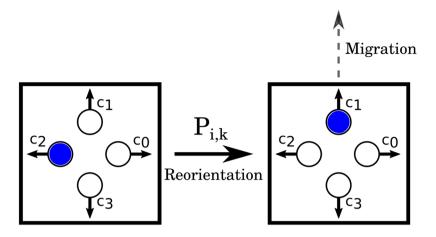
Figure 2. LGCA dynamics. At each time step a particle is assigned an orientation $\vec{c_i}$ with a probability $P_{i,k}$. Subsequently, the particle is translocated to the nearest neighbor in the direction of its orientation.

characteristics: first, particle reorientation and migration are separated into a probabilistic and a deterministic step, respectively. Secondly, to each node, b velocity channels are associated which can be occupied by at most one particle (exclusion principle). The set of velocity channels is given by $\vec{c_j} = \left(\cos\frac{2\pi j}{b}, \; \sin\frac{2\pi j}{b} \right), j \in \{0, 1, \ldots, b-1\}$ (see Fig. 2). Particles move in discrete time steps of duration τ to neighboring nodes located a distance ε away in the lattice. At each time step particles adopt the orientation $\vec{c_i}$ with a probability $P_{i,k}$, called the reorientation probability, where $t = k\tau, k \in \mathbb{N}$ is the elapsed simulation time. Subsequently, cells will be deterministically translocated to the nearest neighbor located in the direction of $\vec{c_i}$; see Fig. 2 for a sketch of LGCA dynamics.

If we were to have N different particles in a single node, then the probability of the particles adopting the orientations $\vec{c}_{n^1}, \vec{c}_{n^2}, \ldots, \vec{c}_{n^N}$ would be given by

$$P_{n^1,\ldots,n^N,k} = \prod_{\ell=1}^{N} P_{n^\ell,k}.$$

(1)

Classical random walk

Here we define an LGCA model for unbiased random walk[44]. In this model, at all timesteps, all orientations are chosen with equal probability[45,46]. This means that the reorientation probability is given by

$$P_{i,k} = \frac{1}{b}.$$

(2)

In order to characterize the movement of a particle in this model we calculate time-dependent expressions for the VACF and the MSD, which measure the persistence, in terms of memory decay in time, and the spatial exploratory power of a moving particle, respectively. In LGCA models space, time and particle velocities are discrete so that the VACF is given by[47]

$$g(k) = \left\langle \vec{c}_{i_0} \cdot \vec{c}_{i_k} \right\rangle = \sum_{i_k=1}^{b} P_{i_k,k} \left[\vec{c}_{i_0} \cdot \vec{c}_{i_k} \right],$$

(3)

where $\vec{c}_{i_k}$ is the orientation of the particle at the k-th time step. The MSD is calculated by

$$\langle r_k^2 \rangle = \sum_{r_k} r_k^2 P_{r_k,k},$$

(4)

where r_k is the norm of the particle displacement at time step k defined as $\vec{r_k} = \vec{x_k} - \vec{x_0}$, where $\vec{x_k}$ is the position of the particle at time step k. The probability $P_{r_k,k}$ can be calculated from Eq. (2) by noticing that $\vec{r_k} = \varepsilon \sum_k \vec{c}_{i_k}$. In this simple random walk model Eq. (3) reduces to a sum of cosines over homogeneously distributed angles. Hence the VACF is given by $g(k) = \delta_{0,k}$, where $\delta_{i,j}$ is the Kronecker delta. In the limit $\tau \to 0$ the VACF tends to

$$g(t) = \delta(t),$$

(5)

where $\delta(t)$ is the Dirac delta function. Equation (5) means that the movement of the particle is uncorrelated as soon as it starts moving, i.e. the orientation of the particle at any time step k is completely independent from its previous orientation, which is the Markov property. On the other hand, simple combinatorics can be used to calculate the particle's MSD yielding $\langle r_k^2 \rangle = k\varepsilon^2$. We can rewrite this expression by using the general definition of the diffusion coefficient
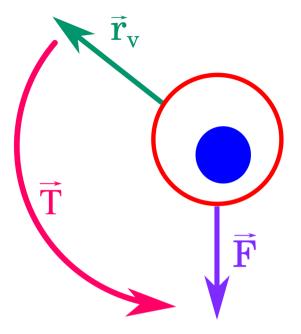
Figure 3. Reorientation of a biological cell moving persistently. A cell which moves in a direction $\vec{r_v}$ feels an intracellular force $\vec{F}$ and reorients towards the direction of the force due to the torque $\vec{T}$.

$$D = \lim_{t \to \infty} \frac{\langle r_t^2 \rangle}{2dt},\tag{6}$$

where d is the dimension of space. For a memoryless random walk this equation boils down to $D_{rw} = \frac{\varepsilon^2}{2d\tau}$ with an MSD of $\langle r_k^2 \rangle = 2dD_{rw}k\tau$. Given that τ is the time step length and that $t = k\tau$ is the elapsed time, the MSD is

$$\langle r_t^2 \rangle = 2dD_{rw}t.\tag{7}$$

Equation (7) shows that the classical random walk model trivially yields a normal diffusion process, where the MSD increases linearly in time[48].

Persistent random walk

The assumption of the classical random walk in Eq. (2) that all the directions of movement are equally probable is generally not true. We will now construct an LGCA model motivated by a simple biophysical model for a single, persistently moving cell.

Rule derivation. Biological cells move by exerting forces to propel themselves. In the case of eukaryotic cells such as fibroblasts, movement is achieved by crawling over the substrate. Crawling is performed by polymerization of the actin cytoskeleton at the leading edge propelling the cell in this direction. We can identify the direction of the actin concentration gradient with the direction of movement of a cell $\vec{r_v}$. Furthermore, intracellular forces due to actin activity $\vec{F}$ point towards the direction of new actin polymerization. The cell then rotates/reorients/repolarizes due to the torque defined by $\vec{T} = \vec{r_v} \times \vec{F}$, whose norm is given by

$$T = r_v F \sin(\theta_F - \theta_r),\tag{8}$$

where $T = \|\vec{T}\|$, $r_v = \|\vec{r_v}\|$, $F = \|\vec{F}\|$, and θ_r and θ_F are the angular components of $\vec{r_v}$ and $\vec{F}$, respectively; see Fig. 3. In the overdamped regime, characteristic of the cellular environment, the intracellular force $\vec{F}$ will not cause the cell to rotate indefinitely but rather will cause the cell to rotate until $\vec{r_v}$ and $\vec{F}$ are parallel. Taking this into account, it is possible to rewrite Eq. (8) as $T = r_v F \sin(\theta_{rt} - \theta_r)$, where θ_{rt} is the direction of motion of the cell after it has finished its reorientation.

The torque is given in terms of an energy of rotation $U(\theta_r, \theta_{rt})$ as $\vec{T} = -\frac{\partial U(\theta_r, \theta_{rt})}{\partial \theta_r}$. Using this equation, the energy of rotation is then given by

$$U(\theta_r, \theta_{rt}) = -v\cos(\theta_{rt} - \theta_r),\tag{9}$$

where $v = r_v F$ is the amplitude of the torque generated inside the cell. Having defined the energy of rotation, Eq. (9), we can describe the cell's reorientation by a Langevin equation[49] as $\frac{\partial \theta_r}{\partial t} = -\gamma \frac{U(\theta_r, \theta_{rt})}{\partial \theta_r} + \xi(t)$ with relaxation constant γ and a zero-mean, delta correlated noise term $\xi(t)$ such that $\langle \xi(t)\xi(t') \rangle = 2D_\theta \delta(t - t')$, where D_θ is the rotational diffusion coefficient. Based on this we can immediately derive the LGCA reorientation probabilities[50].

These probabilities for a single cell then read

$$P_{i_k,k} = \frac{1}{Z}\exp\left[\beta\left(\overrightarrow{c_{i_{k-1}}} \cdot \overrightarrow{c_{i_k}}\right)\right],$$

(10)

where $\overrightarrow{c_{i_{k-1}}}$ is the orientation of the cell at the previous time step, Z is the normalization constant (also known as the partition function) and $\beta = \frac{\tilde{\gamma}}{D_\theta}$ is the sensitivity, where $\tilde{\gamma} = \upsilon\gamma$ is the effective relaxation constant.

Model analysis and results. Using Eq. (10) we can calculate the VACF and MSD for this model. By using the properties of the partition function Z in Eq. (10) the VACF at every time step k is (see Sec. B in the Supplementary Information)

$$g(k) = \exp(\alpha k),$$

(11)

where the exponent α depends heavily on the lattice dimension and geometry. In particular, in a 2D square lattice we have $\alpha = \ln\left[\tanh\left(\frac{\beta}{2}\right)\right]$. In all geometries the exponent is $\alpha < 0$ over its domain $\beta > 0$ (see again Sec. B in the Supplementary Information).

Equation (11) can be generalized to continuous time and space by employing the relations between the time and space scalings, namely the diffusion coefficient in the random walk limit D_{rw} and the instantaneous cell speed $v = \frac{\varepsilon}{\tau}$, where ε is the lattice spacing and τ the time step length. Taking the limit $\tau \to 0$ yields the VACF in continuous time and space

$$g(t) = \exp\left(\frac{\alpha v^2}{2dD_{rw}}t\right).$$

(12)

When time is discrete the MSD is given by[51,52] (see Sec. A in the Supplementary Information) $\langle r^2 \rangle = 2dD_{rw}k\tau + \left\langle \sum_{i=1}^{k}\sum_{j=1}^{k}v^2\tau^2\cos(\theta_i - \theta_j)(1 - \delta_{ij})\right\rangle$. Calculating the expected value on the right hand side we obtain (see again Sec. A in the Supplementary Information)

$$\langle r_k^2 \rangle = 2dD_{rw}k\tau + 2v^2\sum_{i=1}^{k}(k\tau - i\tau)g(i)\tau.$$

(13)

Using Eq. (11) and taking the limit of small time step length we get

$$\langle r_t^2 \rangle = 2dD_{rw}t + 2v^2\int_0^t (t - \tau)e^{\frac{v^2\alpha}{2dD_{rw}}\tau}\mathrm{d}\tau,$$

(14)

which can be easily integrated to obtain the MSD for continuous time

$$\langle r_t^2 \rangle = 2dD_{rw}t\left(1 - \frac{2}{\alpha}\right) + \left(\frac{2\sqrt{2}\,dD_{rw}}{v\alpha}\right)^2\left[\exp\left(\frac{v^2\alpha t}{2dD_{rw}}\right) - 1\right].$$

(15)

Equation (14) agrees with the formal solution of the MSD for an overdamped Langevin equation with colored noise[53]. Correspondingly, Eq. (10) coincides with a Langevin process where the noise is not white but colored whose correlation is given by Eq. (12). Using Eq. (15) we find that $\langle r^2 \rangle \propto t$ when $t \to \infty$ rather quickly.

Comparing Eqs (5) to (12) in Fig. 4 we see that in this second model the velocities are no longer delta correlated but now decay exponentially in time. On the other hand, for long times both Eqs (7) and (15) yield normal diffusion. However, for short times cells performing persistent random walks move superdiffusively, contrary to cells performing classical random walks; see again Fig. 4. In Fig. 5i, Eqs (12) and (15) are compared with results from LGCA simulations, where we see that the theory adequately predicts the observed simulation results. Details on the computational implementation are found in Sec. H of the Supplementary Information.

Time-correlated random walk

Due to the exponential decay of correlations the previous model did not show superdiffusion at long time scales. It turns out that finding a homogeneous, isotropic Markovian model that shows superdiffusion and power-law decaying correlations is not possible (for a proof see Sec. C of the Supplementary Information).

Theorem 1 The velocity autocorrelation function of a particle whose orientations are given by a homogeneous, symmetric Markov chain is either delta-correlated, i.e. $g_k = \delta_{0,k}$, where δ is the Kronecker delta; alternating, i.e. $g_k = (-1)^k a^k, a \in \mathbb{R}^+$; or exponentially decaying, i.e. $g_k = e^{\alpha k}, \alpha \leq 0$.

To reproduce superdiffusion and power law decaying autocorrelations, we will construct a non-homogeneous model by assuming that the time dependency of the VACF is a known power law.

Rule derivation. We now assume that the VACF $g(t)$ is known. In particular, if the movement is power law-correlated the VACF has the form[53]

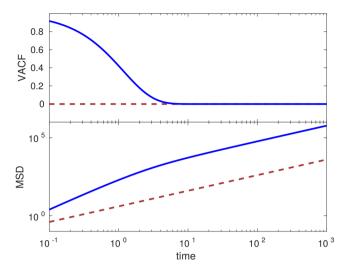$$g(t) = C_0\left(\frac{\Delta}{t}\right)^\phi, \ t \geq \Delta,$$

(16)

Figure 4. Comparison between random walk and persistent random walk models. Shown are theoretical results for VACF (top) and MSD (bottom) in the random walk (dashed maroon line) and persistent (solid blue line) models. The parameter values are $D_{rw} = 1$, $v = 16$, and $\beta = 5$.

where $\Delta > 0$ and $0 < \phi < \infty$ and assume that $\Delta \ll 1$, to disregard the movement at short times, where Eq. (16) diverges. The rate of decay of the VACF is proportional to the exponent ϕ. The crossover time Δ specifies the time at which $g(t) = C_0$. The walk is positively correlated if $C_0 > 0$, and anti-correlated if $C_0 < 0$. Because the process is non-homogeneous, $P_{i_k,k}$ in Eq. (3) explicitly depends on the velocity channels $\overrightarrow{c_i}$ and the current time step k. Combining Eqs (3) and (16) we obtain the following relation[47]:

$$\sum_{i=1}^{b} P_{i_k,k}\left[\overrightarrow{c_{i_0}} \cdot \overrightarrow{c_{i_k}}\right] = g(k). \tag{17}$$

It is possible to derive the reorientation probabilities $P_{i_k,k}$ by expanding Eq. (17) for every time step. Additionally, in order to reduce the number of equations, we make the following assumptions:

- The reorientation probabilities are independent, that is, the probability of following a certain trajectory is $P_{i_1,i_2,\cdots,i_k} = \prod_{j=1}^{k} P_{i_j,j}$.
- There is symmetry around the initial orientation, i.e. if $\overrightarrow{c_{i_k,k}} \cdot \overrightarrow{c_{i_0}} = \overrightarrow{c_{j_k,k}} \cdot \overrightarrow{c_{i_0}}$ then $P_{i_k,k} = P_{j_k,k}$, $i_k \neq j_k$.

Using these assumptions we can derive the general expression for the reorientation probabilities determining a certain VACF (see Sec. D in the Supplementary Information)

$$P_{i_k,k} = \frac{1 + d\left[\overrightarrow{c_{i_0}} \cdot \overrightarrow{c_{i_k}}\right]g(k)}{b}, \tag{18}$$

where d is the dimension of space and b is the number of lattice directions given by the lattice geometry. If the VACF follows a power law, then Eq. (18) is always valid if the crossover time Δ is smaller than the time step length, as the divergence of Eq. (16) is avoided. If $\Delta \gg 0$, we assume that the movement at short times is completely correlated, i.e. the VACF is given by
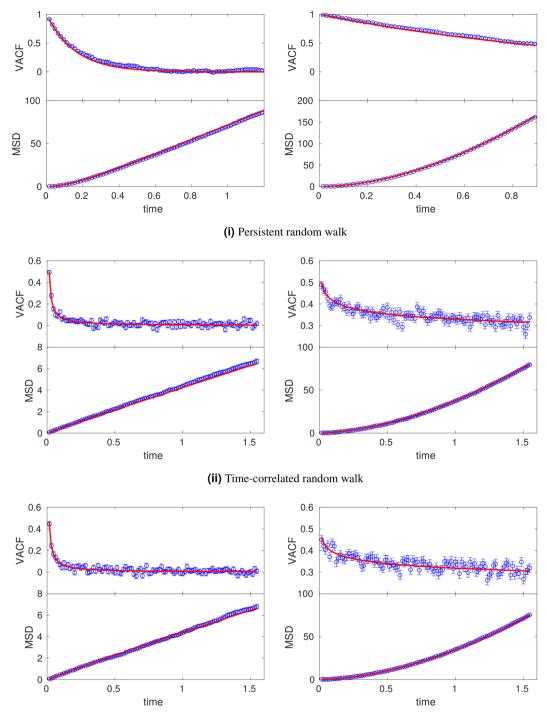
$$g(t) = \begin{cases} 1 & t \leq t^\star \\ C_0\left(\dfrac{\Delta}{t}\right)^{\phi} & t > t^\star \end{cases}, \tag{19}$$

where t^* is such that $C_0\left(\frac{\Delta}{t^*}\right)^{\phi} = 1$. We can then define a piecewise reorientation probability

$$P_{i_k,k} = \begin{cases} \delta(\overrightarrow{c_{i_k}} - \overrightarrow{c_{i_0}}) & k \leq \omega \\ \dfrac{1 + d[\overrightarrow{c_{i_0}} \cdot \overrightarrow{c_{i_k}}]g(k)}{b} & k > \omega \end{cases}, \tag{20}$$

where $\omega\tau = t^*$ is the duration of ballistic motion.

Model analysis and results. For this model the VACF is obviously known as the reorientation probabilities were calculated specifically to reproduce it. We now calculate the MSD for this model. The reorientation

(i) Persistent random walk

(ii) Time-correlated random walk

(iii) Generalized time-correlated random walk

Figure 5. Comparison between persistent, time-correlated and generalized time-correlated random walk models. Shown are simulation results (circles, mean $\pm$ standard error of the mean) and theoretical prediction (solid line) for VACF (top row) and MSD (bottom row). Parameters are $v = 16$ and $D_{rw} = 1$ in all cases. i) Sensitivity values: $\beta = 3$ (left), $\beta = 5$ (right). ii) and iii) $C_0 = 0.5$, $\Delta = 0.016$, and exponents: $\phi = 1$ (left), $\phi = 0.1$ (right).

probabilities given by Eq. (18) only depend on the initial cell orientation $\overrightarrow{c_{i_0}}$ and on the time step k, so they are independent of other orientations at other times. Because of this independence of orientations the MSD is given by (see Sect. A in the Supplementary Information)

$$\langle r_k^2 \rangle = 2dD_{rw}[k\tau - 2\sum_{i=1}^{k}g^2(i)\tau] + 2v^2\sum_{i=1}^{k}\sum_{j=i}^{k}g(i)g(j)\tau^2$$

(21)

yielding a Taylor-Green-Kubo formula[47,54]. In the limit of small time step lengths $\tau \to 0$ the MSD is given by

$$\langle r_t^2 \rangle = 2dD_{rw}\left[t - 2\int_{\Delta}^{t}g^2(\tau)d\tau\right] + 2v^2\int_{\Delta}^{t}\int_{\tau}^{t}g(\tau)g(k)dkd\tau.$$

(22)

Equation (22) shows that there is a correction to the random walk diffusion coefficient as well as a new term depending on the particle speed. If $g(t)$ is given by Eqs (16) and (22) can be integrated yielding

$$\langle r_t^2 \rangle = 2dD_{rw}\left\{t + \frac{2C_0^2\Delta}{1-2\phi}\left[1 - \left(\frac{\Delta}{t}\right)^{2\phi-1}\right]\right\}$$
$$+ 2\left(\frac{vC_0\Delta}{\phi-1}\right)^2\left[\frac{1}{2}\left(\frac{\Delta}{t}\right)^{2\phi-2} - \left(\frac{\Delta}{t}\right)^{\phi-1} + \frac{1}{2}\right], \quad \phi \neq 1, \phi \neq \frac{1}{2}$$

(23a)

$$\langle r_t^2 \rangle = 2dD_{rw}\left[t + 2C_0^2\Delta\ln\left(\frac{\Delta}{t}\right)\right]$$
$$+ 2v^2C_0^2[2t\Delta + 2\Delta^2 - 4\Delta(\Delta t)^{\frac{1}{2}}], \quad \phi = \frac{1}{2}$$

(23b)

$$\langle r_t^2 \rangle = 2dD_{rw}\left[t + 2C_0^2\left(\frac{\Delta^2}{t} - \Delta\right)\right]$$
$$+ \left[vC_0\Delta\ln\left(\frac{\Delta}{t}\right)\right]^2, \quad \phi = 1.$$

(23c)

These expressions for the MSD are valid when $\Delta \to 0$. For long crossover times when reorientation probabilities are given by Eq. (20), the MSD is (see Sec. E in the Supplementary Information):

$$\langle r_t^2 \rangle = \begin{cases} (vt)^2 & t \leq t^\star, \\ 2dD_{rw}\left[(t - t^\star) - 2\int_{t^\star}^{t}g^2(\tau)d\tau\right] & \\ +v^2\left[2\int_{t^\star}^{t}\int_{\tau}^{t}g(\tau)g(k)dkd\tau + t^{\star 2} + 2t^\star\int_{t^\star}^{t}g(\tau)d\tau\right] & t > t^\star. \end{cases}$$

(24)

Figure 5ii shows a comparison of Eqs (16) and (22) with LGCA simulations. Details on the computational implementation are found in Sec. H of the Supplementary Information. From Eq. (23) we see that in general $\langle r_t^2 \rangle \sim t \pm t^{1-2\phi} + t^{2(1-\phi)} - t^{1-\phi}$, which defines three regimes:

1. $\phi < \frac{1}{2}$: superdiffusive regime, arising from the term $t^{2(1-\phi)}$
2. $\frac{1}{2} < \phi < 1$: subdiffusive regime, as the term $t^{2(1-\phi)}$ dominates at short times only
3. $\phi > 1$: normal diffusive regime, as the linear term is the dominating term

Figure 6 shows that, while the VACF decays rapidly in the persistent model, in the time-correlated model the VACF decays much more slowly. Additionally, the movement is superdiffusive in both models are short times, however this behavior is long-lasting in the time-correlated model.

Generalized time-correlated random walk

The reorientation probabilities derived for the time-correlated random walk are only valid for certain time ranges, due to the divergence of the VACF when $t \to 0$. We will now derive a generalized model which is valid on both short and long time scales. For this purpose we use what is called the maximum caliber formalism[55], which we introduce briefly.

Rule derivation. The maximum caliber formalism has been proven successful to derive models for dynamic systems from data. The procedure consists in maximizing the entropy over a path of system evolutions, with the constraint of reproducing certain observables. The procedure of entropy maximization does not only ensure that the resulting model contains as few assumptions as possible, but is also considered the only method of correctly obtaining unknown probability distributions from known data[55]. The procedure is as follows:

Let the path entropy, or caliber, be defined as $\mathcal{C} = -\sum_{\Gamma}P_{\Gamma}\ln P_{\Gamma}$, where Γ is a possible path followed by the system during its evolution. The probability of following such a path is given by P_{Γ}. In the case of a single random walker, the path is the entire history of particle velocities $\Gamma = i_0i_1i_2\ldots i_k$ up to the last time step k. Furthermore, we
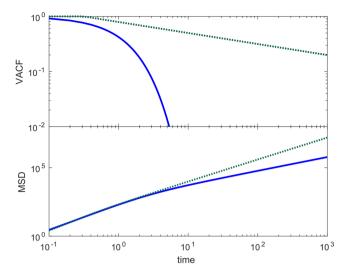
Figure 6. Comparison between persistent and time-correlated random walk models. Shown are theoretical results for VACF (top) and MSD (bottom) in the persistent (solid blue line) and time-correlated (dotted green line) models. The parameter values are $D_{rw} = 1$, $v = 16$, $\beta = 5$, $C_0 = 0.5$, $\Delta = 10$, and $\phi = 0.2$.

constrain the unknown probabilities by a normalization constant and the observed VACF (in this case Eq. (17)). Then the problem translates into optimizing the functional[56] $\widetilde{C}[P_\Gamma] = -\sum_\Gamma P_\Gamma \ln P_\Gamma + \sum_{j=1}^k \beta(j)[\sum_\Gamma P_\Gamma (\overrightarrow{c_{i_0}} \cdot \overrightarrow{c_{i_j}}) - g(j)] + \lambda(\sum_\Gamma P_\Gamma - 1)$, where $\beta(j)$ and λ are Lagrange multipliers to be determined. The Lagrange multiplier $\beta(j)$ is given by $\beta(j) = dg(j)$ (see Sec. F in the Supplementary Information), and λ determines the normalization constant.

Using the expression for $\beta(j)$ we obtain the reorientation probability

$$P_{i_k,k} = \frac{1}{z}\exp\left[dg(k)\left(\overrightarrow{c_{i_0}} \cdot \overrightarrow{c_{i_k}}\right)\right],$$

(25)

where z is the normalization constant for the reorientation probability.

If one required not only that the VACF was observed but also that the autocorrelation function would decay similarly independently of the start and end points, i.e. $\langle \overrightarrow{c_{i_m}} \cdot \overrightarrow{c_{i_n}} \rangle = \langle \overrightarrow{c_{i_j}} \cdot \overrightarrow{c_{i_l}} \rangle$, if $n - m = l - j$, the problem would be similar with the exception that we would now have $\frac{k(k+1)}{2}$ constraints if the trajectory consists of k time steps. Analogously, the probabilites would then be given by

$$P_\Gamma - = \frac{1}{Z}\exp\left[\sum_{j=1}^k \sum_{m=j-1}^{k-1} dg(j-m)\left(\overrightarrow{c_{i_m}} \cdot \overrightarrow{c_{i_j}}\right)\right].$$

(26)

Model analysis and results. The VACF can be easily calculated by using the properties of the partition function Z. Given a distribution $P(x) = \frac{1}{Z}\exp[-\beta H(x)]$, the expected value of the function $H(x)$ is

$$\langle H \rangle = -\frac{\partial}{\partial \beta}\ln Z.$$

(27)

Combining Eqs (3), (25) and (27) the VACF is given by $g(k) = \frac{\partial}{\partial \beta(k)}\ln z$, where $\beta(k) = \beta_0(k)$ in the case of of probabilities given by Eq. (26). In the case of a 2D square lattice the partition function can be easily calculated as $z = 2\{1 + \cosh[\beta(k)]\}$. Taking the logarithm and differentiating we obtain $g(k) = \frac{\sinh[\beta(k)]}{1 + \cosh[\beta(k)]} = \tanh\left[\frac{\beta(k)}{2}\right]$. If we consider power law correlations and take the limit $\tau \to 0$ we obtain

$$g(t) = \tanh\left[C_0\left(\frac{\Delta}{t}\right)^\phi\right].$$

(28)

Eqs (16) and (28), corresponding to the time-correlated random walk and generalized time-correlated random walk, respectively, are visually compared in Fig. 7. A Taylor expansion of $g(t)$ around $\left(\frac{\Delta}{t}\right)^\phi = 0$ (i.e. valid for $\Delta \to 0$ or $k \to \infty$) shows that up to second order terms $g(t) \approx C_0\left(\frac{\Delta}{t}\right)^\phi$, as expected. For $\left(\frac{\Delta}{t}\right)^\phi \gg 0$ the VACF decays as a power law as well (see Sec. G in the Supplementary Information), $g(t) \approx G(t) = \tanh(C_0)\left(\frac{\Delta}{t}\right)^{2C_0\phi \operatorname{csch}(2C_0)}$, and the difference between both functions behaves as $g(t) - G(t) \propto \left(\frac{\phi}{\Delta}\right)^2$. When the probabilites are given by Eq. (25)
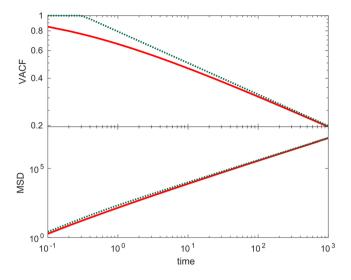
Figure 7. Comparison between generalized and time-correlated random walk models. Shown are theoretical results for VACF (top) and MSD (bottom) in the time-correlated (dotted green line) and generalized time-correlated (solid red line) models. The parameter values are $D_{rw} = 1$, $v = 16$, $C_0 = 0.5$, $\Delta = 10$, $\phi = 0.2$.

particle orientations are independent of one another, as it was the case in the time-correlated model, and the MSD is given by Eq. (21). Therefore, in the limit $\tau \to 0$ the MSD is

$$
\begin{aligned}
\langle r_t^2 \rangle =\ & 2dD_{rw}\left\{ t - 2\int_0^t \tanh^2\left[\frac{\beta(\tau)}{2}\right]d\tau \right\} \\
& + 2v^2 \int_0^t \int_\tau^t \tanh\left[\frac{\beta(\tau)}{2}\right]\tanh\left[\frac{\beta(k)}{2}\right]dkd\tau.
\end{aligned}
\tag{29}
$$

If we consider power law correlations and expand in Taylor series around $\Delta^\phi = 0$ up to second order, we recover Eq. (22) when $g(t)$ is given by Eq. (16). Therefore the MSD in this model follows similar regimes as those of the time-correlated model Eq. (16). Equations (24) and (29) are visually compared in Fig. 7, and see Fig. 5iii for a comparison of Eqs (28) and (29) with LGCA simulations. Details on the computational implementation are found in Sec. H of the Supplementary Information.

When the probabilities are given by Eq. (26) instead the MSD is given by Eq. (13), as all orientation pairs i and j are correlated, where the correlation is given by $g(i - j)$ thus not depending specifically on the values of i and j but on their difference only. Therefore, for a 2D square lattice and a power-law decaying VACF in the limit $\tau \to 0$ the MSD is given by

$$
\langle r_t^2 \rangle = 2dD_{rw}t + 2v^2 \int_0^t (t - \tau)\tanh\left[C_0\left(\frac{\Delta}{\tau}\right)^\phi \right]d\tau.
\tag{30}
$$

Again we see that this expression agrees with the formal solution of the MSD for an overdamped Langevin equation with colored noise[53]. In this case, however, the noise correlation is not decaying exponentially. If we expand the hyperbolic tangent on the right hand side around $\left(\frac{\Delta}{t}\right)^\phi = 0$ (i.e. $\Delta \to 0$ or $t \to \infty$) up to the second term and integrate we obtain the MSD as

$$
\langle r_t^2 \rangle = 2dD_{rw}t + \frac{2C_0}{1 - \phi}(v\Delta)^2 \left\{ \frac{1}{2 - \phi}\left[\left(\frac{\Delta}{t}\right)^{\phi - 2} - 1\right] + 1 - \frac{t}{\Delta} \right\}, \quad \phi \neq 1, \phi \neq 2
\tag{31a}
$$

$$
\langle r_t^2 \rangle = 2dD_{rw}t + 2C_0(v\Delta)^2 \left\{ \frac{t}{\Delta}\left[\ln\left(\frac{t}{\Delta}\right) - 1\right] + 1 \right\}, \quad \phi = 1
\tag{31b}
$$

$$
\langle r_t^2 \rangle = 2dD_{rw}t + 2C_0(v\Delta)^2\left[\frac{t}{\Delta} - \ln\left(\frac{t}{\Delta}\right) - 1 \right], \quad \phi = 2.
\tag{31c}
$$

We conclude from Eq. (31) that

$$
\langle r_t^2 \rangle \sim t \pm t^{2 - \phi}.
$$

When $\phi < 1$ we have $2 - \phi > 1$ and the process is superdiffusive. When $1 < \phi < 2$ we have $0 < 2 - \phi < 1$, at short times the term $t^{2-\phi}$ dominates, and the process is subdiffusive while at long times the linear term dominates yielding normal diffusion. Finally, when $\phi > 2$ we have $2 - \phi < 0$ and the process is completely normal diffusive.

Summary and Discussion

The goal of our study was to design a simple model for a single particle moving with memory in absence of any environmental cue. We chose a cellular automaton, specifically, an LGCA because it is a flexible and computationally efficient framework and has the potential to analyze collective behavior in populations of moving particles or cells. After having introduced an LGCA model for unbiased random walk, we have derived three different novel time-correlated LGCA models for single particle migration.

The subsequent persistent random walk LGCA model was derived from a biophysicallymotivated Langevin equation for particle reorientation in an overdamping environment. We showed that in this model the VACF decays exponentially. Furthermore, we proved that a particle in this model moves superdiffusively only at short times while it diffuses normally in the long time limit. This behavior as well as the expression we found for the MSD agree with that found by Othmer[57] using the telegrapher's equation and also to the one by Chechkin et al.[53] for exponential noise.

The time-correlated random walk model was derived by assuming that the specific form of the VACF is known. We also assumed that reorientation probabilities were completely independent, and that particles have no preference in turning left or right. We considered the specific case of a power law-decaying VACF and showed that the MSD exhibits two transitions when $g(t) \propto t^{-\frac{1}{2}}$ and $g(t) \propto t^{-1}$. For small exponents the particle moves superdiffusively on every time scale. At intermediate exponents there are nonlinear contributions dominating at short times. For large exponents, all nonlinear contributions vanish in the long time limit resulting in normal diffusion.

Finally, we derived a generalized LGCA model by maximizing the diffusing particle's path entropy while retaining the constraint of reproducing a certain VACF. In this model the reorientation probability Eq. (25) is similar to the reorientation probability of the persistent random walk Eq. (10), with some differences: in Eq. (25) the particle's orientation is compared to its initial orientation while in Eq. (10) the particle's orientation is compared to the particle's orientation at the previous time step. Furthermore, in Eq. (10) we have a constant parameter β while in Eq. (27) this parameter decays with time, i.e. $\beta \propto g(t)$. We recall that Eq. (10) results from considering a Langevin equation for the particle's reorientation. The parameter β depends on the magnitude of the reorienting force, the relaxation constant γ (related to friction) and the angular diffusion constant D_θ. A time-dependent parameter $\beta(k)$ would be obtained when considering a generalized Langevin equation resulting from either a time-dependent reorientation force or friction if these values changed much more slowly than the time needed for the particle to be displaced. Taking this into account, Eq. (10) describes the movement of a particle when reorientations can be performed almost instantaneously compared to the time required for the particle to move in space. Eq. (25) on the other hand describes movement when the particle keeps moving but needs considerably more time to change its initial orientation. When in addition the VACF is required to be invariant under time translations we showed that the corresponding MSD time regimes match to those found by Chechkin et al.[53] for power law-correlated noise.

We have verified our analytical results of all constructed models by comparing them to LGCA computer simulations. In order to derive the analytical VACF and MSD expressions, we have considered the limit $\tau \to 0$. In this limit, the macroscopic time t remains small even after several time steps. It stands to reason that, for $t \to \infty$, $\tau \ll t$, the difference between our analytical expressions and simulations becomes negligible.

In their present form our LGCA models assume (i) the particle has constant instantaneous speed v; (ii) the particle moves to a neighboring site at every time step; and (iii) the particle moves on a regular lattice, which impacts the specific expression of the VACF. All these models could be extended by considering different instantaneous speeds, as well as waiting times between subsequent displacements, by using multispeed LGCA and adding rest (zero velocity) channels, respectively. Effects of the lattice regularity on the single particle movement can be compensated by choosing the sensitivity β appropriately in the persistent random walk model as well as the crossover time Δ in the generalized time-correlated model. In the time-correlated model the VACF does not depend on the lattice geometry.

Our new models could also be extended to account for external forces acting on the particle, independent from its intrinsic anomalous movement. For extending the models, we can consider that particle reorientations are caused by internal correlations of individual cells and by particle interactions. The probability $P^{corr}_{n^1,\dots,n^N,k}$ of having N particles in a node with a certain orientation due to internal particle orientations has already been introduced in Eq. (1), while individual particle reorientation probabilities $P_{n^\ell,k}$ would be given according to one of the models introduced in this work. On the other hand, the probability $P^{int}_{n^1,\dots,n^N,k}$ of having N particles in a node with a certain orientation due to particle interactions would be a function of other particles' positions and orientations (see[50], for examples of such probabilities). If we assume that both probabilities are independent, then the reorientation probability for all particles would simply be $P^{tot}_{n^1,\dots,n^N,k} = P^{corr}_{n^1,\dots,n^N,k} \cdot P^{int}_{n^1,\dots,n^N,k}$. Such an extension could be useful for studying physical systems such as plasma gases[14]. Furthermore, it would be interesting to construct LGCA models generating Lévy walks exhibiting non-Gaussian probability density functions[5,8]. More importantly, and tracing back to the original motivation of our work, due to the computational efficiency of LGCAs, our schemes could be applied to model large groups of interacting cells to study the impact of persistence and time correlations in single-cell dynamics on collective phenomena. Highly promising examples are coordination and swarming in bacteria[34], pluripotent cells during early development[25], and the emergence of phase transitions in collective cell migration[28–31]. Moreover, the non-cellular microenvironment is crucial for cell migration phenomena. Recently, the impact of complex environments on cell dissemination has been studied with a cellular automaton model[58]. It would be interesting to extend the models introduced here to analyze the impact of anomalous dynamics and complex microenvironments on cell dissemination and cancer invasion.

The LGCA modeling framework followed in this work is characterized by simplifying the concept of a moving particle to movement in discrete time steps between discrete nodes on a regular lattice, possessing only a finite, discrete set of velocities. On one hand, this "discrete approach" is decidedly more simplified and abstract than "continuous approaches" such as continuous time random walks or fractional diffusion equations. On the other hand, as we have shown in the present work, the LGCA offers an advantage not only in computational efficiency and straightforward multiparticle extension, but also in ease of model analysis. This gets rapidly complex in the aforementioned continuous approaches[1–8] but remains feasible in the LGCA, even when dealing with systems of interacting particles[35,38,42].

In the era of "Big Data", there is an abundance of biological data. Single or collective cell motility can be measured *in vitro* or *in vivo* via various experimental methods such as *in vivo* two-photon imaging[59] or cell cytometry[60], respectively. In this regard, there is a need for "data-driven" modeling frameworks. Our work comes timely to fulfill this scope by proposing the "data-driven" modeling of single particle superdiffusive behavior without prior knowledge of the mechanisms at work. Such an approach is vital for the study of phenomena whose driving mechanisms are currently unknown or challenging to model[61–63].

References

1. Metzler, R. & Klafter, J. The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Phys. Rep.* **339**, 1–77 (2000).
2. Coffey, W., Kalmykov, Y. & Waldron, J. The Langevin Equation (World Scientific, Singapore, 2004).
3. Metzler, R. & Klafter, J. The restaurant at the end of the random walk: Recent developments in the description of anomalous transport by fractional dynamics. *J. Phys. A: Math. Gen.* **37**, R161–R208 (2004).
4. Klages, R., Radons, G. & Sokolov, I. (eds) Anomalous transport (Wiley-VCH, Berlin, 2008).
5. Klafter, J. & Sokolov, I. First Steps in Random Walks: From Tools to Applications (Oxford University Press, Oxford, 2011).
6. Höfling, F. & Franosch, T. Anomalous transport in the crowded world of biological cells. *Rep. Prog. Phys.* **76**, 046602/1–50 (2013).
7. Metzler, R., Jeon, J.-H., Cherstvy, A. G. & Barkai, E. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys.* **16**, 24128–24164 (2014).
8. Zaburdaev, V., Denisov, S. & Klafter, J. Lévy walks. *Rev. Mod. Phys.* **87**, 483–529 (2015).
9. Shlesinger, M., Zaslavsky, G. & Klafter, J. Strange kinetics. *Nature* **363**, 31–37 (1993).
10. Klafter, J., Shlesinger, M. F. & Zumofen, G. Beyond Brownian motion. *Phys. Today* **49**, 33–39 (1996).
11. Sokolov, I., Klafter, J. & Blumen, A. Fractional kinetics. *Phys. Today* **55**, 48–54 (2002).
12. Scher, H. & Montroll, E. Anomalous transit-time dispersion in amorphous solids. *Phys. Rev. B* **12**, 2455–2477 (1975).
13. Kukla, V. *et al.* NMR studies of single-file diffusion in unidimensional channel zeolites. *Science* **272**, 702–704 (1996).
14. Balescu, R. Aspects of anomalous transport in plasmas, vol. 18 of Series in Plasma Physics (CRC Press, London, 2005).
15. Barthelemy, P., Bertolotti, J. & Wiersma, D. A Lévy flight for light. *Nature* **453**, 495–498 (2008).
16. Viswanathan, G., da Luz, M., Raposo, E. & Stanley, H. The Physics of Foraging (Cambridge University Press, Cambridge, 2011).
17. Brockmann, D., Hufnagel, L. & Geisel, T. The scaling laws of human travel. *Nature* **439**, 462–465 (2006).
18. Upadhyaya, A., Rieu, J., Glazier, J. & Sawada, Y. Anomalous diffusion and non-Gaussian velocity distributions of hydra cells in cellular aggregates. *Physica A* **293**, 549–558 (2001).
19. Thurner, S., Wick, N., Hanel, R., Sedivy, R. & Huber, L. Anomalous diffusion on dynamical networks: a model for interacting epithelial cell migration. *Physica A: Statistical Mechanics and its Applications* **320**, 475–484 (2003).
20. Dieterich, P., Klages, R., Preuss, R. & Schwab, A. Anomalous dynamics of cells migration P. *Natl. Acad. Sci. USA* **105**, 459–463 (2008).
21. Takagi, H., Sato, M., Yanagida, T. & Ueda, M. Functional analysis of spontaneous cell movement under different physiological conditions. *PLoS ONE* **3**, e2648/1–8 (2008).
22. Bödeker, H., Beta, C., Frank, T. & Bodenschatz, E. Quantitative analysis of random ameboid motion. *Europhys. Lett.* **90**, 28005/1–5 (2010).
23. Harris, T. *et al.* Generalized Lévy walks and the role of chemokines in migration of effector cd8+ T cells. *Nature* **486**, 545–548 (2012).
24. Metzner, C. *et al.* Superstatistical analysis and modelling of heterogeneous random walks. *Nat. Comm.* **6**, 7516–7523 (2015).
25. Barbaric, I. *et al.* Time-lapse analysis of human embryonic stem cells reveals multiple bottlenecks restricting colony formation and their relief upon culture adaptation. *Stem Cell Reports* **3**, 142–155 (2014).
26. Rieu, J., Upadhyaya, A., Glazier, J., Ouchi, N. & Sawada, Y. Diffusion and deformations of single hydra cells in cellular aggregates. *Biophys. J.* **79**, 1903–1914 (2000).
27. Krummel, M., Bartumeus, F. & Gérard, A. T cell migration, search strategies and mechanisms. *Nat. Rev. Immun.* **16**, 193–201 (2016).
28. Méhes, E. & Vicsek, T. Collective motion of cells: from experiments to models. *Integr. Biol.* **6**, 831–854 (2014).
29. McCann, C., Kriebel, P., Parent, C. & Losert, W. Cell speed, persistence and information transmission during signal relay and collective migration. *J. Cell Sci.* **123**, 1724–1731 (2010).
30. Malinverno, C. *et al.* Endocytic reawakening of motility in jammed epithelia. *Nat. Mat.* **16**, 587–597 (2017).
31. Bi, D., Yang, X., Marchetti, M. & Manning, M. Motility-driven glass and jamming transitions in biological tissues. *Phys. Rev. X.* **6**, 021011/1–13 (2016).
32. Fodor, É. *et al.* How far from equilibrium is active matter? *Phys. Rev. Lett.* **117**, 038103/1–6 (2016).
33. Fedotov, S. & Korabel, N. Emergence of Lévy walks in systems of interacting individuals. *Phys. Rev. E* **95**, 030107/1–6 (2017).
34. Ariel, G. *et al.* Swarming bacteria migrate by Lévy walk. *Nat. Commun.* **6**, 8396/1–6 (2015).
35. Mente, C., Voß-Böhme, A. & Deutsch, A. Analysis of individual cell trajectories in lattice-gas cellular automaton models for migrating cell populations. *B. Math. Biol.* **77**, 660–697 (2015).
36. Chen, J. C. & Kim, A. S. Brownian dynamics, molecular dynamics, and Monte Carlo modeling of colloidal systems. *Adv. Colloid Interfac.* **112**, 159–173 (2004).
37. Weitz, S., Deutsch, A. & Peruani, F. Self-propelled rods exhibit a phase-separated state characterized by the presence of active stresses and the ejection of polar clusters. *Phys. Rev. E* **92**, 012322 (2015).
38. Deutsch, A. & Dormann, S. Cellular Automaton Modeling of Biological Pattern Formation:Characterization, Applications, and Analysis (Birkhauser, Boston, 2017).
39. Hatzikirou, H., Brusch, L. & Deutsch, A. From cellular automaton rules to a macroscopic mean-field description. *Acta Phys. Pol. B Proc. Suppl.* **3**, 399–416 (2010).
40. Reher, D., Klink, B., Deutsch, A. & Voß-Böhme, A. Cell adhesion heterogeneity reinforces tumour cell dissemination: novel insights from a mathemtical model. *Biol. Direct* **12**, 18 (2017).
41. Böttger, K. *et al.* An emerging Allee effect is critical for tumor initiation and persistence. *PLoS Comput. Biol.* **11**, e1004366 (2015).
42. Böttger, K., Hatzikirou, H., Chauvière, A. & Deutsch, A. Investigation of the migration/proliferation dichotomy and its impact on avascular glioma invasion. Math. Model. *Nat. Phenom.* **7**, 105–135 (2012).
43. Hatzikirou, H., Basanta, D., Simon, M., Schaller, K. & Deutsch, A. 'Go or Grow': the key to the emergence of invasion in tumour progression? *Math. Med. Biol.* **29**, 49–65 (2012).

44. Pearson, K. Mathematical contributions to the theory of evolution - a mathematical theory of random migration. *Biometric ser.* **3**, 54 (1906).
45. Barber, M. N. & Ninham, B. W. Random and restricted walks - Theory and applications. (Gordon and Breach, New York, 1970).
46. Lawniczak, A. T. Lattice gas automata for diffusive-convective transport dynamics. Center for Nonlinear Studies, Newsletter No. 136 (1997).
47. Klages, R. & Korabel, N. Understanding deterministic diffusion by correlated random walks. *J. Phys. A: Math. Gen.* **35**, 4823–4836 (2002).
48. Reif, F. Fundamentals of statistical and thermal physics (McGraw-Hill, Auckland, 1965).
49. Peruani, F., Deutsch, A. & Bär, M. A mean-field theory for self-propelled particles interacting by velocity alignment mechanisms. *Eur. Phys. J.-Spec. Top.* **157**, 111–122 (2008).
50. Nava-Sedeño, J. M., Hatzikirou, H., Peruani, F. & Deutsch, A. Extracting cellular automaton rules from physical Langevin equation models for single and collective cell migration. *J. Math. Biol.* 1–26 (2017).
51. Schüring, A. Molekulardynamik-Simulationen und Sprungmodelle zur Diffusion in Zeolithen. (Technische Universität Leipzig, 2003).
52. Shalchi, A. Applicability of the Taylor-Green-Kubo formula in particle diffusion theory. *Phys. Rev. E* **83**, 046402 (2011).
53. Chechkin, A., Lenz, F. & Klages, R. Normal and anomalous fluctuation relations for Gaussian stochastic dynamics. *J. Stat. Mech.: Theor. Exp.* **2012**, L11001/1–13 (2012).
54. Korabel, N., Klages, R., Chechkin, A., Sokolov, I. & Gonchar, V. Fractal properties of anomalous diffusion in intermittent maps. *Phys. Rev. E* **75**, 036213–1–14 (2007).
55. Pressé, S., Ghosh, K., Lee, J. & Dill, K. A. Principles of maximum entropy and maximum caliber in statistical physics. *Rev. Mod. Phys.* **85**, 1115–1141 (2013).
56. Hazoglou, M. J., Walther, V., Dixit, P. D. & Dill, K. A. Communication: Maximum caliber is a general variational principle for nonequilibrium statistical mechanics. *J. Chem. Phys.* **143**, 051104 (2015).
57. Othmer, H. G., Dunbar, S. R. & Alt, W. J. Models of dispersal in biological systems. *J. Math. Biol.* **26**, 263–298 (1988).
58. Talkenberger, K., Cavalcanti-Adam, E. A., Deutsch, A. & Voß-Böhme, A. Amoeboid-mesenchymal migration plasticity promotes invasion only in complex heterogeneous microenvironments. *Sci. Rep* (2017).
59. Cahalan, M. D., Parker, I., Wei, S. H. & Miller, M. J. Real-time imaging of lymphocytes *in vivo*. *Curr. Opin. Immunol.* **15**, 372–377 (2003).
60. del Álamo, J. C. *et al.* Spatio-temporal analysis of eukaryotic cell motility by improved force cytometry. *P. Natl. Acad. Sci. USA* **104**, 13343–13348 (2007).
61. Zienkiewicz, A., Barton, D. A. W., Porfiri, M. & di Bernardo, M. Data-driven stochastic modelling of zebrafish locomotion. *J. Math. Biol.* **71**, 1081–1105 (2015).
62. Christopher, R. *et al.* Data-driven computer simulation of human cancer cell. *Ann. NY Acad. Sci.* **1020**, 132–153 (2004).
63. Sisan, D. R., Halter, M., Hubbard, J. B. & Plant, A. L. Predicting rates of cell state change caused by stochastic fluctuations using a data-driven landscape model. *P. Natl. Acad. Sci. USA* **109**, 19262–19267 (2012).

Acknowledgements

Author Contributions

All authors formulated the mathematical model. J.M.N.S. performed the analysis and simulations. All authors interpreted results. J.M.N.S. wrote the manuscript with contributions from all authors. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-17317-x.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cellular automaton models for time-correlated random walks: derivation and analysis

J. M. Nava-Sedeño, H. Hatzikirou, R. Klages, A. Deutsch

Supplementary information

MSD in correlated systems

Here we derive the general expression for the MSD when there are temporal correlations. We use the derivation by Schüring[51]. To start with, we consider the "MSD" as defined by Shalchi[52]

$$\langle r_i r_j \rangle = \int_{t_0}^{t} \int_{t_0}^{t} \langle v_i(\tau) v_j(\xi) \rangle \, \mathrm{d}\xi \, \mathrm{d}\tau$$

which in LGCA notation are written as

$$\langle r_i r_j \rangle = \sum_{n=1}^{k} \sum_{m=1}^{k} v^2 \tau^2 \left\langle [\hat{e}_i \cdot \vec{c}_{i_n}] [\hat{e}_j \cdot \vec{c}_{i_m}] \right\rangle, \tag{S1}$$

where $\hat{e}_x$ and $\hat{e}_y$ are the two orthonormal unit vectors in Cartesian coordinates. In 2D the diagonal elements are given by:

$$\begin{aligned}
\langle r_x^2 \rangle &= \sum_{i=1}^{k} \sum_{j=1}^{k} v^2 \tau^2 \left\langle \cos(\theta_i) \cos(\theta_j) \right\rangle \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{v^2 \tau^2}{2} \left\langle \cos(\theta_i - \theta_j) + \cos(\theta_i + \theta_j) \right\rangle
\end{aligned} \tag{S2}$$

and

$$\begin{aligned}
\langle r_y^2 \rangle &= \sum_{i=1}^{k} \sum_{j=1}^{k} v^2 \tau^2 \left\langle \sin(\theta_i) \sin(\theta_j) \right\rangle \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{v^2 \tau^2}{2} \left\langle \cos(\theta_i - \theta_j) - \cos(\theta_i + \theta_j) \right\rangle,
\end{aligned} \tag{S3}$$

where $\theta_k = \arg\left[\vec{c}_{i_k}\right]$. Adding them up gives the MSD

$$\langle r^2 \rangle = \langle r_x^2 \rangle + \langle r_y^2 \rangle = \sum_{i=1}^{k} \sum_{j=1}^{k} v^2 \tau^2 \left\langle \cos(\theta_i - \theta_j) \right\rangle, \tag{S4}$$

which is just a Taylor-Green-Kubo formula[47,54]. When $i = j$ we have $\cos(\theta_i - \theta_j) = 1$, so by taking these terms out of the sum we get

$$\langle r^2 \rangle = k v^2 \tau^2 + \sum_{i=1}^{k} \sum_{j=1}^{k} v^2 \tau^2 \left\langle \cos(\theta_i - \theta_j) \right\rangle (1 - \delta_{ij}). \tag{S5}$$

On the one hand $v^2 \tau^2 = \varepsilon^2$ and on the other $\varepsilon^2 = 2dD\tau$, so using these relations on the first term on the right hand side yields

$$\langle r^2 \rangle = 2dDk\tau + \sum_{i=1}^{k} \sum_{j=1}^{k} v^2 \tau^2 \left\langle \cos(\theta_i - \theta_j) \right\rangle (1 - \delta_{ij}). \tag{S6}$$

Because the cosine is an even function $\cos(\theta_i - \theta_j) = \cos(\theta_j - \theta_i)$, which means that we are adding two identical terms for every $i \neq j$ (this condition is already satisfied due to the Kronecker delta). This situation allows us to rewrite the limits of the interior sum if we take care of counting each term twice. Furthermore, using trigonometric identities it is possible to expand the cosine on the right hand side,

$$\begin{aligned}
\langle r^2 \rangle = {}& 2dDk\tau + 2 \sum_{i=1}^{k} \sum_{j=i}^{k} v^2 \tau^2 \left\langle \cos(\theta_i) \cos(\theta_j) \right\rangle (1 - \delta_{ij}) \\
&+ 2 \sum_{i=1}^{k} \sum_{j=i}^{k} v^2 \tau^2 \left\langle \sin(\theta_i) \sin(\theta_j) \right\rangle (1 - \delta_{ij}).
\end{aligned} \tag{S7}$$

If the orientations of the particle are completely uncorrelated from one another, then it is possible to write the sums on the right as

$$\langle r^2 \rangle = 2dDk\tau + 2\sum_{i=1}^{k}\sum_{j=i}^{k} v^2\tau^2 \langle \cos(\theta_i) \rangle \langle \cos(\theta_j) \rangle (1 - \delta_{ij})$$
$$+ 2\sum_{i=1}^{k}\sum_{j=i}^{k} v^2\tau^2 \langle \sin(\theta_i) \rangle \langle \sin(\theta_j) \rangle (1 - \delta_{ij}). \tag{S8}$$

If the reorientation probabilities are even functions of the angle θ_k, then we have that $\langle \sin(\theta_i) \rangle = 0$, and so the second sum on the right hand side disappears leaving

$$\langle r^2 \rangle = 2dDk\tau + 2\sum_{i=1}^{k}\sum_{j=i}^{k} v^2\tau^2 \langle \cos(\theta_i) \rangle \langle \cos(\theta_j) \rangle (1 - \delta_{ij}). \tag{S9}$$

We can choose our coordinate system such that $\theta_0 = 0$ without loss of generality. With this choice of the coordinate system, we can rewrite the MSD as

$$\langle r^2 \rangle = 2dDk\tau + 2\sum_{i=1}^{k}\sum_{j=i}^{k} v^2\tau^2 \langle \cos(\theta_i - \theta_0) \rangle \langle \cos(\theta_j - \theta_0) \rangle (1 - \delta_{ij}). \tag{S10}$$

Using the definition of the VACF this becomes

$$\langle r^2 \rangle = 2dDk\tau + 2\sum_{i=1}^{k}\sum_{j=i}^{k} v^2\tau^2 g(i)g(j)(1 - \delta_{ij}). \tag{S11}$$

Expanding the difference on the right hand side we get

$$\langle r^2 \rangle = 2dDk\tau + 2\sum_{i=1}^{k}\sum_{j=i}^{k} v^2\tau^2 g(i)g(j) - 2\sum_{i=1}^{k}\sum_{j=i}^{k} v^2\tau^2 g(i)g(j)\delta_{ij}. \tag{S12}$$

The last sum on the right hand side can be simplified, so we get

$$\langle r^2 \rangle = 2dDk\tau + 2\sum_{i=1}^{k}\sum_{j=i}^{k} v^2\tau^2 g(i)g(j) - 2\sum_{i=1}^{k} v^2\tau^2 g^2(i). \tag{S13}$$

Finally, using the relation between the intantaneous particle velocity v and the diffusion constant in the random walk limit D_{rw}, and reordering terms, we obtain

$$\langle r^2 \rangle = 2dD_{rw}\left[k\tau - 2\sum_{i=1}^{k} g^2(i)\tau \right] + 2v^2 \sum_{i=1}^{k}\sum_{j=i}^{k} g(i)g(j)\tau^2. \tag{S14}$$

VACF and MSD derivation in the persistent random walk

First we analytically derive the expected form of the VACF for a single particle in an LGCA where the lattice is a 2D square lattice.

VACF
As mentioned before, the orientation probability is given by Eq. (10). The VACF is formally defined as $g(t) = \langle \vec{v}_0 \cdot \vec{v}_t \rangle$, where v_0 and v_t are the velocities of the particle at time 0 and t, respectively. Using this definition, we can calculate the VACF of a stochastically moving particle as

$$g(t) = \int P(\vec{v}, t)(\vec{v}_0 \cdot \vec{v})\, d\vec{v},$$

where $P(\vec{v}, t)$ is the probability of the particle having a velocity $\vec{v}$ at time t.

In an LGCA particle velocities are given by the velocity channels they are located in, which belong to a finite set of unit vectors depending on the lattice dimension and geometry. Furthermore, time is also discrete with time steps of length τ such

that at time step k time has elapsed by $k\tau$. We can then rewrite the definition of the velocity autocorrelation in the case of an LGCA in the following way[47]:

$$g(t) = \langle \vec{v}_0 \cdot \vec{v}_t \rangle = \sum_{i=1}^{b} P_{i_k,k} \left[\vec{c}_{i_0} \cdot \vec{c}_{i_k} \right],$$ (S15)

where $\vec{c}_{i_0}$ is the orientation of the particle at time step $k = 0$ and $\vec{c}_{i_k}$ is the orientation of the particle at time step k. To calculate the VACF, we start by defining a function as:

$$H = -\vec{c}_{i_{k-1}} \cdot \vec{c}_{i_k},$$ (S16)

where $\vec{c}_{i_k}$ is the particle orientation at time step k. Having defined this function, we can rewrite Eq. (10) as follows:

$$P_{i_k,k} = \frac{e^{-\beta H(\vec{c}_{i_k},k)}}{Z},$$ (S17)

where the partition function is defined as

$$Z = \sum_{i_k} e^{-\beta H(\vec{c}_{i_k},k)}.$$ (S18)

The expected value of the function is given by

$$\langle H \rangle = \left\langle -\vec{c}_{i_{k-1}} \cdot \vec{c}_{i_k} \right\rangle,$$ (S19)

that is, the energy of the system is the single-step correlation. Due to the distribution of the reorientation probabilites the total energy can be calculated by the well-known relation

$$\langle H \rangle = -\frac{\partial}{\partial \beta} \ln Z.$$ (S20)

Using the last two equations, we get an expression for the single step correlation:

$$\left\langle \vec{c}_{i_{k-1}} \cdot \vec{c}_{i_k} \right\rangle = \frac{\partial}{\partial \beta} \ln Z.$$ (S21)

In this single-particle model, the partition function can be easily calculated. For a 2D square lattice the partition function reads

$$Z = 2 \left[1 + \cosh(\beta) \right].$$ (S22)

Substituting Supplementary Eq. (S22) into Supplementary Eq. (S21), we have

$$\begin{aligned}
\left\langle \vec{c}_{i_{k-1}} \cdot \vec{c}_{i_k} \right\rangle &= \frac{\partial}{\partial \beta} \left\{ \ln(2) + \ln \left[1 + \cosh(\beta) \right] \right\} \\
&= \frac{\sinh(\beta)}{1 + \cosh(\beta)} = \tanh \left(\frac{\beta}{2} \right).
\end{aligned}$$ (S23)

The particle orientations $\vec{c}_i$ are normalized vectors. This allows us to rewrite the single step correlation as

$$\left\langle \vec{c}_{i_{k-1}} \cdot \vec{c}_{i_k} \right\rangle = \left\langle \cos \left(\theta_k - \theta_{k-1} \right) \right\rangle,$$ (S24)

and the VACF as

$$g(k) = \left\langle \cos \left(\theta_k - \theta_0 \right) \right\rangle,$$ (S25)

where $\theta_k = \arg \left[\vec{c}_k \right]$. Supplemenary Eq. (S25) can be rewritten by adding zeroes in the following way:

$$g(k) = \left\langle \cos \left(\theta_k - \theta_0 \right) \right\rangle = \left\langle \cos \left[\theta_k - \theta_0 + \sum_{i=1}^{k-1} \left(\theta_i - \theta_i \right) \right] \right\rangle,$$

which after rearranging terms, has the form

$$g(k) = \left\langle \cos\left(\sum_{i=1}^{k} \theta_i - \theta_{i-1}\right) \right\rangle. \tag{S26}$$

Using trigonometric identities and the linearity of the expected value operator we can expand this expression to

$$g(k) = \left\langle \prod_{i=1}^{k} \cos\left(\theta_i - \theta_{i-1}\right) \right\rangle + f\left\{ \langle \cos\left(\theta_n - \theta_{n-1}\right) \sin\left(\theta_m - \theta_{m-1}\right) \rangle \right\}, \tag{S27}$$

where f is a sum of expected values of products of sines and cosines. Because the model is Markovian, the i-th orientation is only correlated with the $(i-1)$-th orientation. This allows us to write

$$g(k) = \prod_{i=1}^{k} \langle \cos\left(\theta_i - \theta_{i-1}\right) \rangle + f\left\{ \langle \cos\left(\theta_n - \theta_{n-1}\right) \rangle \langle \sin\left(\theta_m - \theta_{m-1}\right) \rangle \right\}. \tag{S28}$$

Now, because the reorientation probabilities Eq. (10) are even functions with respect to $\theta_k = \arg[\vec{c}(k)]$, the expected values become

$$\langle \sin\left(\theta_m - \theta_{m-1}\right) \rangle = 0,$$

which in turn implies

$$f\left\{ \langle \cos\left(\theta_n - \theta_{n-1}\right) \rangle \langle \sin\left(\theta_m - \theta_{m-1}\right) \rangle \right\} = 0.$$

Using these relations we find that the VACF is given by

$$g(k) = \prod_{i=1}^{k} \langle \cos\left(\theta_i - \theta_{i-1}\right) \rangle. \tag{S29}$$

Using Supplementary Eqs. (S23) and (S24) in Supplementary Eq. (S29) yields

$$g(k) = \left[\tanh\left(\frac{\beta}{2}\right) \right]^k, \tag{S30}$$

which can be written as

$$g(k) = e^{\alpha k}, \tag{S31}$$

if we define the exponent α as

$$\alpha = \ln\left[\tanh\left(\frac{\beta}{2}\right) \right]. \tag{S32}$$

The exponent α depends on the lattice dimension and geometry, as follows:

- In 1D the exponent is given by:

$$\alpha = \ln\left[\tanh\left(\beta\right) \right]. \tag{S33}$$

- In 2D with a triangular lattice the exponent is:

$$\alpha = \ln\left(\frac{e^{\frac{\beta}{2}} - e^{-\beta}}{e^{-\beta} + 2e^{\frac{\beta}{2}}} \right). \tag{S34}$$

- In 2D with a square lattice the exponent is given by:

$$\alpha = \ln\left[\tanh\left(\frac{\beta}{2}\right) \right]. \tag{S35}$$

- In 2D with an hexagonal lattice it takes the form:

$$\alpha = \ln \left[\frac{2\sinh\left(\frac{3\beta}{4}\right)\cosh\left(\frac{\beta}{4}\right)}{\cosh(\beta) + 2\cosh\left(\frac{\beta}{2}\right)} \right].$$

(S36)

- With a cubic 3D lattice the exponent reads:

$$\alpha = \ln \left[\frac{\sinh(\beta)}{\cosh(\beta) + 2} \right].$$

(S37)

Mean square displacement

To calculate the MSD of particles performing persistent random walks, we start with Supplementary Eqs. (S6) and (S24) and rewrite the sum limits taking into account that the cosine is an even function to obtain

$$\langle r^2 \rangle = 2dDk\tau + 2v^2 \sum_{n=1}^{k} \sum_{m=n}^{k} \langle \vec{c}_{i_n} \cdot \vec{c}_{i_m} \rangle \tau^2 (1 - \delta_{nm}).$$

(S38)

The expected value on the right hand side is the $(m-n)$-step correlation. From the previous VACF calculation we know that in the Markovian model

$$\langle \vec{c}_{i_n} \cdot \vec{c}_{i_m} \rangle := g(m-n) = \left[\tanh\left(\frac{\beta}{2}\right) \right]^{|m-n|},$$

(S39)

which can be substituted in the expression of the MSD to obtain

$$\langle r^2 \rangle = 2dDk\tau + 2v^2 \sum_{n=1}^{k} \sum_{m=n}^{k} g(m-n) \tau^2 (1 - \delta_{nm}).$$

(S40)

Because the sums on the right hand side only depend on the interval length $|n-m|$ and not on the specific values of the indices n and m we can replace both sums by a sum over all posible interval lengths. There are $k-j$ ways to divide an interval of k time steps (because the sums start from $n=1$) into intervals of size j. Taking all into account, the MSD becomes

$$\langle r^2 \rangle = 2dDk\tau + 2v^2 \sum_{j=1}^{k} (k-j)g(j)\tau^2,$$

(S41)

which can also be written as

$$\langle r^2 \rangle = 2dDk\tau + 2v^2 \sum_{j=1}^{k} (k-j) e^{j\alpha} \tau^2,$$

where α is given by Supplementary Eq. (S32). We now distribute the two multiplying time steps τ on the second term on the right hand side, and multiply by one the exponent of the exponential function thus leaving it unchanged:

$$\langle r^2 \rangle = 2dDk\tau + 2v^2 \sum_{j=1}^{k} (k\tau - j\tau) e^{\frac{\alpha}{\tau} j\tau} \tau.$$

We now use the definitions of the diffusion coefficient and the particle speed to obtain the following expression for the time step length:

$$\tau = \frac{2dD}{v^2},$$

and use it to substitute for τ on the denominator of the exponent

$$\langle r^2 \rangle = 2dDk\tau + 2v^2 \sum_{j=1}^{k} (k\tau - j\tau) e^{\frac{v^2 \alpha}{2dD} j\tau} \tau.$$

VACF in homogeneous Markovian models

We will now consider a general Markovian model for a single moving particle. The model is then a Markov chain of particle orientations, i.e. the particle can transition between different orientations at each time step.

Definition 1. *The state space of the Markov chain is $\mathscr{E} = \{\vec{c}_0, \vec{c}_{\pm 1}, \cdots, \vec{c}_{\pm n}, \cdots, \vec{c}_N\}$, where the 2N different states are given by*

$$\vec{c}_n = \left(\cos\left(\frac{\pi}{n}\right), \sin\left(\frac{\pi}{n}\right) \right), \; n = 1, \cdots, N-1,$$
$$\vec{c}_0 = (1,0),$$
$$\vec{c}_N = (-1,0).$$

Definition 2. *The state space subset $\mathscr{E}_0$ is defined as*

$$\mathscr{E}_0 := \left\{ \vec{c}_{\pm 1}, \cdots, \vec{c}_{\pm(N-1)} \right\}.$$

If the space is isotropic, then it is reasonable to require that the probability of the particle turning left or right be identical. Furthermore, we assume that the probability of turning does not depend on the specific time step, i.e. that the Markov process is homogeneous.

Definition 3. *The Markov chain is the stochastic process $\{X(k) : k \in \mathbb{N}\}$ where the reorientation probabilities are given by*

$$P\left(X(k+1) = \vec{c}_m \mid X(k) = \vec{c}_n\right) := P\left(\theta := \arg(\vec{c}_n, \vec{c}_m)\right),$$

where $P(\theta) = P(-\theta)$ for $0 <| \theta |< \pi$, and the initial condition $X(0) = \vec{c}_0$.

We will use the following shorthand notation: $P(0) := p_0$, $P(\frac{\pi}{n}) = P(-\frac{\pi}{n}) := p_n$, and $P(\pi) := p_N$.

Definition 4. *The rotation matrix $\mathbf{A_n}$ is given by*

$$\mathbf{A_n} = \begin{pmatrix} \cos\left(\frac{\pi}{n}\right) & -\sin\left(\frac{\pi}{n}\right) \\ \sin\left(\frac{\pi}{n}\right) & \cos\left(\frac{\pi}{n}\right) \end{pmatrix},$$

such that

$$\mathbf{A_n} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\left(\frac{\pi}{n}\right) \\ \sin\left(\frac{\pi}{n}\right) \end{pmatrix} = \vec{c}_n$$

and

$$\mathbf{A_n} \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \end{pmatrix} = \begin{pmatrix} \cos\left(\frac{\pi}{n}\right)\cos(\phi) - \sin\left(\frac{\pi}{n}\right)\sin(\phi) \\ \sin\left(\frac{\pi}{n}\right)\cos(\phi) + \cos\left(\frac{\pi}{n}\right)\sin(\phi) \end{pmatrix} = \begin{pmatrix} \cos(\frac{\pi}{n} + \phi) \\ \sin(\frac{\pi}{n} + \phi) \end{pmatrix}.$$

Definition 5. *The velocity autocorrelation function (VACF) is given by*

$$g_k = \langle X(0) \cdot X(k) \rangle = \sum_{\vec{v} \in \mathscr{E}} (\vec{c}_0 \cdot \vec{v}) P^k(\vec{v})$$

Theorem 2. *The velocity autocorrelation function of a particle whose orientations are given by a homogeneous, symmetric Markov chain is either delta-correlated, i.e. $g_k = \delta_{0,k}$, where δ is the Kronecker delta; alternating, i.e. $g_k = (-1)^k a^k$, $a \in \mathbb{R}^+$; or exponentially decaying, i.e. $g_k = e^{\alpha k}$, $\alpha \leq 0$.*

Proof. The proof is by induction.

$$g_1 = \sum_{\vec{v} \in \mathscr{E}} (\vec{c}_0 \cdot \vec{v}) P(\vec{v}) = \sum_{\vec{v} = \vec{c}_0, \vec{c}_N} (\vec{c}_0 \cdot \vec{v}) P(\vec{v}) + \sum_{\vec{v} \in \mathscr{E}_0} (\vec{c}_0 \cdot \vec{v}) P(\vec{v}) = P(0) - P(\pi) + \sum_{i=1}^{N-1} \left[(\vec{c}_0 \cdot \vec{c}_i) P\left(\frac{\pi}{i}\right) + (\vec{c}_0 \cdot \vec{c}_{-i}) P\left(-\frac{\pi}{i}\right) \right]$$

$$p_0 - p_N + \sum_{i=1}^{N-1} \left[\cos\left(\frac{\pi}{i}\right) p_i + \cos\left(-\frac{\pi}{i}\right) p_i \right] = p_0 - p_N + 2 \sum_{i=1}^{N-1} \cos\left(\frac{\pi}{i}\right) p_i := a.$$

Using the Chapman-Kolmogorov equation, we can calculate the VACF at the time step $k+1$

$$g_{k+1} = \sum_{\vec{v} \in \mathscr{E}} (\vec{c}_0 \cdot \vec{v}) P^{k+1}(\vec{v}) = \sum_{\vec{v} \in \mathscr{E}} (\vec{c}_0 \cdot \vec{v}) \sum_{\vec{u} \in \mathscr{E}} P^k(\vec{u}) P(\vec{v} \mid \vec{u}) = \sum_{\vec{u} \in \mathscr{E}} P^k(\vec{u}) \sum_{\vec{v} \in \mathscr{E}} (\vec{c}_0 \cdot \vec{v}) P(\vec{v} \mid \vec{u}).$$

We now expand the second sum on the right hand side of the equation

$$
\begin{aligned}
\sum_{\vec{v} \in \mathscr{E}} (\vec{c}_0 \cdot \vec{v}) P(\vec{v} \mid \vec{u}) &= (\vec{c}_0 \cdot \vec{u}) p_0 - (\vec{c}_0 \cdot \vec{u}) p_N + \sum_{i=1}^{N-1} \left[(\vec{c}_0 \cdot \mathbf{A_i} \vec{u}) P(\mathbf{A_i} \vec{u} \mid \vec{u}) + (\vec{c}_0 \cdot \mathbf{A_{-i}} \vec{u}) P(\mathbf{A_{-i}} \vec{u} \mid \vec{u}) \right] \\
&= (\vec{c}_0 \cdot \vec{u}) p_0 - (\vec{c}_0 \cdot \vec{u}) p_N + \sum_{i=1}^{N-1} \left[\begin{pmatrix} \cos\left(\frac{\pi}{i}\right) \\ -\sin\left(\frac{\pi}{i}\right) \end{pmatrix}^T \vec{u} p_i + \begin{pmatrix} \cos\left(\frac{\pi}{i}\right) \\ \sin\left(\frac{\pi}{i}\right) \end{pmatrix}^T \vec{u} p_i \right] \\
&= (\vec{c}_0 \cdot \vec{u}) p_0 - (\vec{c}_0 \cdot \vec{u}) p_N + \sum_{i=1}^{N-1} \begin{pmatrix} 2\cos\left(\frac{\pi}{i}\right) \\ 0 \end{pmatrix}^T \vec{u} p_i \\
&= (\vec{c}_0 \cdot \vec{u}) p_0 - (\vec{c}_0 \cdot \vec{u}) p_N + 2 \sum_{i=1}^{N-1} \cos\left(\frac{\pi}{i}\right) (\vec{c}_0 \cdot \vec{u}) p_i \\
&= (\vec{c}_0 \cdot \vec{u}) \left[p_0 - p_N + 2 \sum_{i=1}^{N-1} \cos\left(\frac{\pi}{i}\right) p_i \right] = (\vec{c}_0 \cdot \vec{u}) a.
\end{aligned}
$$

Inserting this expression back into the VACF yields

$$
g_{k+1} = \sum_{\vec{u} \in \mathscr{E}} P^k(\vec{u}) \sum_{\vec{v} \in \mathscr{E}} (\vec{c}_0 \cdot \vec{v}) P(\vec{v} \mid \vec{u}) = \sum_{\vec{u} \in \mathscr{E}} P^k(\vec{u}) (\vec{c}_0 \cdot \vec{u}) a = g_k a = a^k a = a^{(k+1)}.
$$

We can rewrite a as $a = \sum_\theta p(\theta) \cos\theta$, where $\theta = \arg(\vec{c}_0, \vec{c}_n), \forall \vec{c}_n \in \mathscr{E}$. Using the fact that $0 \le p(\theta) \le 1$ and $\sum_\theta p(\theta) = 1$, we have

$$
-1 \le \cos(\theta) \le 1 \implies -p(\theta) \le p(\theta)\cos(\theta) \le p(\theta) \implies -1 \le \sum_\theta p(\theta)\cos\theta \le 1 \therefore -1 \le a \le 1.
$$

We have three cases:

- $-1 \le a < 0$, then $a = -1 \mid a \mid$ and $g_k = a^k = (-1)^k \mid a \mid^k$.
- $a = 0$, then $g_k = a^k = 0, k \ne 0$.
- $0 < a \le 1$ then $g_k = a^k = e^{k\ln(a)} = e^{\alpha k}$, where $\alpha = \ln(a)$. $0 < a \le 1 \implies -\infty < \alpha \le 0$.

$\square$

Time correlated random walk: rule derivation for different dimensions and geometries

One dimension

We will now sketch our method for obtaining the reorientation probabilities $P_{i_k,k}$ in 1D. We start by expanding $g(k)$ for the first two time steps after $k\tau = t \ge \Delta$ (see Eq. (16)) for a 1D lattice. We will denote by the subscript f the lattice direction parallel to the original orientation of the particle. Similarly, the subscript r denotes the direction opposite to the original orientation of the particle. Numerical subscripts denote the time step at which the reorientation probability is evaluated.

Time step $k = 1$ Only two trajectories are possible after one time step. Their probabilities are given by $P_{f,1}$ and $P_{r,1}$. The normalization condition for these probabilties reads

$$
P_{f,1} + P_{r,1} = 1. \tag{S42}
$$

We now expand the VACF:

$$
\sum_{i=1}^{1} P_{i_k,k} \left[\vec{c}_{i_0} \cdot \vec{c}_{i_k} \right] = P_{f,1} - P_{r,1} = g(1). \tag{S43}
$$

We can substitute $P_{r,1}$ from Supplementary Eq. (S42) into Supplementary Eq. (S43)

$$
P_{f,1} - P_{r,1} = P_{f,1} - (1 - P_{f,1}) = 2P_{f,1} - 1 = g(1).
$$

Rearranging terms we obtain the probability for having the same orientation as originally to

$$
P_{f,1} = \frac{1 + g(1)}{2}. \tag{S44}
$$

Substituting Supplementary Eq. (S44) into Supplementary Eq. (S42) we obtain the probability for the particle to turn around after the first time step,

$$P_{r,1} = \frac{1 - g(1)}{2}. \tag{S45}$$

After inspection of Supplementary Eqs. (S44) and (S45), and recalling that in the 1D lattice $c_1 = 1$ and $c_2 = -1$, both probabilities can be written as a single expression,

$$P_{i_1,1} = \frac{1 + \left[c_{i_0} \cdot c_{i_1}\right] g(1)}{2}, \tag{S46}$$

where i is a placeholder variable for either f or r.

Time step $k = 2$ After two time steps, we have four different possible paths for the particle, with four different probabilities. If we assume the probabilities at each time can be written as $P_{ff} = P_{f,1}P_{f,2}$, we can expand the VACF to obtain:

$$P_{f,1}P_{f,2} - P_{f,1}P_{r,2} + P_{r,1}P_{f,2} - P_{r,1}P_{r,2} = \left(P_{f,2} - P_{r,2}\right)\left(P_{f,1} + P_{r,1}\right) = g(2),$$

which by employing Supplementary Eq. (S42) can simplified to:

$$P_{f,2} - P_{r,2} = g(2). \tag{S47}$$

Given that the probabilities in the previous time step were normalized, it is sufficient to require that the probabilities in the current time step be normalized:

$$P_{f,2} + P_{r,2} = 1. \tag{S48}$$

Inspecting Supplementary Eqs. (S47) and (S48) and comparing them with Supplementary Eqs. (S42) and (S43) we can see that they are identical except for the evaluation of $g(k)$. Therefore, for the second time step it holds that

$$P_{i_2,2} = \frac{1 + \left[c_{i_0} \cdot c_{i_2}\right] g(2)}{2}. \tag{S49}$$

Any k It is easy to see that for further times we can always assume that the probabilities are uncorrelated so that only the last orientation in the particle's orientation history is relevant for the calculation. If we do, Supplementary Eqs. (S46) and (S49) can be generalized for any time step k in the following way:

$$P_{i_k,k} = \frac{1 + \left[c_{i_0} \cdot c_{i_k}\right] g(k)}{2}. \tag{S50}$$

Two dimensions: Triangular lattice
We will repeat the calculation we did in 1D now in 2D for two different lattice geometries to identify possible dependencies on the lattice dimension and/or geometry.

Time step $k = 1$ We have three possible lattice directions with lattice vectors given by either $\vec{c}_1 = (1,0)$, $\vec{c}_2 = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$, $\vec{c}_3 = \left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right)$, or $\vec{c}_1 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$, $\vec{c}_2 = (-1,0)$, $\vec{c}_3 = \left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right)$ on alternating nodes. In the first time step there are three possible paths given by $P_{r,1}$, $P_{a,1}$, and $P_{u,1}$, where $P_{r,1}$ is the probability to reverse orientation. The normalization condition is, in this case, given by

$$P_{r,1} + P_{u,1} + P_{a,1} = 1, \tag{S51}$$

while the VACF is given by

$$\frac{1}{2}\left(P_{u,1} + P_{a,1}\right) - P_{r,1} = g(1). \tag{S52}$$

We need to make an assumption to continue, as there are more variables than equations. We assume that the probability of turning left or right is identical,

$$P_{u,1} = P_{a,1} := P_{f,1}. \tag{S53}$$

Under this assumption we can rewrite Supplementary Eq. (S51) as

$$P_{r,1} + 2P_{f,1} = 1, \tag{S54}$$

and Supplementary Eq. (S52) as

$$P_{f,1} - P_{r,1} = g(1). \tag{S55}$$

Substituting $P_{f,1}$ from Supplementary Eq. (S55) into Supplementary Eq. (S54) we obtain

$$P_{r,1} + 2(P_{r,1} + g(1)) = 3P_{r,1} + 2g(1) = 1,$$

which, after rearranging, gives the expression for the probability of the particle to go back:

$$P_{r,1} = \frac{1 - 2g(1)}{3}. \tag{S56}$$

Using Supplementary Eq. (S56) in Supplementary Eq. (S54) we obtain the probability

$$P_{f,1} = \frac{1 + g(1)}{3}. \tag{S57}$$

Examining Supplementary Eqs. (S53), (S56) and (S57) we can summarize them as

$$P_{i_1,1} = \frac{1 + 2\left[\vec{c}_{i_0} \cdot \vec{c}_{i_1}\right] g(1)}{3}. \tag{S58}$$

Time step $k = 2$ In this case there are 9 different possible orientation histories with 9 different probabilities. If we now denote by f and r the lattice directions parallel and antiparallel to the original particle orientation, respectively, and by u and a the remaining lattice directions and assume that the probabilities are uncorrelated, we require that probabilites at the present time step are normalized:

$$P_{u,2} + P_{a,2} + P_{f,2} = 1, \tag{S59}$$

while the VACF has the form

$$P_{u,1}P_{f,2} - \frac{1}{2}P_{u,1}P_{u,2} - \frac{1}{2}P_{u,1}P_{a,2} + P_{a,1}P_{f,2}$$

$$- \frac{1}{2}P_{a,1}P_{u,2} - \frac{1}{2}P_{a,1}P_{a,2} + P_{r,1}P_{f,2} - \frac{1}{2}P_{r,1}P_{u,2} - \frac{1}{2}P_{r,1}P_{a,2}$$

$$= \left[P_{f,2} - \frac{1}{2}\left(P_{u,2} + P_{a,2}\right)\right](P_{u,1} + P_{a,1} + P_{r,1}) = g(2),$$

which, by Supplementary Eq. (S51), is simplified to:

$$P_{f,2} - \frac{1}{2}\left(P_{u,2} + P_{a,2}\right) = g(2). \tag{S60}$$

To continue, we impose the isotropy condition Supplementary Eq. (S53) denoting by $P_{r,2}$ the probabilities $P_{u,2}$ and $P_{a,2}$. With these assumptions the normalization condition reads

$$P_{f,2} + 2P_{r,2} = 1, \tag{S61}$$

while the VACF is now

$$P_{f,2} - P_{r,2} = g(2). \tag{S62}$$

Inserting $P_{f,2}$ from Supplementary Eq. (S61) into Supplementary Eq. (S62) we obtain the probability $P_{r,2}$:

$$P_{r,2} = \frac{1 - g(2)}{3} \tag{S63}$$

and, using the normalization condition Supplementary Eq. (S61) we obtain the probability $P_{f,2}$:

$$P_{f,2} = \frac{1 + 2g(2)}{3}. \tag{S64}$$

These probabilities can be written in the general form

$$P_{i_2,2} = \frac{1 + 2\left[\vec{c}_{i_0} \cdot \vec{c}_{i_2}\right] g(2)}{3}. \tag{S65}$$

Any k From Supplementary Eqs. ([S58](#)) and ([S65](#)) we can see that for any further time step k and making the same assumptions as before the probabilities are given by

$$P_{i_k,k} = \frac{1 + 2 \left[\vec{c}_{i_0} \cdot \vec{c}_{i_k} \right] g(k)}{3}. \tag{S66}$$

Two dimensions: Square lattice

Time step $k = 1$ There are four possible lattice directions with lattice vectors $\vec{c}_1 = (1,0)$, $\vec{c}_2 = (0,1)$, $\vec{c}_3 = (-1,0)$, and $\vec{c}_4 = (0,-1)$. Therefore there are four possible probabilities, so the normalization condition reads

$$P_{r,1} + P_{f,1} + P_{u,1} + P_{a,1} = 1, \tag{S67}$$

where $P_{u,1}$ and $P_{a,1}$ are the probabilities of going in the two directions orthogonal to the original orientation of the particle. We now expand the VACF to obtain

$$P_{f,1} - P_{r,1} = g(1). \tag{S68}$$

Right from the start we have more variables than equations, so we need to make one more assumption in order to continue with the derivation. To simplify we assume the following:

$$P_{u,1} = P_{a,1} := \frac{1}{4}. \tag{S69}$$

With this assumption the normalization condition becomes

$$P_{f,1} + P_{r,1} = \frac{1}{2}. \tag{S70}$$

Inserting Supplementary Eq. ([S70](#)) into Supplementary Eq. ([S68](#)) we obtain

$$P_{f,1} - \left(\frac{1}{2} - P_{f,1} \right) = 2P_{f,1} - \frac{1}{2} = g(1).$$

Rearranging terms we obtain the probability

$$P_{f,1} = \frac{1 + 2g(1)}{4}. \tag{S71}$$

Inserting Supplementary Eq. ([S71](#)) into Supplementary Eq. ([S70](#)) we obtain the remaining probability

$$P_{r,1} = \frac{1 - 2g(1)}{4}. \tag{S72}$$

Supplementary Equations ([S69](#)), ([S71](#)) and ([S72](#)) can then be summarized as

$$P_{i_1,1} = \frac{1 + 2 \left[\vec{c}_{i_0} \cdot \vec{c}_{i_1} \right] g(1)}{4}. \tag{S73}$$

Time step $k = 2$ There are now 16 different possible histories for the traveling particle. As before we assume that the probabilities are uncorrelated which, together with Supplementary Eq. ([S67](#)), allows us to write the normalization condition as

$$P_{r,2} + P_{f,2} + P_{u,2} + P_{a,2} = 1, \tag{S74}$$

and the VACF now is

$$\begin{aligned} &P_{f,1}P_{f,2} + P_{u,1}P_{f,2} + P_{r,1}P_{f,2} + P_{a,1}P_{f,2} - \\ &(P_{f,1}P_{r,2} + P_{u,1}P_{r,2} + P_{r,1}P_{r,2} + P_{a_1}P_{r,2}) = \\ &\left(P_{f,2} - P_{r,2} \right) \left(P_{f,1} + P_{u,1} + P_{r,1} + P_{a,1} \right) = g(2), \end{aligned}$$

which, by using Supplementary Eq. ([S67](#)), is simplified to

$$P_{f,2} - P_{r,2} = g(2). \tag{S75}$$

We see that Supplementary Eqs. ([S67](#)) and ([S74](#)), and ([S68](#)) and ([S75](#)) are practically identical. Therefore, by making the same assumptions, we arrive at the following expression for the probabilities at $k = 2$:

$$P_{i_2,2} = \frac{1 + 2 \left[\vec{c}_{i_0} \cdot \vec{c}_{i_2} \right] g(2)}{4}. \tag{S76}$$

Any k If we continue making the assumptions we have done until now, the probabilities can be generalized in a straightforward way as

$$P_{i_k,k} = \frac{1 + 2\left[\vec{c}_{i_0} \cdot \vec{c}_{i_k}\right] g(k)}{4}. \tag{S77}$$

Three dimensions: cubic lattice

Time step $k = 1$ In this case we have six lattice directions given by $\vec{c}_1 = (1,0,0)$, $\vec{c}_2 = (0,1,0)$, $\vec{c}_3 = (0,0,1)$, $\vec{c}_4 = (-1,0,0)$, $\vec{c}_5 = (0,-1,0)$, and $\vec{c}_6 = (0,0,-1)$. We denote the lattice direction parallel to the initial orientation with the subindex f, the contrary direction by r and the rest by u, a, d, and s. The normalization condition is

$$P_{f,1} + P_{u,1} + P_{r,1} + P_{d,1} + P_{a,1} + P_{s,1} = 1. \tag{S78}$$

The VACF is given by

$$P_{f,1} - P_{r,1} = g(1). \tag{S79}$$

Similarly as in the case of the square lattice, we impose the following condition which allows deriving the reorientation probabilities:

$$P_{u,1} = P_{a,1} = P_{s,1} = P_{d,1} = \frac{1}{6}, \tag{S80}$$

which enables us to simplify the normalization condition in the following way:

$$P_{f,1} + P_{r,1} = \frac{1}{3}. \tag{S81}$$

Using Supplementary Eq. (S81) to substitute $P_{r,1}$ into Supplementary Eq. (S79) we obtain

$$P_{f,1} - (\frac{1}{3} - P_{f,1}) = 2P_{f,1} - \frac{1}{3} = g(1),$$

which, after rearranging terms yields the probability

$$P_{f_1} = \frac{1 + 3g(1)}{6}. \tag{S82}$$

Now, using Supplementary Eq. (S81) we can obtain the remaining probability

$$P_{r,1} = \frac{1 - 3g(1)}{6}. \tag{S83}$$

Examining Supplementary Eqs. (S80), (S82) and (S83) we arrive at the general expression

$$P_{i_1,1} = \frac{1 + 3\left[\vec{c}_{i_0} \cdot \vec{c}_{i_1}\right] g(1)}{6}. \tag{S84}$$

Any k As done before we can continue the process for further times and, making the same assumptions, we arrive at an equation as Supplementary Eq. (S84) for any time k:

$$P_{i_k,k} = \frac{1 + 3\left[\vec{c}_{i_0} \cdot \vec{c}_{i_k}\right] g(k)}{6}. \tag{S85}$$

Any dimension, any lattice geometry, any time

Now that probabilities were derived for several dimensions, geometries, and times, we can see from Supplementary Eqs. (S50), (S66), (S77) and (S85) that the general form of the probabilities is given by

$$P_{i_k,k} = \frac{1 + d\left[\vec{c}_{i_0} \cdot \vec{c}_{i_k}\right] g(k)}{b}, \tag{S86}$$

where d is the spatial dimension and b is the number of nearest neighbors.

MSD of the piecewise process

We will now calculate the MSD using probabilities such that the VACF is a power-law decaying piecewise function defined as

$$g(t) = \begin{cases} 1 & t \leq t^\star \\ C_0 \left(\frac{\Delta}{t}\right)^\phi & t > t^\star \end{cases}, \tag{S87}$$

where $t^\star$ is such that $C_0 \left(\frac{\Delta}{t^\star}\right)^\phi = 1$. It is straightforward to see that the probabilities which define such a VACF obey

$$P_{i_k,k} = \begin{cases} \delta_{i,i_0} & k \leq \omega \\ \frac{1+d[\vec{c}_{i_0}\cdot\vec{c}_{i_k}]g(k)}{b} & k > \omega \end{cases} \tag{S88}$$

where i_0 is the index of the velocity channel the particle started in and ω is such that $t^\star = \omega\tau$. It is easy to see that in the first ω time steps the MSD is defined by

$$\langle r^2 \rangle (k) = k^2 \varepsilon^2,$$

or, using the definition of the particle speed and taking the limit $\tau \to 0$:

$$\langle r^2 \rangle (t) = (vt)^2. \tag{S89}$$

We will now calculate the MSD for time steps greater than ω. The calculation will be made for a 1D lattice, but the results are identical for any dimension and lattice geometry. To ease notation, we will omit any subindices refering to time steps $k \leq \omega$, as we know that, given Supplementary Eq. (S88), only those trajectories where the first ω orientations of the particle are identical to the original orientation of the particle have non-zero probabilities.

$\omega+1$ At the first time step after ω time steps have elapsed, we find that the MSD is given by

$$\langle r^2 \rangle (\omega+1) = r_f^2 P_{f,\omega+1} + r_r^2 P_{r,\omega+1},$$

where the displacements are $r_f^2 = (\omega+1)^2\varepsilon^2$ and $r_r^2 = (\omega-1)^2\varepsilon^2$. Using Supplementary Eq. (S88) and substituting the square displacements we obtain

$$\langle r^2 \rangle (\omega+1) = (\omega+1)^2\varepsilon^2 \left[\frac{1+g(\omega+1)}{2}\right] + (\omega-1)^2\varepsilon^2 \left[\frac{1-g(\omega+1)}{2}\right]$$

$$= \frac{\varepsilon^2}{2} \left\{(\omega^2+2\omega+1)[1+g(\omega+1)] + (\omega^2-2\omega+1)[1-g(\omega+1)]\right\}$$

$$= \frac{\varepsilon^2}{2} \left\{2\omega^2+2+2\omega[1+g(\omega+1)-1+g(\omega+1)]\right\}$$

which reduces to

$$\langle r^2 \rangle (\omega+1) = \varepsilon^2[\omega^2+1+2\omega g(\omega+1)]. \tag{S90}$$

$\omega+2$ Now, the MSD can be expanded in the following way:

$$\langle r^2 \rangle (\omega+2) = \varepsilon^2 \left\{(\omega+2)^2 \left[\frac{1+g(\omega+1)}{2}\right]\left[\frac{1+g(\omega+2)}{2}\right]\right.$$

$$+ \omega^2 \left[\frac{1+g(\omega+1)}{2}\right]\left[\frac{1-g(\omega+2)}{2}\right]$$

$$+ \omega^2 \left[\frac{1-g(\omega+1)}{2}\right]\left[\frac{1+g(\omega+2)}{2}\right]$$

$$\left. + (\omega-2)^2 \left[\frac{1-g(\omega+1)}{2}\right]\left[\frac{1-g(\omega+2)}{2}\right]\right\}$$

$$\frac{\varepsilon^2}{4} \left\{(\omega^2+4\omega+4)[1+g(\omega+1)+g(\omega+2)+g(\omega+1)g(\omega+2)]\right.$$

$$+ \omega^2[1+g(\omega+1)-g(\omega+2)-g(\omega+1)g(\omega+2)]$$

$$+ \omega^2[1-g(\omega+1)+g(\omega+2)-g(\omega+1)g(\omega+2)]$$

$$\left. + (\omega^2-4\omega+4)[1-g(\omega+1)-g(\omega+2)+g(\omega+1)g(\omega+2)]\right\},$$

which reduces to:

$$\langle r^2 \rangle (\omega+2) = \varepsilon^2 \left\{\omega^2+2+2g(\omega+1)g(\omega+2)+2\omega[g(\omega+1)+g(\omega+2)]\right\}. \tag{S91}$$

Any k   We can proceed for further k and will arrive at the following expression for any $k > \omega$:

$$
\langle r^2 \rangle (\omega + k) = \varepsilon^2 \left[\omega^2 + k + 2 \sum_{i=1}^{k} \sum_{j=i}^{k} g(\omega + i) g(\omega + j) \right.
$$
$$
\left. - 2 \sum_{i=1}^{k} g^2(\omega + 1) + 2\omega \sum_{i=1}^{k} g(\omega + i) \right]
\tag{S92}
$$

which by using the definition of the diffusion coefficient and particle speed can be converted to

$$
\langle r^2 \rangle (\omega + k) = 2dD \left[k\tau - 2 \sum_{i=1}^{k} g^2(\omega + i)\tau \right] + v^2 \left[2 \sum_{i=1}^{k} \sum_{j=i}^{k} g(\omega + i) g(\omega + j) \tau^2 \right.
$$
$$
\left. + (\omega\tau)^2 + 2\omega\tau \sum_{i=1}^{k} g(\omega + i)\tau \right]
\tag{S93}
$$

which in the limit $\tau \to 0$ is

$$
\langle r^2 \rangle (t) = 2dD \left[(t - t^\star) - 2 \int_{t^\star}^{t} g^2(\tau)\mathrm{d}\tau \right]
$$
$$
+ v^2 \left[2 \int_{t^\star}^{t} \int_{\tau}^{t} g(\tau) g(k)\mathrm{d}k\mathrm{d}\tau + t^{\star 2} + 2t^\star \int_{t^\star}^{t} g(\tau)\mathrm{d}\tau \right].
\tag{S93}
$$

Combining Supplementary Eqs. (S89) and (S93) we obtain the MSD of a particle with a piecewise power-law decaying VACF:

$$
\langle r^2 \rangle (t) = \begin{cases} (vt)^2 & t \leq t^\star, \\ 2dD \left[(t - t^\star) - 2 \int_{t^\star}^{t} g^2(\tau)\mathrm{d}\tau \right] \\ + v^2 \left[2 \int_{t^\star}^{t} \int_{\tau}^{t} g(\tau) g(k)\mathrm{d}k\mathrm{d}\tau + t^{\star 2} + 2t^\star \int_{t^\star}^{t} g(\tau)\mathrm{d}\tau \right] & t > t^\star. \end{cases}
\tag{S94}
$$

Generalized time-correlated random walk: rule derivation

We maximize the caliber

$$
\mathscr{C} = -\sum_{\Gamma} P_{\Gamma} \ln P_{\Gamma},
\tag{S95}
$$

subject to observing a certain VACF, which translates into the Lagrange multiplier problem

$$
\tilde{\mathscr{C}} [P_{\Gamma}] = -\sum_{\Gamma} P_{\Gamma} \ln P_{\Gamma} + \sum_{i=1}^{k} \beta(i) \left[\sum_{\Gamma} P_{\Gamma} \left(\vec{c}_{n_0} \cdot \vec{c}_{n_i} \right) - g(i) \right] + \lambda \left(\sum_{\Gamma} P_{\Gamma} - 1 \right),
\tag{S96}
$$

This yields the trajectory probabilites

$$
P_{\Gamma} = \frac{1}{Z} \exp \left[\sum_{i=1}^{k} \beta(i) \left(\vec{c}_{n_0} \cdot \vec{c}_{n_i} \right) \right],
\tag{S97}
$$

where $Z = \exp(1 - \lambda)$ is called the dynamical partition function which, by optimizing the functional with respect to λ (i.e., $\frac{\partial \tilde{\mathscr{C}}}{\partial \lambda} = 0$), is given by $Z = \sum_{\Gamma} \exp \left[\sum_{i=1}^{k} \beta(i) \left(\vec{c}_{n_0} \cdot \vec{c}_{n_i} \right) \right]$. Optimizing with respect to $\beta(i)$ yields our original constraint

$$
g(k) = \sum_{\Gamma} P_{\Gamma} \left(\vec{c}_{n_0} \cdot \vec{c}_{n_i} \right).
\tag{S98}
$$

Solving for $\beta(i)$ using Supplementary Eqs. (S97) and (S98) can be quite challenging, so we expand P_{Γ} in a Taylor series around $\beta(i) = 0$, which reduces to $g(k) \approx \sum_{\Gamma} \frac{1}{Z} \left[1 + \sum_{i=0}^{k} \beta(i) \left(\vec{c}_{n_0} \cdot \vec{c}_{n_i} \right) \right] \left(\vec{c}_{n_0} \cdot \vec{c}_{n_k} \right)$, where the dynamical partition function is

simplified as

$$Z \approx \sum_{\Gamma} \left[1 + \sum_{i=0}^{k} \beta(i) \left(\vec{c}_{n_0} \cdot \vec{c}_{n_i} \right) \right] = b^k + \sum_{\Gamma} \sum_{i=1}^{k} \beta(i) \left(\vec{c}_{n_0} \cdot \vec{c}_{n_i} \right) = b^k + \sum_{i=1}^{k} \beta(i) \sum_{\Gamma} \left(\vec{c}_{n_0} \cdot \vec{c}_{n_i} \right)$$

$$= b^k + \sum_{i=1}^{k} \beta(i) \sum_{\Gamma} \cos \theta_i = b^k,$$

where b is the number of lattice directions and θ_i is the angle between the original particle orientation and the particle orientation at time step i. $\sum_{\Gamma} \cos \theta_i = 0$ because the lattice directions and hence the possible particle orientations are symmetrically and homogeneously distributed. Substituting Z, using the same notation as previously, employing trigonometric identities, and denoting the spatial dimension by d, we proceed with the calculation:

$$g(k) \approx b^{-k} \left[\sum_{\Gamma} \cos \theta_k + \sum_{\Gamma} \sum_{i=1}^{k} \beta(i) \cos \theta_i \cos \theta_k \right] = b^{-k} \sum_{\Gamma} \sum_{i=1}^{k} \beta(i) \cos \theta_i \cos \theta_k = b^{-k} \sum_{i=1}^{k} \sum_{\Gamma} \beta(i) \cos \theta_i \cos \theta_k$$

$$= b^{-k} \left[\beta(1) \sum_{\Gamma} \cos \theta_1 \cos \theta_k + \beta(2) \sum_{\Gamma} \cos \theta_2 \cos \theta_k + \cdots + \beta(k-1) \sum_{\Gamma} \cos \theta_{k-1} \cos \theta_k + \beta(k) \sum_{\Gamma} \cos^2 \theta_k \right]$$

$$= b^{-k} \beta(k) \sum_{\Gamma} \cos^2 \theta_k = \frac{\beta(k)}{2b^k} \sum_{\Gamma} \left[1 + \cos(2\theta_k) \right] = \frac{\beta(k)}{2b^k} \left[b^k + \sum_{\Gamma} \cos(2\theta_k) \right] = \frac{\beta(k)}{2b^k} \left[b^k + \frac{b^k}{d}(2-d) \right]$$

$$= \frac{\beta(k)}{2} \left[1 + \frac{2-d}{d} \right] = \frac{\beta(k)}{d},$$

which determines the Lagrange multiplier

$$\beta(k) = d g(k). \tag{S99}$$

So the generalized probabilities are finally

$$P_\Gamma = \frac{1}{Z} \exp \left[\sum_{i=1}^{k} d g(i) \left(\vec{c}_{n_0} \cdot \vec{c}_{n_i} \right) \right], \tag{S100}$$

which is the probability for the whole trajectory. Due to the exponential form of this probability, we can decompose the trajectory probability into reorientation probabilities:

$$P_\Gamma = \prod_{i=1}^{k} P_{x,k}, \tag{S101}$$

given by:

$$P_{n_k,k} = \frac{1}{z} \exp \left[d g(k) \left(\vec{c}_{n_0} \cdot \vec{c}_{n_k} \right) \right], \tag{S102}$$

where z is the normalization constant for the reorientation probability.

Generalized time-correlated random walk: VACF decay analysis

Eq. (28) is at first sight, different from a simple power law decay. We now assess how similar Eq. (28) is to a simple power law decay for intermediate times. The easiest and most insightful way to achieve this is to expand both Eq. (28) and a generic power law in a Taylor series, and to compare the Taylor coefficients. We expand around $t = \Delta$. We will denote Eq. (28) by $C(t)$. The power law function has the following form:

$$G(t) = G_1 \left(\frac{\Delta}{t} \right)^\gamma, \tag{S103}$$

where the constants G_1 and γ are unspecified. First, we calculate the first two derivatives of $C(t)$:

$$\frac{dC(t)}{dt} = -\phi C_0 \Delta^\phi t^{-\phi-1} \operatorname{sech}^2\left[C_0\left(\frac{\Delta}{t}\right)^\phi\right] \tag{S104a}$$

$$\frac{d^2C(t)}{dt^2} = \phi C_0 \Delta^\phi t^{-2\phi-2} \operatorname{sech}^2\left[C_0\left(\frac{\Delta}{t}\right)^\phi\right] \left\{(\phi+1)t^\phi - 2\phi C_0 \Delta^\phi \tanh\left[C_0\left(\frac{\Delta}{t}\right)^\phi\right]\right\}, \tag{S104b}$$

with which we can calculate its Taylor series up to the second order term:

$$C(t) = \tanh(C_0) - \frac{\phi C_0}{\Delta}\operatorname{sech}^2(C_0)(t-\Delta)+$$
$$\frac{1}{2!}\frac{\phi C_0 \operatorname{sech}^2(C_0)}{\Delta^2}\{1+\phi[1-2C_0\tanh(C_0)]\}(t-\Delta)^2 + \mathcal{O}(t^3). \tag{S105}$$

We now proceed in the same way with the power law:

$$\frac{dG(t)}{dt} = -G_1\Delta^\gamma \gamma t^{-\gamma-1} \tag{S106a}$$

$$\frac{d^2G(t)}{dt^2} = G_1\Delta^\gamma \gamma(1+\gamma)t^{-\gamma-2} \tag{S106b}$$

and expand in a Taylor series around $t=\Delta$:

$$G(t) = G_1 - \frac{G_1\gamma}{\Delta}(t-\Delta) + \frac{1}{2!}\frac{G_1\gamma}{\Delta^2}(1+\gamma)(t-\Delta)^2 + \mathcal{O}(t^3). \tag{S107}$$

To determine G_1 and γ we equate the zeroth and first order terms of Supplementary Eqs. (S105) and (S107), which yields

$$G_1 = \tanh(C_0) \tag{S108a}$$

$$\gamma = \phi C_0 \frac{\operatorname{sech}^2(C_0)}{\tanh(C_0)}, \tag{S108b}$$

so that the Taylor series expansion is determined by

$$G(t) = \tanh(C_0) - \frac{\phi C_0}{\Delta}\operatorname{sech}^2(C_0)(t-\Delta)+$$
$$\frac{1}{2!}\frac{\phi C_0 \operatorname{sech}^2(C_0)}{\Delta^2}\left[1+\phi C_0\frac{\operatorname{sech}^2(C_0)}{\tanh(C_0)}\right](t-\Delta)^2 + \mathcal{O}(t^3). \tag{S109}$$

To estimate the similarity between both decays, we calculate the difference between Supplementary Eqs. (S105) and (S109) up to second order terms:

$$C(t) - G(t) \approx \frac{1}{2!}\frac{\phi C_0 \operatorname{sech}^2(C_0)}{\Delta^2}\{1+\phi[1-2C_0\tanh(C_0)]\}(t-\Delta)^2-$$
$$\frac{1}{2!}\frac{\phi C_0 \operatorname{sech}^2(C_0)}{\Delta^2}\left[1+\phi C_0\frac{\operatorname{sech}^2(C_0)}{\tanh(C_0)}\right](t-\Delta)^2 = (t-\Delta)^2\cdot$$
$$\frac{1}{2!}\frac{\phi C_0 \operatorname{sech}^2(C_0)}{\Delta^2}\left\{1+\phi[1-2C_0\tanh(C_0)]-1-\phi C_0\frac{\operatorname{sech}^2(C_0)}{\tanh(C_0)}\right\}$$
$$= (t-\Delta)^2\frac{1}{2!}\frac{\phi^2 C_0 \operatorname{sech}^2(C_0)}{\Delta^2}\left[1-2C_0\tanh(C_0)-C_0\frac{\operatorname{sech}^2(C_0)}{\tanh(C_0)}\right]$$
$$= (t-\Delta)^2\frac{1}{2!}\frac{\phi^2 C_0 \operatorname{sech}^2(C_0)}{\Delta^2}\left\{1-C_0\left[\frac{2\sinh(C_0)}{\cosh(C_0)}-\frac{1}{\cosh(C_0)\sinh(C_0)}\right]\right\}$$
$$= (t-\Delta)^2\frac{1}{2!}\frac{\phi^2 C_0 \operatorname{sech}^2(C_0)}{\Delta^2}\left\{1-C_0\left[\frac{2\sinh^2(C_0)+1}{\cosh(C_0)\sinh(C_0)}\right]\right\}$$

which, after using hyperbolic identities, can be simplified to

$$C(t) - G(t) \approx (t-\Delta)^2\frac{1}{2!}\frac{\phi^2 C_0 \operatorname{sech}^2(C_0)}{\Delta^2}[1-2C_0\coth(2C_0)] \propto \left(\frac{\phi}{\Delta}\right)^2. \tag{S110}$$

LGCA simulations

Persistent random walk

Simulations were performed with only one particle with the reorientation probability given by Eq. (10) whose displacement and orientation were tracked at every time step. The lattice spacing was set to $\varepsilon = 0.25$, and the time step to $\tau = 0.015625$. The total simulation consisted of 100 time steps. The sensitivity (related to the internal force required for reorientation) was varied from $\beta = 3$ to $\beta = 5$. Simulations were repeated 1000 times for each sensitivity in order to obtain statistically relevant results. Simulation results for low and high sensitivities are shown in Fig. 5i.

As expected, correlations die off more slowly with increasing sensitivity. On the other hand, the MSD quickly starts behaving linearly, except for times close to zero, where it behaves almost ballistically. The region where displacement is almost ballistic increases with increasing sensitivity.

Additionally, we observe that the derived continuous time expressions agree perfectly with the discrete LGCA simulations.

Time-correlated random walk

Simulations were performed with only one particle with the reorientation probability given by Eq. (18). The lattice spacing was set to $\varepsilon = 0.25$, and the time step to $\tau = 0.015625$. The constant C_0 was set to 0.5, and the crossover time was equal to the time step length, $\Delta = \tau = 0.015625$. The total simulation consisted of 1000 time steps. Three different exponents were evaluated: $\phi = 0.1$, $\phi = 1$, and $\phi = 9$. Simulations were repeated 1000 times for each exponent, in order to obtain statistically relevant results. Simulation results for small and large exponents are shown in Fig. 5ii, as well as a plot of Eqs. (16) and (22) (integrated with MATLAB). We can see that Eqs. (16) and (22) match the simulation data perfectly. We also observe that for low values of the exponent ϕ the particle moves superdiffusively while for large values the particle diffuses normally.

Generalized time-correlated random walk

Simulations were performed with only one particle with probabilities given by Eq. (18). The lattice spacing was set to $\varepsilon = 0.25$ and the time step to $\tau = 0.015625$. The constant C_0 was set to 0.5 and the crossover time was equal to the time step length, $\Delta = \tau = 0.015625$. The total simulation consisted of 100 time steps. Two different exponents were evaluated, $\phi = 0.1$ and $\phi = 1$. Simulations were repeated 1000 times for each exponent, in order to obtain statistically relevant results. Simulation results for small, and large exponents are shown in Fig. 5iii as well as a plot of Eqs. (28) and (29) (integrated with MATLAB). We can see that Eq. (28) and (29) match the simulation data perfectly. Comparing Figs. 5ii and 5iii, it is evident that the VACF in both cases is quite similar, as expected given the small value of Δ used in these simulations. We also observe that for low values of the exponent ϕ the particle moves superdiffusively while for large values the particle diffuses normally.