Letter

Mutual information as a measure of mixing efficiency in viscous fluids

Yihong Shi [1], Ramin Golestanian [1,2,3,*] and Andrej Vilfan [1,4,†]

[1] *Max Planck Institute for Dynamics and Self-Organization (MPI-DS), Am Fassberg 17, 37077 Göttingen, Germany*

[2] *Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford OX1 3PU, United Kingdom*

[3] *Institute for the Dynamics of Complex Systems, University of Göttingen, 37077 Göttingen, Germany*

[4] *Jožef Stefan Institute, 1000 Ljubljana, Slovenia*

(Received 19 October 2023; accepted 18 May 2024; published 3 June 2024)

Because of the kinematic reversibility of the Stokes equation, fluid mixing at the microscale requires an interplay between advection and diffusion. Here, we introduce mutual information between particle positions before and after mixing as a measure of mixing efficiency. We demonstrate its application in a Couette flow in an annulus and show that nonuniform rotation sequences can lead to more efficient mixing. We also determine mutual information from Brownian dynamics simulations using data compression algorithms. Our results show that mutual information provides a universal and assumption-free measure of mixing efficiency in microscale flows.

DOI: 10.1103/PhysRevResearch.6.L022050

Designing protocols that force an out-of-equilibrium system into equilibrium faster than the natural relaxation rate is a pertinent topic in a number of classical [1,2] and quantum systems [3]. A prime example of forced equilibration is the mixing of fluids if the initial state contains a nonequilibrium concentration or temperature distribution. Fluid mixing at the microscale is of paramount importance in biological organisms and in artificial systems. Examples range from the uptake of oxygen, nutrients, or chemical signals in aquatic organisms to microreactors and "lab-on-a-chip" applications [4–8]. In biology, mixing is frequently accomplished by cilia which drive long-range flows, but also localized regions of chaotic advection [9–14]. A particular challenge to microscale mixing is posed by the time reversibility of flows at low Reynolds numbers [15,16]. Mixing therefore requires an interplay between advection (stirring) and diffusion [17,18]. Although most examples work with the mixing of two distinct fluids, the formalisms that are used apply equally to other scalar quantities such as temperature [17].

The measures that quantify the mixing efficiency can broadly be classified as global and local [8,17]. Global measures typically start by imposing a pattern, e.g., by distributing the solute in a part of the fluid volume. After mixing, the intensity of segregation can be defined as the variance of solute concentration (the L^2 norm) [19,20], its entropy [21–25], the mean distance to the closest particle from the other population [26], or Sobolev norms [20,27]. The same global mixing measures can also be applied to quantify unmixing in cases of spontaneous phase segregation [23]. The limitation of these measures is that the result will depend on the choice of the initial distribution. Local measures typically characterize the amount of stretching or the Lyapunov exponents [16,28,29].

In this Letter, we introduce mutual information as a universal measure of mixing in fluids at low Reynolds numbers with strong interplay between advection and diffusion. As a simple model system, we test our method in a two-dimensional (2D) Couette flow and show that the mixing efficiency depends in a nontrivial way on the time sequence of rotation (see Fig. 1).

Because the theoretically smallest compressed size of a data set is given by its Shannon entropy, lossless data compression algorithms can be used to estimate the entropy of a distribution [30–33]. Order in two-dimensional systems can be estimated with lossless image compression algorithms such as PNG [34] or GIF with LZW compression [35]. However, it has also been argued that compression algorithms lead to poor estimates in two-dimensional systems with long-range correlations [36,37]. We estimate mutual information with different compression algorithms and achieve good accuracy with the neural-network based PAQ8PX [38].

The fluid motion at low Reynolds number is governed by the Stokes equation $\mu\Delta\mathbf{v} = \nabla p$ and incompressibility condition $\nabla \cdot \mathbf{v} = 0$ with the fluid velocity $\mathbf{v}$, pressure p, and viscosity μ. We assume that the boundary shape does not change during the mixing process, although an extension to shape-changing compartments is straightforward. Therefore, the normal velocity has to vanish at the boundary, $\mathbf{n} \cdot \mathbf{v} = 0$. Without inertia the fluid motion at any time t is fully determined by the instantaneous velocities at the boundaries. The particles in the fluid to be mixed—we assume that they are all equivalent—are subject to Brownian motion in addition to flow-driven advection. The time evolution of the probability

*ramin.golestanian@ds.mpg.de

†andrej.vilfan@ds.mpg.de

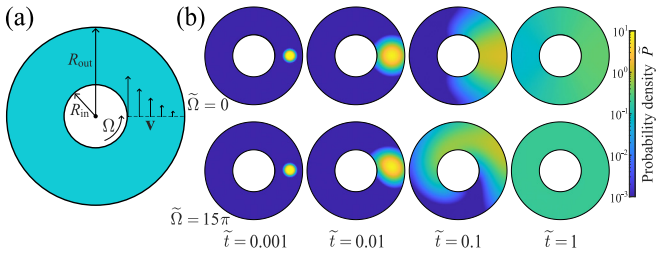

FIG. 1. (a) 2D Couette flow in an annulus geometry. The outer boundary is stationary and the inner is rotating with angular velocity Ω. (b) Time evolution of the probability density $\widetilde{P}(\mathbf{x}, \tilde{t}|\mathbf{x}_0)$, with an initial position $\mathbf{x}_0$ in the middle of the annulus. The top row shows the process with pure diffusion ($\widetilde{\Omega} = 0$) and the bottom row with additional uniform advection ($\widetilde{\Omega} = 15\pi$) (see Movies 1 and 2 in Supplemental Material [39]).

density $P(\mathbf{x}, t)$ is determined by the advection-diffusion equation (equivalent to a Fokker-Planck equation)

$$\partial_t P + \nabla \cdot (\mathbf{v}P) = D\Delta P, \tag{1}$$

with the diffusion constant D. The zero-flux condition implies $\mathbf{n} \cdot \nabla P(\mathbf{x}, t) = 0$ at the boundary. The conditional probability $P(\mathbf{x}, t|\mathbf{x}_0)$ is obtained by solving Eq. (1) with the initial condition $\mathbf{x} = \mathbf{x}_0$ at $t = 0$. Because the flow is divergence free, the stationary solution is always given by a uniform density, $P(\mathbf{x}, t) \equiv 1/V$.

We quantify the mixing efficiency using the mutual information between the initial and the final position of a particle in the fluid. Mutual information provides the strictest possible measure of mixing: Zero mutual information means that the final position of a particle is unrelated to its initial position and therefore also to the position of any other particle in the fluid. Unlike many other criteria, it does not require any assumptions about the initial spatial distribution of the fluid components to be mixed. Mutual information is defined as the sum of entropies of initial and final position distributions, reduced by the joint entropy of the initial and final position together [40]:

$$I[\mathbf{x}_0; \mathbf{x}_t] = S[\mathbf{x}_0] + S[\mathbf{x}_t] - S[\mathbf{x}_0, \mathbf{x}_t]. \tag{2}$$

Here, $S[\mathbf{x}_0]$ is the entropy of the initial position variable, $S[\mathbf{x}_t]$ is the entropy of the position variable at time t, and $S[\mathbf{x}_0, \mathbf{x}_t]$ is the joint entropy of these two variables. The final distribution at time t can also be expressed as $P(\mathbf{x}, t) = \int P(\mathbf{x}, t|\mathbf{x}_0)P(\mathbf{x}_0)d\mathbf{x}_0$. Mutual information can equivalently be expressed with the conditional entropy as

$$I[\mathbf{x}_0; \mathbf{x}_t] = S[\mathbf{x}_t] - S[\mathbf{x}_t|\mathbf{x}_0], \tag{3}$$

where $S[\mathbf{x}_t] = -\int P(\mathbf{x}, t) \log P(\mathbf{x}, t)d\mathbf{x}$, and $S[\mathbf{x}_t|\mathbf{x}_0] = -\int P(\mathbf{x}_0)[\int P(\mathbf{x}, t|\mathbf{x}_0) \log P(\mathbf{x}, t|\mathbf{x}_0)d\mathbf{x}]d\mathbf{x}_0$ is the conditional entropy with the knowledge of the initial position.

Mutual information, as defined above, depends on the distribution of initial positions $P(\mathbf{x}_0)$. This distribution determines the statistical weight that is given to the mixing efficiency in different regions of the fluid. It should not be confused with any imposed pattern in the fluid to be "erased" by mixing, as frequently used in other mixing efficiency criteria. In the following, we assume a homogeneous weight $P(\mathbf{x}_0) = 1/V$. As this is the stationary solution, it also leads to $P(\mathbf{x}, t) = 1/V$ at all later times. We note, however, that as a possible alternative, one could also use an initial distribution that maximizes $I(t)$, following the spirit of Shannon's channel capacity theorem, which would provide a quantitatively stricter measure.

Interestingly, the mixing efficiency is invariant upon time reversal of the mixing sequence, defined by $\bar{\mathbf{v}}(\mathbf{x}, t) = -\mathbf{v}(\mathbf{x}, t_F - t)$, where t_F is the duration of the mixing process. This follows from a general property of the Fokker-Planck equation with divergence-free flux densities [41], for which $\bar{P}(\mathbf{x}, t|\mathbf{x}_0) = P(\mathbf{x}_0, t|\mathbf{x})$. With a uniform distribution $P(\mathbf{x}_0)$, we have $S[\mathbf{x}_t] = S[\mathbf{x}_0]$ and from Eq. (3) it follows immediately that $\bar{I}(t_F) = I(t_F)$.

We now demonstrate the application of mutual information as measure for mixing efficiency in 2D Couette flow. The outer circular boundary with radius R_{out} is stationary and the inner boundary with radius R_{in} rotates with an angular velocity $\Omega(t)$, as shown in Fig. 1(a). In the following we nondimensionalize all variables using the length scale $L = (R_{\text{in}} + R_{\text{out}})/2$ and timescale $T = L^2/D$ such that $\tilde{t} = t/T$, $\tilde{r} = r/L$, etc. In the examples shown below, we use $\tilde{R}_{\text{in}} = 0.6$ and $\tilde{R}_{\text{out}} = 1.4$. The flow in the annulus is solved by [42]

$$\tilde{v} = \widetilde{\Omega}\tilde{r}\frac{\tilde{r}^{-2} - \widetilde{R}_{\text{out}}^{-2}}{\widetilde{R}_{\text{in}}^{-2} - \widetilde{R}_{\text{out}}^{-2}}. \tag{4}$$

In order to numerically determine the conditional entropy $S[\mathbf{x}_t|\mathbf{x}_0]$ in Eq. (3), we use a spectral method to calculate $P(\mathbf{x}, t|\mathbf{x}_0)$ for any initial position $\mathbf{x}_0$. Two examples of the solution for $P(\mathbf{x}, t|\mathbf{x}_0)$, one with pure diffusion and one with uniform rotation, are shown in Fig. 1(b). The conditional entropy $S[\mathbf{x}_t|\mathbf{x}_0]$, needed to determine the mutual information $I(t)$, is eventually obtained by integrating $P(\mathbf{x}, t|\mathbf{x}_0) \log P(\mathbf{x}, t|\mathbf{x}_0)$ over $\tilde{r}_0$, $\tilde{r}$, and θ, and making use of the rotational symmetry of the system with regard to θ_0. The decay of mutual information with pure diffusion and with different advection rates is shown in Fig. 2. In all cases $I(\tilde{t})$ monotonically decreases with time (as in any Markovian process, information can only be lost, but never recovered), but we see that the decay is faster in the presence of rotation, proving that the interplay between advection and diffusion accelerates mixing. At short times, the effect of advection is small and the particle dynamics can be approximated as free diffusion, giving $S[\mathbf{x}_t|\mathbf{x}_0] = \log(4\pi Dt) + 1$ [18]. Inserting the conditional entropy into Eq. (3) and switching to nondimensional units gives an approximation for the mutual information $I(\tilde{t}) = -\log(4\pi\tilde{t}/\tilde{A}) - 1$, where $\tilde{A}$ is the dimensionless surface area [Fig. 2(a), dashed line]. While giving the correct slope, the value of I is slightly underestimated because of neglected boundaries.

A more conventional way of quantifying mixing consists of studying the dissolution of an initial pattern [19]. For example, the initial state can consist of two different fluids. A comparison between the decay of concentration variance and mutual information $I(\tilde{t})$ is shown in Fig. 3. It shows that the slowest of the patterns has the same final exponential decay rate as $I(\tilde{t})$ while others are faster, in line with the argument that mutual information provides the strictest possible measure of mixing efficiency. The calculation of concentration variance is described in the Supplemental Material [39].

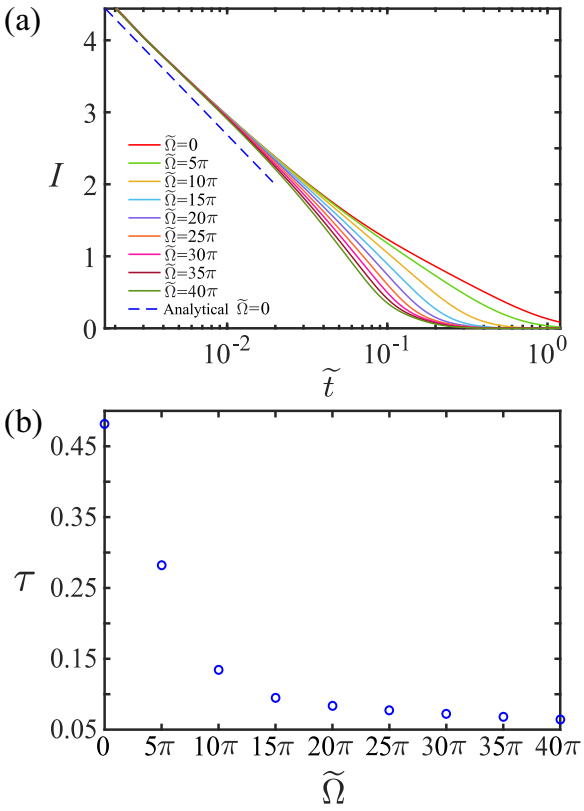

FIG. 2. (a) Temporal decay of mutual information between particle positions in the initial state and the state after time $\tilde{t}$ for different constant rotation rates. The dashed line shows the approximation for free diffusion, $I(\tilde{t}) = -\log(4\pi\tilde{t}/\tilde{A}) - 1$. (b) The time constant of final relaxation for the processes shown in (a), $I \propto \exp(-\tilde{t}/\tau)$.

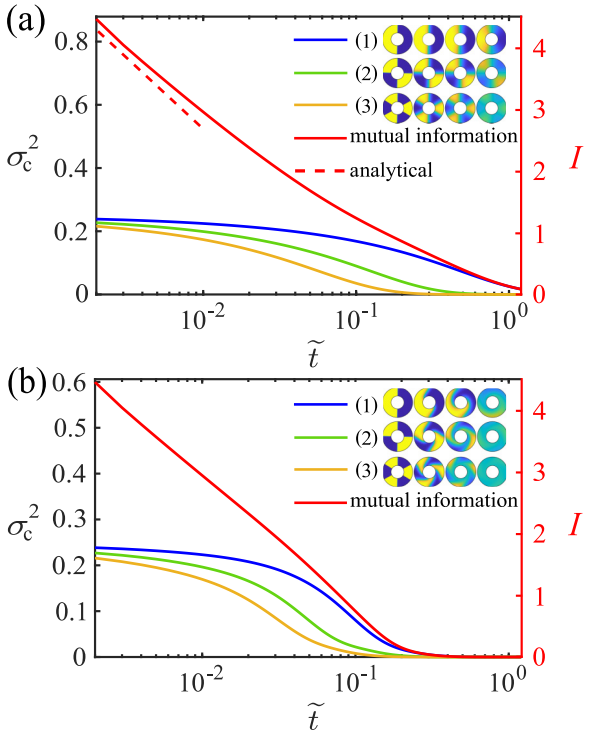

FIG. 3. Temporal decay of mutual information I between particle positions in the initial state and the state after time $\tilde{t}$ (right axis; red) compared with the variance of concentration σ_c^2 for three different initial patterns (left axis; color). The insets show the time evolution of concentration for each initial pattern (see also Fig. S1 [39]). (a) Processes with pure diffusion ($\widetilde{\Omega} = 0$). The dashed line shows the approximation for free diffusion. (b) Processes with both diffusion and advection ($\widetilde{\Omega} = 20\pi$).

Because mixing in viscous fluids requires an interplay between advection and diffusion, we expect that the mixing efficiency depends on the time sequence of rotation $\widetilde{\Omega}(\tilde{t})$ and not just the total angle $\int_0^{\tilde{t}_F} \widetilde{\Omega} d\tilde{t}$ [17]. For example, rotation at the very beginning or the end of the time interval just reorders the positions, but does not reduce the mutual information. We also know, as shown above for any mixing process, that I is invariant if we reverse the mixing sequence in time, $\Omega(t) \to \Omega(t_F - t)$.

We investigated the mixing efficiency of time-dependent advection under two different conditions: (i) for a fixed total rotation $\int_0^{\tilde{t}_F} \widetilde{\Omega} d\tilde{t}$ with the additional condition that the sense of rotation be constant ($\widetilde{\Omega} > 0$) and (ii) for a fixed total viscous dissipation, $\int_0^{\tilde{t}_F} \widetilde{\Omega}^2 d\tilde{t}$. In each scenario, we compared the mixing efficiency of uniform rotation ($\widetilde{\Omega} = \text{const}$) with a Gaussian profile, centered around the middle of the time interval (Supplemental Material Movie 3 [39]) and the globally optimal sequence. We determined the latter with a global optimizer [GlobalSearch in MATLAB (MathWorks, Inc.)], using the velocities in discrete intervals as optimization variables. For a constant rotation, the dependence of I_F on the Gaussian width σ is usually monotonic [Fig. 4(a)] and optimal mixing is achieved by a sharp, discrete rotation at midtime [Fig. 4(d)]. For certain parameters, especially for large total rotation angles and long times, the dependence becomes nonmonotonic [Fig. 4(b)]. The optimal sequence then consists of two symmetrically arranged peaks [Fig. 4(e)]. If the total viscous dissipation is kept constant, mixing sequences with strong nonuniformity naturally become less efficient [Fig. 4(c)]. Then the optimal velocity profile becomes approximately parabolic with a maximum in the middle and dropping to zero at the beginning and end of the interval. Overall, the analysis shows that the velocity profiles that maximize the mixing efficiency are always nonuniform, regardless of whether the total rotation or the total viscous dissipation is kept constant. The symmetry of mixing efficiency upon time reversal is also reflected in the symmetry of the optimal profiles. The exact dependence, however, is complex and can be nonmonotonic in certain parameter regions. A nonmonotonic dependence can also be seen if we quantify mixing using the variance of concentration σ_c^2 (Fig. S2 [39]), but the outcome depends on the initial pattern, in contrast to the universal nature of mutual information.

As an alternative approach, mutual information can also be estimated using lossless data compression algorithms. The (information) entropy of a data set in principle gives a lower bound on its compressed size. By creating a file with initial and final positions of particles $(\mathbf{x}_0, \mathbf{x}_t)$ and compressing it, we obtain an upper bound on the joint entropy $S[\mathbf{x}_0, \mathbf{x}_t]$, and consequently, a lower bound on the mutual information $I(t)$ using Eq. (2). As each position is a two-component

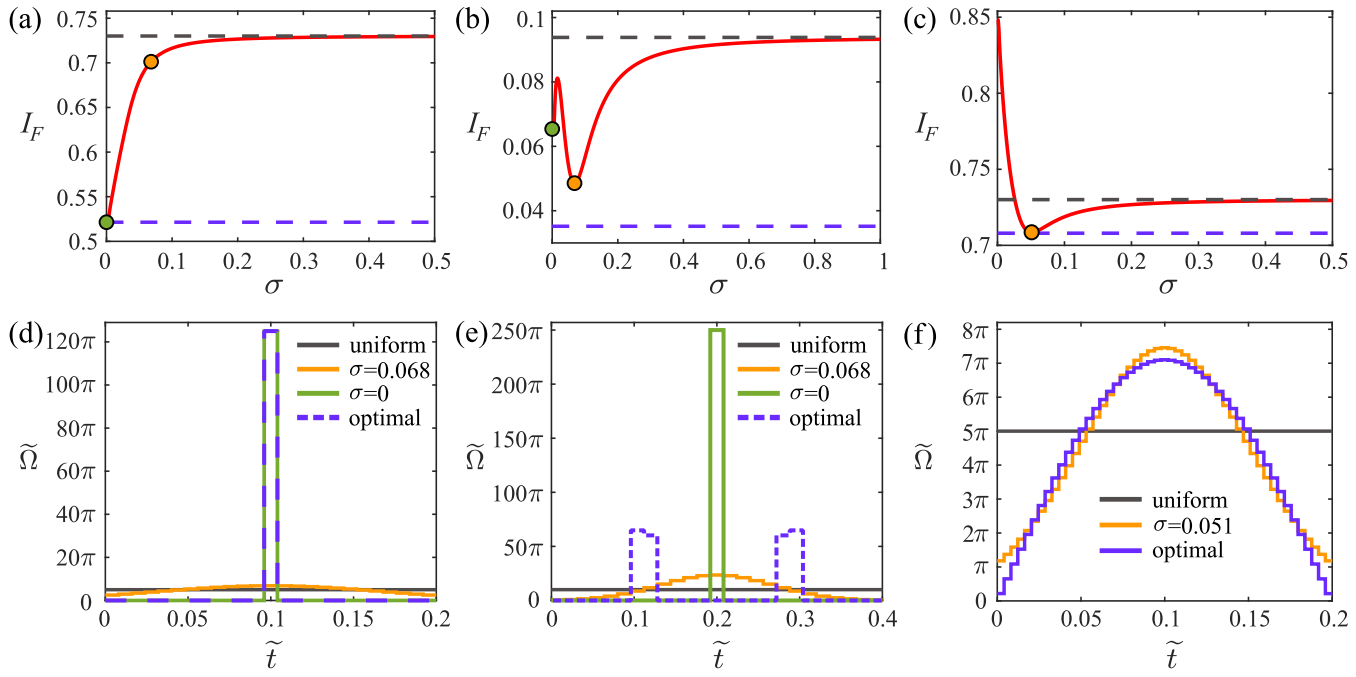

FIG. 4. Mixing efficiency I_F of nonuniform sequences $\widetilde{\Omega}(\tilde{t})$ with (a), (b) a fixed total rotation angle and (c) fixed dissipation. (a) Mixing efficiency of the rotation sequence with a Gaussian dependence $\widetilde{\Omega} \sim \exp[-(\tilde{t}-\tilde{t}_F/2)^2/2\sigma^2]$ with standard deviation σ, discretized with 25 segments, for $\int \widetilde{\Omega} d\tilde{t} = \pi$, $\tilde{t}_F = 0.2$ (red line). The blue dashed line shows the optimal, the gray dashed line shows the uniform sequence, the green point shows the sequence with $\sigma = 0$, and the orange point shows the sequence with $\sigma = 0.068$. (b) As in (a), but with $\int \widetilde{\Omega} d\tilde{t} = 4\pi$, $\tilde{t}_F = 0.4$. (c) Mixing efficiency of Gaussian sequences, discretized with 49 segments, with equal total dissipation $\int \widetilde{\Omega}^2 d\tilde{t} = 5\pi^2$ and $\tilde{t}_F = 0.2$. The orange point shows the sequence with $\sigma = 0.051$. (d)–(f) show the sequences used in (a)–(c), respectively. The optimal sequence shown in (e) becomes bimodal.

vector, the problem is equivalent to computing entropy of a distribution in a four-dimensional space, larger than in most uses of compression algorithms for entropy evaluation reported to date.

We generate a data set by randomly selecting an initial position $\mathbf{x}_0$ within the fluid area. We then simulate the Brownian trajectory with the Euler-Maruyama method with each next position determined as

$$\mathbf{x}_{t+\Delta t} = \mathbf{R}(\omega(|\mathbf{x}_t|)\Delta t) \cdot \mathbf{x}_t + \boldsymbol{\xi}_t. \tag{5}$$

Here, $\mathbf{R}(\Delta\theta)$ denotes a rotation matrix with the rotation angle $\Delta\theta$, $\omega = v/r$ is the local rotation rate and $\boldsymbol{\xi}_t$ is a vector where each component is a Gaussian-distributed random variable with mean zero and standard deviation $\sqrt{2D\Delta t}$. The combination of rotational motion and noise in Cartesian coordinates is chosen to avoid spurious drift caused by the integration procedure. Positions outside the radial range $[R_{\min}, R_{\max}]$ were reflected at the boundary. Both positions were expressed in polar coordinates and the quantities $r_0^2 - R_{\min}^2$, θ_0, $r_F^2 - R_{\min}^2$, θ_F were normalized, converted to 8-bit unsigned integers, and written to a file in this order. Using r^2 as a variable ensures a homogeneous radial distribution. The simulation is repeated with $N = 10^5$, 10^6, or 10^7 starting points, giving a file size of 400 kB, 4 MB, or 40 MB, respectively.

The optimally compressed size of this file (C_t) is a measure of the joint entropy $S[\mathbf{x}_0, \mathbf{x}_t]$ up to a constant that depends on the level of coarse graining. Likewise, the compressed size of an equivalent file with uncorrelated random initial and final positions (C_∞) is a measure of the sum of the entropies of the initial and the final distributions $S[\mathbf{x}_0] + S[\mathbf{x}_t]$. Following Eq. (2), the mutual information can be estimated as

$$I[\mathbf{x}_0; \mathbf{x}_t] \approx (\log 2)\frac{C_\infty - C_t}{N}, \tag{6}$$

with both file sizes measured in bits.

For data compression, we use three different programs: the commonly used Lempel-Ziv-Markov chain algorithm (LZMA) and BZIP2, as well as the experimental, neural network based algorithm PAQ8PX [38,43,44]. A simple test case comprising a one-dimensional drift-diffusion process shows that all three algorithms qualitatively give the right dependence when 10-bit integers are used, but PAQ8PX consistently gave errors $\lesssim$0.1 with large samples (Supplemental Material Fig. S3 [39]). Figure 5 shows the mutual information $I(\tilde{t})$ of our mixing process estimated using the three compression algorithms and different data-set sizes in comparison with the numerical results, obtained with the spectral solution. The comparison is shown for a diffusive process [Fig. 5(a)] and for constant advection with $\widetilde{\Omega} = 20\pi$ [Fig. 5(b)]. Whereas the standard algorithms practically failed to detect any mutual information, PAQ8PX gave results with an absolute error $\ll$1. With the largest size N, the errors are $\lesssim$0.3. LZMA and BZIP2 failed because they use consecutive byte sequences when searching for repeated patterns, whereas PAQ8PX analyzes data bitwise and also includes predictors for arithmetic numbers. These features allow PAQ8PX to detect nontrivial regularities in the significant digits of the positions and not getting distracted by the largely stochastic lower bits.

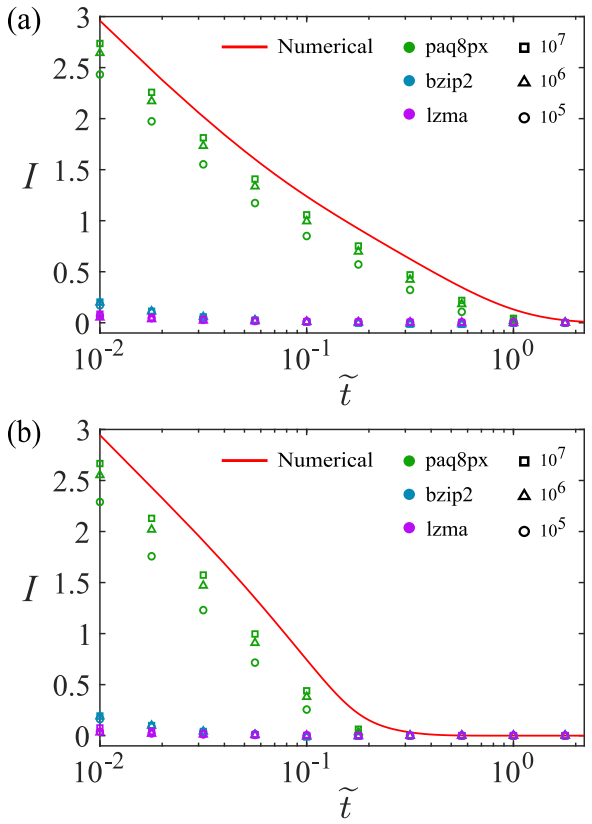

FIG. 5. Mutual information I between the initial state and the state after time $\tilde{t}$, estimated by means of three different compression algorithms (color). Different symbol shapes mark different data-set sizes N. The red line shows the numerical result using the spectral solution. (a) Process with pure diffusion ($\widetilde{\Omega} = 0$). (b) Process with both diffusion and advection ($\widetilde{\Omega} = 20\pi$).

In conclusion, we have introduced mutual information between the initial and the final state as a universal measure for mixing efficiency in microfluidic setups with a strong interplay between advection and diffusion. We have shown that under this measure, the mixing efficiency is symmetric upon time reversal of the actuation sequence. Among all sequences with the same rotation angle, the ones with optimal mixing consist of a fast rotation in the middle of the time interval, or in some cases two symmetrically arranged. We have also demonstrated that advanced neural network based compression algorithms can be applied to estimate mutual information to a high accuracy. The latter can prove useful in more complex flows in which a full solution of the advection-diffusion equation may not be tractable. In some limits, the shear-induced enhancement of mixing shown in our model is similar to Taylor dispersion which requires a relatively thin channel and therefore fast radial diffusion [45]. We also stress that the Couette flow that we chose as a demonstration is far from optimal and that several works have investigated sequences with optimal kinematics under given mixing norms [46–51]. Finding the mixing pattern without geometric restrictions that minimizes mutual information remains a task for future. Furthermore, we expect that our formalism will also be applicable to more complex mixing situations, for example, by active swimmers [52–54], natural or artificial cilia [11–14,55,56], or in active materials and systems where the final distribution is nonuniform or even nonstationary [25].

We thank Aljaž Godec and Reiner Kree for helpful discussions. This work has received support from the Max Planck School Matter to Life and the MaxSynBio Consortium, which are jointly funded by the Federal Ministry of Education and Research (BMBF) of Germany, and the Max Planck Society. A.V. acknowledges support from the Slovenian Research Agency (Grant No. P1-0099).

[1] I. A. Martínez, A. Petrosyan, D. Guéry-Odelin, E. Trizac, and S. Ciliberto, Engineered swift equilibration of a Brownian particle, Nat. Phys. **12**, 843 (2016).

[2] Z. Lu and O. Raz, Nonequilibrium thermodynamics of the Markovian Mpemba effect and its inverse, Proc. Natl. Acad. Sci. USA **114**, 5083 (2017).

[3] M. V. Berry, Transitionless quantum driving, J. Phys. A: Math. Theor. **42**, 365303 (2009).

[4] A. D. Stroock, S. K. W. Dertinger, A. Ajdari, I. Mezić, H. A. Stone, and G. M. Whitesides, Chaotic mixer for microchannels, Science **295**, 647 (2002).

[5] C. J. Campbell and B. A. Grzybowski, Microfluidic mixers: from microfabricated to self-assembling devices, Philos. Trans. R. Soc. London A **362**, 1069 (2004).

[6] R. O. Grigoriev, M. F. Schatz, and V. Sharma, Chaotic mixing in microdroplets, Lab Chip **6**, 1369 (2006).

[7] D. J. Pine, J. P. Gollub, J. F. Brady, and A. M. Leshansky, Chaos and threshold for irreversibility in sheared suspensions, Nature (London) **438**, 997 (2005).

[8] H. Aref, J. R. Blake, M. Budišić, S. S. S. Cardoso, J. H. E. Cartwright, H. J. H. Clercx, K. El Omari, U. Feudel, R. Golestanian, E. Gouillart, G. J. F. van Heijst, T. S. Krasnopolskaya, Y. Le Guer, R. S. MacKay, V. V. Meleshko, G. Metcalfe, I. Mezić, A. P. S. de Moura, O. Piro, M. F. M. Speetjens *et al.*, Frontiers of chaotic advection, Rev. Mod. Phys. **89**, 025007 (2017).

[9] W. Supatto, S. E. Fraser, and J. Vermot, An all-optical approach for probing microscopic flows in living embryos, Biophys. J. **95**, L29 (2008).

[10] N. Uchida and R. Golestanian, Synchronization and collective dynamics in a carpet of microfluidic rotors, Phys. Rev. Lett. **104**, 178103 (2010).

[11] M. Rahbar, L. Shannon, and B. L. Gray, Microfluidic active mixers employing ultra-high aspect-ratio rare-earth magnetic nano-composite polymer artificial cilia, J. Micromech. Microeng. **24**, 025003 (2014).

[12] Y. Ding, J. C. Nawroth, M. J. McFall-Ngai, and E. Kanso, Mixing and transport by ciliary carpets: a numerical study, J. Fluid Mech. **743**, 124 (2014).

[13] S. Jonas, E. Zhou, E. Deniz, B. Huang, K. Chandrasekera, D. Bhattacharya, Y. Wu, R. Fan, T. M. Deserno, M. K. Khokha, and M. A. Choma, A novel approach to quantifying ciliary physiology: microfluidic mixing driven by a ciliated biological surface, Lab Chip **13**, 4160 (2013).

[14] J. C. Nawroth, H. Guo, E. Koch, E. A. C. Heath-Heckman, J. C. Hermanson, E. G. Ruby, J. O. Dabiri, E. Kanso, and M. McFall-Ngai, Motile cilia create fluid-mechanical microhabitats for the active recruitment of the host microbiome, Proc. Natl. Acad. Sci. USA **114**, 9510 (2017).

[15] R. Golestanian, J. M. Yeomans, and N. Uchida, Hydrodynamic synchronization at low Reynolds number, Soft Matter **7**, 3074 (2011).

[16] J. Arrieta, J. H. E. Cartwright, E. Gouillart, N. Piro, O. Piro, and I. Tuval, Geometric mixing, Philos. Trans. R. Soc. London A **378**, 20200168 (2020).

[17] E. Villermaux, Mixing versus stirring, Annu. Rev. Fluid Mech. **51**, 245 (2019).

[18] E. Tang and R. Golestanian, Quantifying configurational information for a stochastic particle in a flow-field, New J. Phys. **22**, 083060 (2020).

[19] P. V. Danckwerts, The definition and measurement of some characteristics of mixtures, Appl. Sci. Res. **3**, 279 (1952).

[20] J.-L. Thiffeault, Using multiscale norms to quantify mixing and transport, Nonlinearity **25**, R1 (2012).

[21] D. D'Alessandro, M. Dahleh, and I. Mezic, Control of mixing in fluid flow: a maximum entropy approach, IEEE Trans. Autom. Control **44**, 1852 (1999).

[22] M. Camesasca, M. Kaufman, and I. Manas-Zloczower, Quantifying fluid mixing with the Shannon entropy, Macromol. Theory Simul. **15**, 595 (2006).

[23] G. B. Brandani, M. Schor, C. E. MacPhee, H. Grubmüller, U. Zachariae, D. Marenduzzo, and C. M. Aegerter, Quantifying disorder through conditional entropy: An application to fluid mixing, PLoS One **8**, e65617 (2013).

[24] R. Kree and A. Zippelius, Chaos and mixing in self-propelled droplets, Phys. Rev. Fluids **4**, 113102 (2019).

[25] J.-L. Thiffeault, Nonuniform mixing, Phys. Rev. Fluids **6**, 090501 (2021).

[26] Z. B. Stone and H. A. Stone, Imaging and quantifying mixing in a model droplet micromixer, Phys. Fluids **17**, 063103 (2005).

[27] G. Mathew, I. Mezić, and L. Petzold, A multiscale measure for mixing, Physica D **211**, 23 (2005).

[28] Y.-K. Tsang, T. M. Antonsen, and E. Ott, Exponential decay of chaotically advected passive scalars in the zero diffusivity limit, Phys. Rev. E **71**, 066301 (2005).

[29] P. Meunier and E. Villermaux, The diffuselet concept for scalar mixing, J. Fluid Mech. **951**, A33 (2022).

[30] T. Schürmann and P. Grassberger, Entropy estimation of symbol sequences, Chaos **6**, 414 (1996).

[31] R. Avinery, M. Kornreich, and R. Beck, Universal and accessible entropy estimation using a compression algorithm, Phys. Rev. Lett. **123**, 178102 (2019).

[32] S. Martiniani, P. M. Chaikin, and D. Levine, Quantifying hidden order out of equilibrium, Phys. Rev. X **9**, 011031 (2019).

[33] M. Zu, A. Bupathy, D. Frenkel, and S. Sastry, Information density, structure and entropy in equilibrium and non-equilibrium systems, J. Stat. Mech. (2020) 023204.

[34] A. Ziepke, I. Maryshev, I. S. Aranson, and E. Frey, Multi-scale organization in communicating active matter, Nat. Commun. **13**, 6727 (2022).

[35] G. Ariel and H. Diamant, Inferring entropy from structure, Phys. Rev. E **102**, 022110 (2020).

[36] E. Brigatti and F. N. M. de Sousa Filho, Comment on "Universal and accessible entropy estimation using a compression algorithm", Phys. Rev. Lett. **129**, 029801 (2022).

[37] F. N. M. de Sousa Filho, V. G. Pereira de Sá, and E. Brigatti, Entropy estimation in bidimensional sequences, Phys. Rev. E **105**, 054116 (2022).

[38] https://github.com/hxim/paq8px (2023).

[39] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.6.L022050 for movies showing the time dependence of probability density, comparison with a conventional mixing norm (variance of concentration), and a 1D test case for compression algorithms.

[40] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, Hoboken, NJ, 2005).

[41] C. Dieball and A. Godec, Coarse graining empirical densities and currents in continuous-space steady states, Phys. Rev. Res. **4**, 033243 (2022).

[42] J. Happel and H. Brenner, *Low Reynolds Number Hydrodynamics* (Springer, Dordrecht, 1983).

[43] M. V. Mahoney, Adaptive weighing of context models for lossless data compression, Tech. Rep. CS-2005-16 (Florida Institute of Technology, 2005), https://cs.fit.edu/media/TechnicalReports/cs-2005-16.pdf.

[44] B. Knoll and N. de Freitas, A machine learning perspective on predictive coding with PAQ8, in *Data Compression Conference (DCC)* (IEEE, New York, 2012), pp. 377–386.

[45] G. I. Taylor, Dispersion of soluble matter in solvent flowing slowly through a tube, Proc. R. Soc. London, Ser. A **219**, 186 (1953).

[46] M. F. Eggl and P. J. Schmid, Mixing enhancement in binary fluids using optimised stirring strategies, J. Fluid Mech. **899**, A24 (2020).

[47] M. F. Eggl and P. J. Schmid, Mixing by stirring: Optimizing shapes and strategies, Phys. Rev. Fluids **7**, 073904 (2022).

[48] O. Gubanov and L. Cortelezzi, Towards the design of an optimal mixer, J. Fluid Mech. **651**, 27 (2010).

[49] Z. Lin, J.-L. Thiffeault, and C. R. Doering, Optimal stirring strategies for passive scalar mixing, J. Fluid Mech. **675**, 465 (2011).

[50] D. P. G. Foures, C. P. Caulfield, and P. J. Schmid, Optimal mixing in two-dimensional plane Poiseuille flow at finite Péclet number, J. Fluid Mech. **748**, 241 (2014).

[51] L. Vermach and C. P. Caulfield, Optimal mixing in three-dimensional plane Poiseuille flow at high Péclet number, J. Fluid Mech. **850**, 875 (2018).

[52] D. Saintillan and M. J. Shelley, Instabilities, pattern formation, and mixing in active suspensions, Phys. Fluids **20**, 123304 (2008).

[53] P. Mueller and J.-L. Thiffeault, Fluid transport and mixing by an unsteady microswimmer, Phys. Rev. Fluids **2**, 013103 (2017).

[54] H. Reinken, S. H. L. Klapp, and M. Wilczek, Optimal turbulent transport in microswimmer suspensions, Phys. Rev. Fluids **7**, 084501 (2022).

[55] J. den Toonder, F. Bos, D. Broer, L. Filippini, M. Gillies, J. de Goede, T. Mol, M. Reijme, W. Talen, H. Wilderbeek, V. Khatavkar, and P. Anderson, Artificial cilia for active microfluidic mixing, Lab Chip **8**, 533 (2008).

[56] A. R. Shields, B. L. Fiser, B. A. Evans, M. R. Falvo, S. Washburn, and R. Superfine, Biomimetic cilia arrays generate simultaneous pumping and mixing regimes, Proc. Natl. Acad. Sci. USA **107**, 15670 (2010).