3D Spatial Exploration by E. coli Echoes Motor Temporal Variability

Nuris Figueroa-Morales,1,2 Rodrigo Soto,3 Gaspard Junot,1 Thierry Darnige,1 Carine Douarche,4 Vincent A. Martinez,5 Anke Lindner,1 and Éric Clément1,*

1PMMH, UMR 7636 CNRS, ESPCI-Paris, PSL Research University, Sorbonne Université, Université de Paris, 7-9 quai Saint-Bernard, 75005 Paris, France
2Department of Biomedical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802, USA
3Departamento de Física, FCFM, Universidad de Chile, Santiago 8370448, Chile
4Université Paris-Saclay, CNRS, FAST, 91405 Orsay, France
5SUPA and School of Physics & Astronomy, The University of Edinburgh, Edinburgh EH9 3FD, United Kingdom

(Received 6 January 2020; accepted 26 February 2020; published 6 April 2020)

Unraveling bacterial strategies for spatial exploration is crucial for understanding the complexity in the organization of life. Bacterial motility determines the spatiotemporal structure of microbial and controls infection spreading and the microbiota organization in guts or in soils. Most theoretical approaches for modeling bacterial transport rely on their run-and-tumble motion. For Escherichia coli, the run-time distribution is reported to follow a Poisson process with a single characteristic time related to the rotational switching of the flagellar motors. However, direct measurements on flagellar motors show heavy-tailed distributions of rotation times stemming from the intrinsic noise in the chemotactic mechanism. Currently, there is no direct experimental evidence that the stochasticity in the chemotactic machinery affects the macroscopic motility of bacteria. In stark contrast with the accepted vision of run and tumble, here we report a large behavioral variability of wild-type E. coli, revealed in their three-dimensional trajectories. At short observation times, a large distribution of run times is measured on a population and attributed to the slow fluctuations of a signaling protein triggering the flagellar motor reversal. Over long times, individual bacteria undergo significant changes in motility. We demonstrate that such a large distribution of run times introduces measurement biases in most practical situations. Our results reconcile the notorious conundrum between run-time observations and motor-switching statistics. We finally propose that statistical modeling of transport properties, currently undertaken in the emerging framework of active matter studies, should be reconsidered under the scope of this large variability of motility features.

DOI: 10.1103/PhysRevX.10.021004 Subject Areas: Biological Physics, Interdisciplinary Physics, Soft Matter

1. INTRODUCTION

The run-and-tumble (R&T) strategy developed by bacteria for exploring their environment is a cornerstone of quantitative modeling of bacterial transport. In this paradigm, bacteria swim straight during a run time, undergo a reorientation process during a tumbling time, and pursue thereafter the next run in a different direction. The now standard vision of the R&T strategy was established in the 1970s for swimming Escherichia coli by Berg and Brown [1,2], based on 3D trajectories obtained via a Lagrangian tracking technique. They proposed that an adapted bacterium would perform, over long times, an isotropic random walk composed of the run-and-tumble phases, both distributed in time as a Poisson process [1–5]. For quantitative analysis, the run-time and tumble-time distributions are often taken as Poisson processes with typical values \( \hat{\tau}_{\text{run}} \sim 1 \) s and \( \hat{\tau}_{\text{tumble}} \sim 1/10 \) s [2,6]. These values change in the presence of chemical gradients, leading to a biased random walk known as chemotaxis.

Alongside the relevance of this result in the context of biology, medicine, or ecology, fluids laden with motile bacteria have become an epitome for active matter, where the organization of active particles recently led scientists to revisit many concepts of out-of-equilibrium statistical physics [7–10]. Suspensions of motile bacteria are systems of choice for these studies [11], and many original phenomena such as anti-Fick’s law migration [12], collective motion [13], viscosity reduction [14–16], enhanced

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.
diffusion [7], or motion rectification [17–20] have been discovered. Most recent theoretical studies on active matter, aimed at understanding the emergence of collective motion or other macroscopic transport processes in bacterial fluids, assume uncorrelated orientational noise, which is the direct consequence of the Poisson character of the R&T process [9].

The simple approach of introducing a Poisson distribution for the run times, although useful for simple qualitative interpretations, is not fully consistent with a growing number of measurements performed on the individual rotary motors [21–25] driving the helix-shaped flagella. For *E. coli*, the forward (run) motion is associated with the counterclockwise (CCW) rotation of the motors, and the tumbles take place when the motors rotate clockwise (CW). The CCW to CW transition is regulated by an internal biochemical process associated with the phosphorylation of the CheY protein.

In a seminal work, Korobkova et al. [21] brought evidence for a heavy tail distribution for the duration of CCW rotations. Importantly, this highlights possible coupling between the stochastic fluctuations in the chemotactic biochemical network and the emergent bacterial motility. Its consequences could affect the macroscopic organization of bacterial populations, chemotactic response to chemical heterogeneity, and genetic and epigenetic feedback of bacterial populations to environmental constraints.

Its potential importance in the context of active matter studies remains overlooked. For multiflagellated bacteria, the correspondence between switching statistics, motor synchronization, flagellar bundling and unbundling dynamics, and, finally, large-scale exploration properties remains unclear. Recently, indirect experimental evidence suggested that the macroscopic motility of free-swimming bacteria is sensitive to the stochasticity borne by the chemotactic biological circuit [26]. Here, we give direct evidence of this sensitivity.

Conceptually, our analysis starts from the extreme sensitivity of the rotational CCW → CW switching to the abundance of the phosphorylated protein CheY-P in the cell. This picture induces a timescale separation, since, at short times, the alternation of CCW and CW rotations keeps a memory of a quasifixed level of CheY-P. This memory is erased at longer times, and we thus expect very different run times and motility features at the macroscopic level.

For the first time, we link the individual motor rotation statistics to the global motility features that we observe in a large number of 3D trajectories of wild-type *E. coli* bacteria. At short observation times, the time persistence of the swimming orientations displays an exponential decay as classically admitted, but with a large distribution of characteristic times within a population of monoclonal bacteria. However, when tracking the cells individually over several tenths of minutes, we identify for each cell a large behavioral variability. The motility data are quantitatively analyzed through a simple model initially proposed by Tu and Grinstein [27] involving the fluctuations of CheY-P triggering the tumbling events. The model is here adapted to render the spatial exploration process. It now explains the occurrence of a large behavioral variability of swimming direction and also why, at short observation times, a large distribution of these is expected over a population. The central outcome of this model is that the persistence time durations naturally follow a log-normal distribution, instead of a standard Poisson distribution. Importantly, we identify a source of measurement bias introduced in most practical situations that is a consequence of such a large distribution of run times. Finally, we discuss the consequences of measuring averaged quantities over a population displaying a large distribution of motility features. This source of measurement bias is relevant in the general framework of experiments on statistical physics of active matter.

II. VARIABILITY OF BACTERIAL MOTILITY IN A POPULATION

To characterize the bacterial motility, we build an automated tracking device suited to follow fluorescent objects and record their 3D trajectories. A swimming bacterium is kept automatically in the center of the visualization field and at the focus of an inverted microscope by a visualization feedback loop acting horizontally on a mechanical stage and vertically on a piezo stage. The method is fully detailed in Ref. [28] by Darnige et al. (see also Sec. VI) and was recently used to investigate the swimming of bacteria in a Poiseuille flow [29].

We first monitor more than 100 swimming *E. coli* from different strains (see Sec. VI) in homogeneous diluted suspensions (concentration $\sim10^3$ bact.mL$^{-1}$) confined between two horizontal glass slides, $250 \mu$m apart. Figure 1(a) shows two typical trajectories from the same batch of monoclonal wild-type *E. coli*. We center our analysis on pieces of tracks exploring the bulk [Fig. 1(b)], i.e., in a measurement region located $10 \mu$m above the surface and of maximum height $H = 130 \mu$m. For this series of experiments, the duration of a track is at minimum 8 s. We name these experiments 1.

The bacterial velocities $\vec{V}(t)$ at each point of the trajectories are obtained by fitting the sequence of coordinates, along *X*, *Y*, and *Z* independently, over segments spanning 0.1 s, using a second-order polynomial. The first derivative of the polynomial evaluated at the center of the segment provides the velocity component. Figure 2 shows an example of a 3D trajectory and its velocity. Typically, the velocity curves for each track are irregular [Fig. 2(b)]. For a single track, the velocity distribution [Fig. 2(c)] shows a peak corresponding to the run phase and a low-velocity tail that might correspond to tumbling events. For the wild-type strain RP437 in a motility buffer, the average of the peak
values for \( V = |\vec{V}(t)| \) over the different tracks is \( \langle V \rangle = 27 \pm 6 \, \mu m/s \).

Standard analysis to extract run-time distributions relies on the identification of tumbling events, usually done by detecting velocity drops and/or abrupt changes in the swimming direction [2,6,30]. However, as shown in Figs. 2(a) and 2(b), abrupt direction changes can take place without a representative velocity decrease, and velocity drops are sometimes not associated with reorientation. This observation is consistent with results from Refs. [30,31]. Moreover, by directly monitoring the flagellar dynamics, Turner et al. [31] identify partial flagellar debundling inducing weak velocity drops and directional changes. We find that, without a direct observation of the flagella, run-and-tumble detection requires the choice of arbitrary criteria. We demonstrate this arbitrariness in the Appendix A.

Here, in order to characterize the motility features, we do not seek to explicitly identify the tumbling events. Instead, we use the orientation correlation function \( C(\Delta t) \) as a direct measurement of the swimming direction persistence. The director vectors pointing along the track are determined as \( \vec{p} = \bar{V}(t)/V(t) \) for each track. For each trajectory, we compute \( C(\Delta t) = \langle \dot{\vec{p}}(t) \cdot \dot{\vec{p}}(t+\Delta t) \rangle = \langle \cos[\theta(\Delta t)] \rangle \), where \( \theta \) is the angle between swimming directors separated by a time lag \( \Delta t \) [Fig. 1(b)]. The brackets denote an average over a time window sliding along the track. To ensure good statistics, the maximum lag time \( \Delta t \) is chosen as one-tenth of the total track duration. The orientation correlation reflects the R&T statistics but advantageously does not require an ad hoc criterion. In Fig. 3(a), 30 orientation correlation functions obtained from separate tracks of different bacteria (RP437 wild type in M9G) are displayed as a function of \( \Delta t \).

From the classical picture of an exponential distribution of run times, the orientation correlation function is expected to decay exponentially with a typical decay time of \( \tau_p \), defining the persistence time of the trajectory. For a characteristic run time of \( \tau_{run} = 1 \, s \) and a distribution of reorientation angles of mean value \( \theta_m = 51^\circ \) [1], one finds \( \tau_p = (\tau_{run}/(1 - \langle \cos(\theta) \rangle)) = 1.5 \, s \) [32]. Recently, a slight dependence of this angle on the swimming speed was demonstrated [33] but is neglected in our study. Taking into account rotational Brownian diffusion during the run phase also leads to an exponential decaying correlation function (see Appendix B). Its contribution represents a slight modification to \( \tau_p \) due to the much longer timescales of Brownian diffusion. The predicted correlation function is
The correlation functions are fitted with an exponential decay \( \exp(-\tau/\tau_p) \) to extract the persistence times \( \tau_p \). The dotted line corresponds to \( \tau_p = 1.5 \) s as expected from Ref. [1]. Inset: Correlation functions as a function of time. If the population displayed a large distribution of persistence times \( \tau_p \), we determine the persistence times \( \tau_p \) from the correlation function. Finally, we consider the strain CR20, a smooth swimmer that tumbles only very rarely. In this case, the time distribution is gathered around the average \( \tau_p = 25 \pm 10 \) s, which is close to the Brownian rotational diffusion constant \( \tau_p = \tau_B = 1/2D_B^p \), as expected.

Therefore, the wide distribution of persistence times for CR20 could arise from the bacterial size distribution. A possible origin of this dispersion on the measurement protocol is discussed in Sec. IV C.

**III. VARIABILITY OF INDIVIDUAL BACTERIAL MOTILITY OVER TIME**

The large diversity of trajectories here observed over short times in bacterial populations leads to the question of its origin. The diversity could arise from a phenotype multiplicity present in the monoclonal population [37,38], where each bacterium is characterized by a mean run time; alternatively, it could be due to temporal variability of the bacterial behavior, with mean run times varying over the course of time. To determine which scenario is taking place, we perform a second series of measurements, experiments II, where we follow individual bacteria over very long times (up to 20 min). In the new configuration, the top and bottom of the measurement chamber are within the observation range or the 3D tracker device. We follow individual bacteria as they alternate between the surfaces and the bulk, as sketched in Fig. 4(a). For the analysis, individual tracks are cut in pieces localized entirely in the bulk (10 \( \mu \)m away from the walls). For each piece, we extract the persistence time from the correlation function. Finally, for each bacterium, we obtain a list of persistence times as a function of time. If the population displayed a large distribution of fixed run-times, one would expect for each bacterium a sequence of persistence times narrowly distributed around a characteristic value, but this value would be different for different bacteria. Figure 4(b) carries a very different message. For each of the tracks tested, the persistence times span a range of the same magnitude as...
for the whole population using shorter tracking times (see Fig. 3).

Previous studies based on 3D Eulerian tracking techniques [33,39], i.e., on a fixed reference frame, or even the Lagrangian tracking technique [6] were limited to short observation times and, consequently, were not able to catch such slow fluctuations of the run time. The fact that for a given bacterium the sequence of persistence times is largely distributed confirms the importance of behavioral variability in the motility process. However, due to tracking time limitations imposed by the bleaching of the fluorescent signal, we are not able to test precisely to what extent the behavioral variability contains features which could vary from one bacterium to the other, stemming from inherent phenotype variations, as identified, for example, by Dufour et al. [40].

IV. MOTILITY AND MOTOR ROTATION STATISTICS

The presence of a behavioral variability, as identified earlier, raises the question of its biochemical origins. Previous results point toward a definite influence of a stochastic process in the chemotactic sensory circuit. At the end of the biochemical cascade, there is a phosphorylation of a CheY protein (CheY-P) promoting a switch in the motor rotation from the CCW state (run phase) to the CW state (tumbling phase). The most accepted picture rendering the CCW=CW transition is a two-state model initially proposed by Khan and Macnab [41], which considers the switching of the rotation direction CCW $\rightarrow$ CW (equivalently, CW $\rightarrow$ CCW) as an activated process regulated by the presence of CheY-P. The double-well Gibbs free energy associated with the transition CCW$\rightleftharpoons$CW depends in a very sensible way on the CheY-P ($Y$) concentration values near the motor, as shown by Cluzel, Surette, and Leibler [42]. This strong sensitivity leads naturally to behavioral variations, as slow fluctuations around the mean value can change the motility features from preferentially tumbling (high CheY-P) to preferentially running (low CheY-P). It also means that at short times the CheY-P level does not change significantly and motility features remain constant. Therefore, at a given moment, the motility features should be largely distributed in a population of bacteria bearing different CheY-P concentrations. This large distribution is in essence what is observed in our experiments in Figs. 3 and 4.

A. Quantitative description of the behavioral variability model

To rationalize and quantify our experimental findings, we adapt the simple but enlightening physical model proposed by Tu and Grinstein [27]. The behavioral variability (BV) model we present here quantifies the role of fluctuations of the phosphorylated protein CheY-P in the regulation of the motor-switching statistics. The key idea is that the observed typical switching time at a given moment depends on the instantaneous CheY-P concentration $|Y|/(t)$. Then, considering concentration fluctuations around a mean value ($\delta Y(t) = |Y|/(t) - |Y|_0$), one obtains a two-state model with a time-varying barrier describing the CCW $\rightarrow$ CW switching process. Tu and Grinstein model the $\delta Y$ fluctuations as an Ornstein-Uhlenbeck process with a memory (relaxation) time $T_Y$, hence yielding a Gaussian distribution for $\delta Y$ values. Note that $T_Y$ is considered to be larger than typical motor-switching times [see Fig. 5(a) for the relevant timescales].

For small fluctuations of concentration, the average switching time can be written as

$$\tau_s = \tau_0 e^{-\Delta_n \delta Y}.$$  

Here, $\delta X$ corresponds to the fluctuations in concentration normalized by the $\delta Y$ standard deviation $\sigma_{Y}$; $\tau_0$ is a typical switching time corresponding to the mean concentration $|Y|_0$ and $\Delta_n = \alpha (\sigma_Y/|Y|_0)$. The parameter $\alpha$ is positive [42] and measures the sensitivity of the switch to variations in $|Y|$, which means that higher concentrations of CheY-P lead to shorter run times. Note that, in principle, the two switching times describing CCW $\rightarrow$ CW (run times) or CW $\rightarrow$ CCW (tumbling times) could be modeled with corresponding parameters $\tau_n$ and $\Delta_n$. However, as the results from Korobkova et al. [21] show that, in contrast with run times, the distribution of tumble times is exponential, meaning that the equivalent of $\Delta_n$ for tumbles is
small. Hence, we consider the tumbling times as a Poissonian process, well described by a single timescale. Let us first consider the CCW → CW switching time distributions. Each observed time belongs to a Poisson distribution with a typical time $\tau_s$ set by the current CheY-P concentration $\frac{1}{2} \frac{Y}{C_1} (t)$ [see Eq. (1)]. As a consequence, the observed switching statistics for an individual bacterium, when observed over a time interval shorter than the memory time, should approximately appear as an effective Poisson process, which is indeed the case, as shown from the collapse of the rescaled orientation correlation functions onto a single exponential decay shown in Fig. 3(a). The model provides a second important outcome. A random choice of a bacterium in a population is like a random choice of $\delta X$, hence defining a typical switching time $\tau_s$ for this bacterium. A Gaussian distribution for $\delta X$, as assumed by the BV model, leads to a Gaussian distribution of $\ln(\tau_s)$ characterized by an average $\ln(\tau_0)$ and a standard deviation $\sigma_{\ln \tau_s} = \Delta_n$, yielding naturally a large log-normal distribution of $\tau_s$ provided the switch sensitivity $\alpha$ is large. Note that the power law distribution discussed by Tu and Grinstein [27] is obtained in the limit of very large $\Delta_n$ and not in contradiction with the above statement. As $\tau_s$ and $\tau_p$ are proportional, the distribution of $\ln(\tau_p)$ should also be Gaussian.

To illustrate this idea, a very long 3D trajectory is synthesized numerically using the switching statistics from the BV model. Figure 5(b) shows a 2D projection (see Sec. VI for technical details and Sec. IV C for the parameter values). The simulated trajectory contains very persistent (inset I) and very nonpersistent (inset II) parts. The colors represent the local values of $\delta X$ illustrating the direct influence of the slow variations of CheY-P concentration on the bacterial motility, hence explaining the observed behavioral variability.

**B. Memory time**

The evolution of persistence times $\tau_p$ along individual trajectories displays large variations. It is shown in Fig. 6(a)

![Figure 6](image.png)

**FIG. 6.** Determination of the memory time from experiments II. (a) Persistence times $\tau_p$ computed over pieces of span 20 s and shifted 5 s for two different bacteria. Gaps larger than 5 s between consecutive points correspond to lapses in which the bacteria are swimming close to surfaces. (b) Persistence times self-correlation function $C_p$ using pieces of 20 s. Points represent the average over the ensemble of bacteria. (c) Correlation time of the persistence times as a function of the lengths of the pieces. We extract the memory time to be $T_Y = 19.0 \pm 1.3$ s.
for the case of two different bacteria continuously tracked for 11 and 17 min. The values of $\tau_p$ for each track are extracted from intervals of span 20 s shifted 5 s along the trajectory. Gaps larger than 5 s between consecutive points correspond to lapses in which the bacterium is swimming close to a surface. Analyzing, for example, the bacterium of the blue longer trajectory, at time 300 s (5 min) it displays a persistence time close to 0.1 s, in contrast with a persistence time close to 5 s around time 1000 s ($\sim$17 min). This temporal variation of $\tau_p$ is considered in the framework of the BV model. The memory time $T_Y$ is then a central parameter of the BV model, as it provides a natural separation between short-time measurements and long-time measurements. Therefore, for a correct statistical interpretation of the results, $\tau_p$ values must be extracted from pieces of tracks not longer than the memory time $T_Y$.

We estimate the memory time $T_Y$ from the long-time tracking data in experiments II, using the following procedure. For each bacterial trajectory, we compute a sequence of $\tau_p$ using intervals of a specific duration. For each sequence of $\tau_p$, we compute the self-correlation function of persistence times, $C_p(t) = \langle \ln \tau_p(t + t') \ln \tau_p(t') \rangle - \langle \ln \tau_p \rangle^2$, where the average is done over $t'$. The average of $C_p$ over the ensemble of trajectories is fitted with an exponential, giving the correlation time [Fig. 6(b)]. With this procedure, we investigate different lengths of intervals [Fig. 6(c)], finding that the correlation times grow with the duration of the interval until saturation at the value of the memory time $T_Y \approx 19.0 \pm 1.3$ s.

C. Comparison with the model

The BV model depends on several parameters: the memory time $T_Y$, the mean switch time and sensitivity $\tau_0$ and $\Delta_n$, respectively, the rotational diffusion coefficient $D^R$, and the dimensionless rotational diffusion coefficient $D^R_{\text{eff}}$ used to model the reorientation during tumble (see Sec. VI for details). We determine $T_Y$ from the experiments, while the rest of the parameters are fitted using the following protocol. A long simulated trajectory is generated and cut in pieces of duration 20 s, similar to the analysis of the experimental tracks, and the persistence time is computed for each piece. We look for the values of the parameters that best agree with the experimental values of the first four moments of the distribution of $\ln \tau_p$. The result is $\tau_0 = 1.53$ s, $\Delta_n = 1.62$, $D^R = 0.025$ s$^{-1}$, and $D^R_{\text{eff}} = 3.86$. Note that the velocity does not appear in the fit, because we compare simulations and experiments using the persistence times, which depend only on the orientations.

Figure 7(a) compares the experimental distribution of $\ln \tau_p$ with the results from simulations using the optimal parameters. The agreement is very good, with two features that need discussion. First, in agreement with the BV model, the distributions are not exactly Gaussian but present a negative excess kurtosis. With 63% probability, the switch times are in the range $[\tau_0 e^{-\Delta_n}, \tau_0 e^{\Delta_n}] = [0.30 \text{ s}, 7.7 \text{ s}]$. Hence, there is no complete separation of timescale with $T_Y$. As a consequence, in each piece, $\delta X$ is not constant, and the measured and simulation distributions result from the mixture of different values of $\tau_p$. Note that shorter pieces would imply too few tumble events and would make it unreliable to use the orientation correlation function. A perfect log-normal distribution could be observed if there was a good separation of timescales, allowing a choice of intervals such that $\tau_0 \ll T_{\text{interval}} \ll T_Y$.

The second feature is the small peak at $\ln \tau_p \approx 3$ in the simulations. This peak corresponds to pieces of the
trajectory where no single tumble took place. The change in orientation is due only to rotational diffusion during a run. Because $τ_B = 1/2D_B^p \approx 20 \text{ s}$ is similar to $T_V$, no complete reorientation occurs in the interval, resulting in a distribution of $τ_p$ for nontumbling swimmers. In fact, the persistence times for the nontumbling bacteria [strain CR20 from experiments I, Fig. 3(b)] coincide with this peak. This feature should also be present in experiments, but, as discussed in Sec. II, $D_B^p$ depends strongly on the bacterial dimensions, which vary within the population. This dispersion of rotational diffusion and other imperfections blur this peak in the experiments, contrary to the simulations, where all swimmers are identical. Note that, despite of the diversity, the fitted value of $D_B^p$ matches closely the prediction made in Sec. II for ellipsoidal swimmers.

Since the pieces of trajectories are of a finite length, the orientation correlation function is not perfectly sampled, and, even for a constant switch time $τ_s$, the persistence times $τ_p$ obtained from an exponential fit would present some dispersion. To test whether the observed dispersion is due only to the data analysis protocol, we perform simulations with a Poisson model. For this test, we look for the best parameters to reproduce the first fourth moments of the distribution of $\ln τ_p$, setting $Δ_n = 0$. The result is $τ_p = 1.18 \text{ s}$, $D_B^p = 0.026 \text{ s}^{-1}$, and $D_{eff}^p = 1.61$. Figure 7(a) presents the resulting distribution, which is far from the experimental one. We conclude that a Poisson process cannot explain the broad distribution of persistence times observed experimentally.

Finally, for consistency, we return to the persistence times obtained in Fig. 3 from experiments I. In this experimental protocol, the trajectories are selected within a certain height ($10–140 \mu\text{m}$ from the surface) and longer than 8 s. The corresponding experimental distribution of $\ln τ_p$ for RP437 bacteria in all media [Fig. 7(b)] displays a clearly positive skewness, which differs strikingly from the experimental measures of Fig. 7(a), done using the same bacterial strain and confinement and a similar chemical environment. This difference originates from a measurement bias built into the analysis of Fig. 7(b) (and Fig. 3). The bias is a consequence of a preferential selection of long trajectories staying essentially in the $x$-$y$ plane, with limited bounds in the vertical direction. The skewness is enhanced by the broad distribution of run times, since very persistent swimmers will likely quit the measurement region in a very short time, hence privileging small persistence times. The curve “BV model” represents the distribution of persistence times from simulations of the BV model that fit the experiments in Fig. 7(a) (experiments I). When this same simulation is analyzed by taking pieces following strictly the experimental constraints, on both duration and vertical spatial exploration, the resulting distribution (“BV model experiment constraints”) compares very well and noticeably without any additional fitting parameter, to the experimental curve in Fig. 7(b) (experiments II).

V. CONCLUSIONS

We have shown that the 3D spatial exploration of an adapted E. coli reflects a behavioral variability that we associate with intrinsic noise in the chemotaxis pathway controlling the run-and-tumble sequence. Our results for free-swimming bacteria are consistent with models describing motor-switching dynamics based on tethered cell measurements. We identified a large log-normal distribution of persistent times stemming from the slow fluctuations of an internal variable accounting for the CheY-P concentration near the motors. In the context of many recent works on statistical physics of active matter, we suggest that this large variability should be included in the description of bacterial fluids. This variability is expected to influence the prediction of averaged quantities like diffusivity, viscosity, or any constitutive relations of macroscopic transport processes.

The broad distribution of run times is likely to introduce measurement biases in practical situations. Here, we reduce the bias by taking pieces of trajectories of equal length, not larger than the memory time. Mixing trajectories of different lengths can result in highly distorted distributions.

The large distribution of motility features has been related to the time bacteria spend close to surfaces. As an example, the existence of a large distribution of motor-switching statistics was found crucial to understand large-scale upstream bacterial contamination of narrow channels [26], where substantial transport occurs along surfaces [43–46].

We expect the chemotactic drift to be sensitive to the distribution of CheY-P concentrations, since a nonlocal spatiotemporal coupling will take place between chemical gradients and bacterial concentration. This sensitivity should be taken into account in future motility modeling. Finally, these findings may also impact quantitative modeling on how bacterial populations react to environmental changes, colonize space, swarm in a biofilm [47], or interact with other communities.

VI. MATERIALS AND METHODS

A. Bacterial strains and culture

We use the wild-type strains RP437 and AB1157 and a smooth swimmer mutant strain CR20 (ΔCheY) expressing yellow fluorescent protein (YFP) from a plasmid. Bacteria are grown overnight at 30 °C in M9G medium [M9 minimal medium supplemented with glucose (4 g/L), casamino acids (1 g/L), MgSO$_4$ (2 mM), and CaCl$_2$ (0.1 mM)] plus the corresponding antibiotics, up to optical density $= 0.5$ at 595 nm. Cells are then washed 3 times by centrifugation at 2000 g for 5 min and suspended in a motility buffer (10 mM potassium phosphate buffer $pH \sim 7.0$, 0.1 mM EDTA, 1 μM L-methionine, and 10 mM sodium L-lactate), supplemented with polyvinylpyrrolidone (PVP-360 kDa 0.002%) and, when indicated, with L-Serine (0.04 g/mL).
B. The 3D Lagrangian tracker

We develop a device for keeping individual microscopic objects—as swimming bacteria—in focus, as they move in microfluidic chambers [28]. The system is based on real-time image processing, determining the displacement of the stage to keep the chosen object at a fixed position in the observation frame. The z displacement of the stage is based on the refocusing of the fluorescent object that keeps the moving object in focus. The algorithm for z determination is designed for not being affected by photobleaching.

The instrument is mounted on an epifluorescent inverted microscope (Zeiss-Observer, Z1) with a high magnification objective (100 × /0.9 DIC Zeiss EC Epiplan-Neofluar), an x-y mechanically controllable stage with a z piezomover from Applied Scientific Instrumentation (ms-2000-flat-top-objective), and a digital camera ANDOR iXon 897 EMCCD. The device works nominally at 30 frames per second on a xyz), and a digital camera ANDOR iXon 897 EMCCD.

The tracking limitations come essentially from the z exploration range, restricted by the working distance of 150 μm of the objective. In the x-y plane, the spatial limitations are virtually nonexistent, since the stage displacement can be as long as 15 cm, which is much bigger than the typical sizes of the sample (a few millimeters). Details of the apparatus are given in Ref. [28], as well as an exhaustive explanation of a method for correcting the mechanical backlash typically affecting these systems and a discussion of the device’s performance and limitations.

C. Experimental geometries and bacteria tracking

We monitor hundreds of single E. coli in a drop of a diluted homogeneous suspension (concentration \( \sim 10^5 \) bact/mL) squeezed between two horizontal glass slides. The drop has typically a diameter of 1 mm. The gap between the two glass plates is 250 μm. For experiments I, displayed in Fig. 3, only pieces of 3D trajectories remaining between the vertical bounds \( z_m = 10 \) μm from the bottom surface and \( z_M = 140 \) μm, the highest possible height, and lasting more than 8 s are taken into account. For the set of very long tracks in Fig. 4, experiments II, the gap between the glass plates is also 250 μm, but the whole trajectories are captured, as they alternate between the bottom and top. For the analysis, only pieces farther than 10 μm from the surfaces are taken into account.

The velocities are determined from second-order Savitzky-Golay filtering of the coordinates over 0.1 s, resulting in uncertainties close to 5% [36]. For each track, the velocity distribution shows a peak corresponding to the mean run velocity and a low-velocity tail corresponding to the contribution of sudden velocity drops (Fig. 2). Peak velocities are on average \( \langle V \rangle = 27 \pm 6 \) μm/s. To compute the correlation function \( C(\Delta t) \), the average is made over time, and the lag time is offset by 0.2 s to avoid the short-time decorrelation due to wobbling [36,48]. The correlation function is then normalized by its value at 0.2 s to yield 1 at the lag time origin.

D. Track simulations using the BV model

Swimmers are described by their position \( \vec{r} \), orientation \( \hat{\rho} \), and the instantaneous value of the normalized fluctuations of the CheY-P concentration \( \delta X = (|Y|/[Y_0])/\sigma_Y \). During the run phase, they obey the equations

\[
\dot{\vec{r}} = V \hat{\rho},
\]

\[
\dot{\hat{\rho}} = \sqrt{D^\rho (I - \hat{\rho} \hat{\rho}) \eta(t)},
\]

\[
\dot{\delta X} = -\delta X/T_Y + \sqrt{2/T_Y} \xi(t),
\]

where \( V \) is the swim velocity, \( D^\rho \) is the rotational diffusion coefficient, \( T_Y \) is the memory time, \( (I - \hat{\rho} \hat{\rho}) \) is a projector orthogonal to \( \hat{\rho} \), \( \xi \) is a white noise of zero mean and correlation \( \langle \xi(t)\xi(s) \rangle = \delta(t-s) \), and \( \eta \) is a white noise vector of zero mean, where the components have correlations \( \langle \eta_i(t)\eta_k(s) \rangle = \delta_{ik}\delta(t-s) \).

The BV model yields a relation between the characteristic switching time for the transition CCW → CW (run to tumble) and the CheY-P concentration. As a simplification, we assume that, due to the small cellular dimensions, all six flagella operate at the same CheY-P concentration and that the reverse of the rotation direction of a single flagellum is enough to trigger a tumble. Hence, the probability to tumble in \( \Delta t \) would be \( 6\Delta t/\tau_s \). To simplify notation, we absorb the factor 6 into \( \tau_0 \), resulting in a tumble probability \( \Delta t \Delta \delta \sigma /\tau_0 \).

The BV model predicts that the characteristic switching time for the transition CCW → CW (tumble to run) is also given from an activated process. But, as the corresponding value of \( \Delta n \) is small, the tumble duration is given by a Poisson process with characteristic time \( \tau_1 \). In addition, the reorientation dynamics during a tumble needs to be modeled. A priori, the link between motor switch and tumble is far from being trivial, as, in principle, one needs to account for the hydrodynamically complex bundling and unbundling process of the multiflagellated E. coli bacteria [49,50]. Here, we rather follow a simple effective approach inspired by Saragosti, Silberzan, and Buguín [51]. We model the reorientation dynamics during tumbling as an effective rotational diffusion process with a coefficient \( D^\rho_{\text{eff}} \).

Defining the dimensionless combination \( \tilde{D}^\rho_{\text{eff}} = D^\rho_{\text{eff}}/\tau_1 \), the dimensionless tumble durations are sorted from an exponential distribution with a typical time equal to one, and, during a tumble, the dynamics is
\[ \dot{r} = 0. \]  
\[ \dot{p} = \sqrt{D_r^i (I - \hat{p} \cdot \hat{p}) \delta(t)}, \]  
\[ \delta X = 0. \]

After the tumble phase, a new run phase starts.

**ACKNOWLEDGMENTS**

The authors thank Dr. Reinaldo García García for useful discussions, Professor Axel Bugnin for bacterial strains, and Professor Igor Aranson for comments on the manuscript. This work was supported by the ANR grant “BacFlow” ANR-15-CE30-0013 and the Franco-Chilean EcosSud Collaborative Program C16E03. N.F.-M. thanks the Pierre-Gilles de Gennes Foundation for financial support. A.L. and N.F.-M. acknowledge support from the ERC Consolidator Grant PaDyFlow under Grant Agreement No. 682367. V.A.M. was funded by ERC AdG 340877 Ministry of Economy, Development and Tourism Matter of the Millennium Scientific Initiative of the Ministry of Economy, Development and Tourism (Chile). V.A.M. was awarded the European Research Council Consolidator Grant PaDyFlow under Grant Agreement No. 1180791 and Millennium Nucleus Physics of Active Matter of the Millennium Scientific Initiative of the Ministry of Economy, Development and Tourism. V.A.M. was funded by the ERC Consolidator Grant PaDyFlow under Grant Agreement No. 682367. R.S. thank the support from Laboratoire International (PHYSAPS) and Joliot-Curie Chair from ESPCI. E.C. (Chile). V.A.M. was funded by ERC AdG 340877 (PHYSAPS) and Joliot-Curie Chair from ESPCI. E.C. and R.S. thank the support from Laboratoire International Associé “Matière: Structure et dynamique” (LIA-MSD), CNRS-France. This work has received the support of Institut Pierre-Gilles de Gennes (Equipement d’Excellence, “Investissements d’avenir,” Program No. ANR-10-EQPX-34).

**APPENDIX A: RUN DURATION ANALYSIS**

Figure 8 displays a series of analyses on a single trajectory, evidencing that tumble detection is criterium dependent. Here, we set a threshold velocity, cutoff \( V \), and identify as runs all the continuous parts of the track where the bacterial velocity is above the prescribed threshold. The plot demonstrates that, by changing the cutoff value for the velocity, we can obtain average run times of duration 1–8 s.

<table>
<thead>
<tr>
<th>( t_r ) (s)</th>
<th>Cutoff ( V ) (( \mu m/s ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>2</td>
</tr>
<tr>
<td>1.5</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4.5</td>
<td>8</td>
</tr>
<tr>
<td>6.5</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>16</td>
<td>20</td>
</tr>
</tbody>
</table>

**APPENDIX B: PERSISTENCE CORRELATION FUNCTION**

The orientation correlation function is defined as

\[ C(\tau) = \langle \hat{p}(t) \cdot \hat{p}(t + \tau) \rangle = \langle \cos(\theta(\tau)) \rangle, \]  

where \( \hat{p} \) is the director vector and the average is done over time \( t \).

To compute the correlation function, we use a kinetic theory approach. The object under study is the distribution function \( f(\hat{p}, t) \), which gives the probability that a bacterium has an orientation \( \hat{p} \) at time \( t \). In this context, the correlation function is obtained assuming that the initial condition at \( t = 0 \) is with the bacterium pointing in a specific direction, say, \( \hat{p}_0 \). Hence, we have to compute

\[ C(\tau) = \langle \hat{p}(\tau) \cdot \hat{p}_0 \rangle, \]  

where now the average is over the distribution function. At the end, another average, over \( \hat{p}_0 \), should be done. In practice, this last average is unnecessary by the isotropy of space, because the first average gives already a value independent of \( \hat{p}_0 \).

The distribution function obeys the kinetic equation

\[ \frac{\partial f}{\partial t} = -Lf, \]  

with

\[ f(\hat{p}, 0) = \delta(\hat{p} - \hat{p}_0) \]  

and \( L \) the evolution operator. Two models must be considered. In the case of Brownian rotational diffusivity,

\[ Lf = -D^\theta \nabla^2_\hat{p} f, \]  

where \( D^\theta \) is the rotational diffusion coefficient and \( \nabla^2_\hat{p} \) is the angular part of the Laplacian. In the case of tumbling with a characteristic switch time \( \tau_s \),

\[ Lf = \frac{1}{\tau_s} \left[ f - \int d\hat{p}' W(\hat{p}', \hat{p}) f(\hat{p}') \right]. \]  

The kernel \( W(\hat{p}', \hat{p}) \) gives the probability that for a swimmer with director \( \hat{p}' \), after tumbling, the new director is \( \hat{p} \). It is normalized to \( \int d\hat{p}' W(\hat{p}', \hat{p}) = 1 \), indicating that some director \( \hat{p} \) must be chosen. If the space is isotropic, the kernel depends only on the relative angle between the directors, that is, \( W(\hat{p}', \hat{p}) = w(\hat{p}' \cdot \hat{p}) \). Finally, if tumbling and diffusion are present, the operator is just the sum of both.

If the space is isotropic, the evolution operator is also isotropic, which in this case implies that it commutes with the angular Laplacian \( \nabla^2_\hat{p} \). Therefore, both operators share eigenfunctions, which are the spherical harmonics \( Y_{lm}(\hat{p}) \). Then, there are eigenvalues \( \lambda_l \).
Using that where bounded from above by \( \tau \). When for tumbling they are proportional to \( 1/\tau_s \) and depend on the kernel model. In summary,

\[
\lambda_i = D_{\beta}^B l(l + 1) + 1/(a_i \tau_s),
\]

(B7)

where \( a_i \) are dimensionless parameters of the order of 1 that depend on the kernel \( w \).

Using the basis of the spherical harmonics, the solution of the kinetic equation (B2) is

\[
f(\hat{p}, t) = \sum_{lm} f_{lm}(0) Y_{lm}(\hat{p}) e^{-\lambda_i t},
\]

(B8)

where \( f_{lm}(0) \) depend on the initial condition (B3).

Now, the correlation function is

\[
C(t) = \langle \hat{p}(t) \cdot \hat{p}_0 \rangle = \int d\hat{p} \hat{p}_0 \cdot \hat{p} f(\hat{p}, t)
\]

\[
= \sum_{lm} f_{lm}(0) e^{-\lambda_i t} \hat{p}_0 \cdot \int d\hat{p} \hat{p} Y_{lm}(\hat{p}).
\]

(B11)

Using that \( \hat{p} \) can be written as a linear combination of \( Y_{lm} \), with \( m = 0, \pm 1 \) and the orthogonality of the spherical harmonics, it is obtained that the integral is not vanishing only for \( l = 1 \). Combining factors, one obtains

\[
C(t) = C_0 e^{-t/\tau_p},
\]

(B12)

where

\[
\tau_p = \frac{a_1 \tau_s}{1 + a_1 \tau_s/\tau_B}
\]

and \( \tau_B = 1/(2D_{\beta}^B) \) is the Brownian decorrelation time.

In the classical picture, where all bacteria have a single value for \( \tau_s \), the decorrelation time \( \tau_p \) is single valued also. When \( \tau_s \) is broadly distributed, the decorrelation time \( \tau_p \) also follows a broad distribution for \( \tau_p \ll \tau_B \) and it is bounded from above by \( \tau_B \). Finally, in the description of Berg and Brown [1], the tumble angles are distributed with a peak at 63°. In this case, \( a_1 = 1/(1 - \langle \cos \theta \rangle) \) [51].


