Information-Optimal Mixing at Low Reynolds Number

Luca Cocconi,[1,*] Yihong Shi,[1] and Andrej Vilfan[2]
[1]Max Planck Institute for Dynamics and Self-Organization, 37077 Göttingen, Germany
[2]Jožef Stefan Institute, 1000 Ljubljana, Slovenia

(Received 5 February 2025; accepted 24 June 2025; published 15 July 2025)

Mutual information between particle positions before and after mixing provides a universal assumption-free measure of mixing efficiency at low Reynolds number that accounts for the kinematic reversibility of the Stokes equation. For a generic planar shear flow with time-dependent shear rate, we derive a compact expression for the mutual information as a nonlinear functional of the shearing protocol and solve the associated extremization problem exactly to determine the optimal control under both linear and nonlinear constraints, specifically total shear and total dissipation per unit volume. Remarkably, optimal protocols turn out to be universal and time-reversal symmetric in both cases. Our results establish a minimum energetic cost of erasing information in a broad class of nonequilibrium drift-diffusive systems.

DOI: 10.1103/cl9m-cmhb

Because of the kinematic reversibility of the Stokes equation [1,2], most compellingly illustrated by G. I. Taylor's Couette cell experiment [3,4], fluid mixing at low Reynolds number requires an interplay between advection (stirring) and diffusion [5,6]. Shear-induced enhancement of diffusive mixing, a phenomenon closely related to Taylor dispersion [7], is fundamental to many biological and artificial systems, from the uptake of oxygen, nutrients, or chemical signals in ciliated aquatic microorganisms to microreactors and "lab-on-a-chip" applications [8–12]. In fact, it represents a fundamental feature of any out-of-equilibrium relaxation process governed by an advection-diffusion equation [5], including the dispersion of pollutants in the upper troposphere and stratosphere [13]. As a result, the design of optimal mixing protocols is a problem of both fundamental and practical importance [14–17] and aligns with a growing interest in applying concepts of optimal control theory to nonequilibrium physics [18–25].

Traditionally, global mixing efficiency has been quantified by imposing an initial pattern (e.g., solute distribution or temperature profile) and characterizing the effect of stirring on the latter through the change in its L^2/Sobolev norms [26,27] or Shannon entropy [14,28,29]. Local mixing may additionally be quantified in terms of Lyapunov exponents [2,30]. More recently, a universal assumption-free (i.e., pattern-independent) metric for global mixing efficiency was introduced in the form of the mutual information between particle positions before and after mixing [15]. In experiments, mutual information can be estimated from tracer data using lossless compression algorithms [31].

Here, we apply this novel metric to the problem of mixing of a fluid by a divergence-free linear shear flow. Defining the time-dependent shear rate as our protocol, we reexpress the mutual information as a nonlinear functional of the latter and solve the optimal control problem exactly to derive optimal protocols under constraints of total shear and total viscous dissipation per unit volume. From these, we derive exact bounds on the mixing efficiency. Our findings demonstrate that mutual information is not only a conceptually elegant framework, but also that its analytical tractability may offer practical advantages in the determination of optimal mixing protocols.

General framework—Let $\mathbf{r}(t) \in \mathbb{R}^d$ denote position in a fluid and $\mathbf{v}(\mathbf{r}, t)$ be a linear shear flow with time-dependent shear rate $\omega(t)$ put under the control of an external operator, $v_i(\mathbf{r}, t) = \omega(t)\mathsf{M}_{ij} r_j$ (Fig. 1). We assume that the flow is divergence-free, $\boldsymbol{\nabla} \cdot \mathbf{v} = 0$, implying $\mathrm{Tr}(\mathsf{M}) = 0$. Following the approach introduced in Ref. [15], we quantify the mixing efficiency of a *protocol* $\{\omega(t)\}_0^T$ of duration T in terms of (minus) the mutual information $I[\mathbf{r}_t; \mathbf{r}_0]$ between the initial and final position of a tagged point particle that is perfectly advected by the flow and simultaneously undergoes Brownian motion with constant diffusivity D. The mutual information can be written in terms of the Gibbs-Shannon entropy S as

$$I[\mathbf{r}_T; \mathbf{r}_0] = S[P(\mathbf{r}_T)] - \int d\mathbf{r}_0 \, P(\mathbf{r}_0) S[P(\mathbf{r}_T|\mathbf{r}_0)], \quad (1)$$

where $P(\mathbf{r}_T)$, $P(\mathbf{r}_0)$, and $P(\mathbf{r}_T|\mathbf{r}_0)$ denote the posterior, prior, and conditional probability densities of the process. For the

*Contact author: luca.cocconi@ds.mpg.de

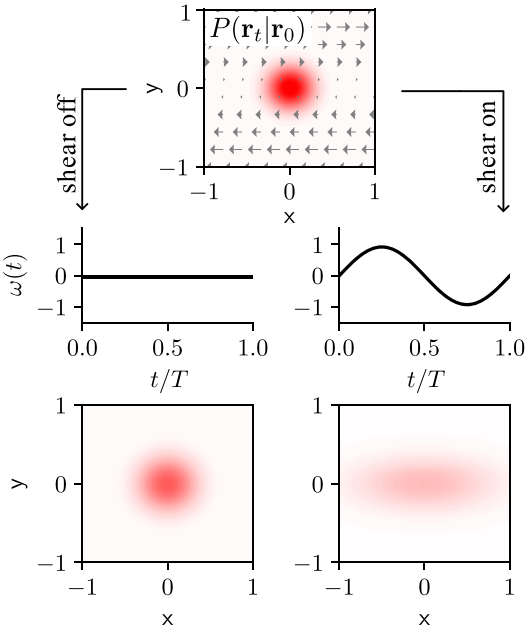

FIG. 1. Enhancement of diffusive mixing in the presence of a time-dependent divergence-free flow, shown here for the case of simple shear, $\mathbf{v}(\mathbf{r},t)=\omega(t)(y,0)$.

natural assumption-free choice of uniform prior distribution, the divergence-free condition implies a uniform posterior. Thus, only the conditional probability depends on the shearing protocol. In particular, the conditional density is governed by the advection-diffusion equation $\partial_t P(\mathbf{r}_t|\mathbf{r}_0)+\boldsymbol{\nabla}\cdot(\mathbf{v}P-D\boldsymbol{\nabla}P)=0$ and may describe the time-dependent solute or temperature distribution given a Dirac-delta initialization $\delta(\mathbf{r}-\mathbf{r}_0)$ at $t=0$. Equivalently, it captures the time-dependent distribution of a tagged particle whose position $\mathbf{r}(t)$ is governed by the linear Langevin dynamics

$$\dot{\mathbf{r}}(t)=\omega(t)\mathsf{M}\mathbf{r}(t)+\sqrt{2D}\boldsymbol{\eta}(t), \tag{2}$$

with Gaussian white noises acting on each coordinate, $\langle\eta_i(t_1)\eta_j(t_2)\rangle=\delta_{ij}\delta(t_1-t_2)$. The conditional probability density associated with Eq. (2) is thus a multivariate Gaussian of the form

$$P(\mathbf{r}_t|\mathbf{r}_0)=\frac{1}{\sqrt{(2\pi)^d\det\Sigma_t}}e^{-\hat{\mathbf{r}}_t^T\Sigma_t^{-1}\hat{\mathbf{r}}_t}, \tag{3}$$

where $\hat{\mathbf{r}}_t=\mathbf{r}_t-\bar{\mathbf{r}}_t$ denotes the deviation from the mean position, which is obtained by solving Eq. (2) with $D=0$, and $\Sigma(t)$ is the covariance matrix, which is independent of the initial position $\mathbf{r}_0$. Using the standard result for the entropy of a multivariate Gaussian distribution we find

$$S[P(\mathbf{r}_t|\mathbf{r}_0)]=\frac{d}{2}\ln(2\pi e)+\frac{1}{2}\ln\det\Sigma(t), \tag{4}$$

which is again independent of $\mathbf{r}_0$. Substituting into Eq. (1), performing the now trivial integral over $\mathbf{r}_0$ and recalling that $S[P(\mathbf{r}_t)]$ is independent of $\omega(t)$, we conclude that maximum mixing efficiency is achieved by protocols that maximize the covariance determinant. Noting that $\Sigma_{ij}(0)=0$, the covariance itself may be written explicitly as [32]

$$\frac{\Sigma(t)}{2D}=\int_0^t dt_1\exp\left[\mathsf{M}\int_{t_1}^t dt_2\omega(t_2)\right]\exp\left[\mathsf{M}^T\int_{t_1}^t dt_2\omega(t_2)\right]. \tag{5}$$

Although the above equation provides a formal solution for $\det\Sigma$, its calculation can be simplified by trading the determinant for another time integral. We combine the Jacobi formula for the determinant of a generic time-dependent matrix $\Sigma(t)$,

$$\frac{d}{dt}\det\Sigma(t)=\det\Sigma(t)\,\mathrm{Tr}\left(\Sigma^{-1}(t)\frac{d\Sigma(t)}{dt}\right), \tag{6}$$

with the Lyapunov equation that follows from the time derivative of Eq. (5),

$$\frac{d\Sigma(t)}{dt}=[\mathsf{M}\Sigma(t)+\Sigma(t)\mathsf{M}^T]\omega(t)+2D\mathbb{I}, \tag{7}$$

where $\mathbb{I}$ is the identity matrix, to obtain

$$\det\Sigma(t)=2D\int_0^t dt_1\ \det\Sigma(t_1)\mathrm{Tr}\Sigma^{-1}(t_1). \tag{8}$$

Equation (8) draws solely on M being traceless, a property that is inherited from the requirement that $\mathbf{v}(\mathbf{r},t)$ is a divergence-free flow.

In the following, we restrict ourselves to the case of a 2D fluid ($d=2$), whereby both Σ and M are 2×2 matrices. Our results will also apply to 3D fluids upon assuming translational invariance along one of the dimensions [33]. We may now write M as a linear superposition,

$$\mathsf{M}=m_r\begin{bmatrix}0&1\\-1&0\end{bmatrix}+\frac{m_{p,1}}{2}\begin{bmatrix}0&1\\1&0\end{bmatrix}+\frac{m_{p,2}}{2}\begin{bmatrix}1&0\\0&-1\end{bmatrix}, \tag{9}$$

where the subscripts r, p refer to rotation and pure shear, respectively. Simple shear along the x axis, for example, corresponds to $2m_r=m_{p,1}$ with $m_{p,2}=0$. For 2×2 matrices, Eq. (8) reduces to

$$\det\Sigma(t)=2D\int_0^t dt_1\,\mathrm{Tr}\Sigma(t_1), \tag{10}$$

where the integrand no longer contains the covariance determinant. $\mathrm{Tr}\Sigma$ is determined from Eq. (5), which can be massaged into a more convenient form by evaluating the traceless matrix exponentials using Putzer's method [34],

$$e^{k\mathsf{M}} = \begin{cases} \cosh(\lambda_{\mathsf{M}} k)\mathbb{I} + \frac{\sinh(\lambda_{\mathsf{M}} k)}{\lambda_{\mathsf{M}}}\mathsf{M} & \text{if } \lambda_{\mathsf{M}} \neq 0 \\ \mathbb{I} + k\mathsf{M} & \text{if } \lambda_{\mathsf{M}} = 0 \end{cases}, \tag{11}$$

where $k \in \mathbb{R}$. The eigenvalue $\lambda_{\mathsf{M}} = \sqrt{-\det \mathsf{M}}$ is non-negative real when $m_{p,1}^2 + m_{p,2}^2 \geq 4m_r^2$ and purely imaginary otherwise. Notice that the case $\lambda_{\mathsf{M}} = 0$ (e.g., simple shear) can be recovered in the limit $\lambda_{\mathsf{M}} \to 0$ and thus need not be treated separately. For $\lambda_{\mathsf{M}} \neq 0$, and using $e^{k\mathsf{M}^T} = (e^{k\mathsf{M}})^T$, we obtain

$$\mathrm{Tr}\Sigma = 2D \int_0^t dt_1 \left[2\cosh^2\left(\lambda_{\mathsf{M}} \int_{t_1}^t dt_2 \omega(t_2)\right) + \frac{\sinh^2(\lambda_{\mathsf{M}} \int_{t_1}^t dt_2 \omega(t_2))}{\lambda_{\mathsf{M}}^2} \mathrm{Tr}(\mathsf{M}\mathsf{M}^T) \right]. \tag{12}$$

Equation (10) then leads to

$$\det \Sigma(T) = 4D^2T^2 + 2D^2\mathcal{A}[\omega], \tag{13}$$

where we used the relation $\cosh^2 = 1 + \sinh^2$ and defined a non-negative actionlike functional $\mathcal{A}[\omega]$ as

$$\mathcal{A}[\omega] \equiv \frac{\gamma_{\mathsf{M}}^2}{\lambda_{\mathsf{M}}^2} \int_0^T dt_1 \int_0^T dt_2 \sinh^2\left(\lambda_{\mathsf{M}} \int_{t_1}^{t_2} dt \omega(t)\right), \tag{14}$$

with $\gamma_{\mathsf{M}}^2 = 2\lambda_{\mathsf{M}}^2 + (\mathsf{M}:\mathsf{M})$ and $\mathsf{M}:\mathsf{M} \equiv \mathrm{Tr}(\mathsf{M}\mathsf{M}^T)$. This functional can be related back to the "excess reduction" in mutual information (equivalently, the increase in mixing efficiency) induced by shearing with respect to a purely diffusive baseline via the expression

$$\Delta I \equiv I[\mathbf{r}_T; \mathbf{r}_0]_{\text{noshear}} - I[\mathbf{r}_T; \mathbf{r}_0] = \frac{1}{2} \ln\left(1 + \frac{\mathcal{A}[\omega]}{2T^2}\right). \tag{15}$$

Perhaps counterintuitively, ΔI does not depend on the diffusivity D. Equations (14) and (15) constitute one of our key results as they reduce the optimal mixing problem to the maximization of a relatively simple functional, which vanishes in the absence of shear. Consistently with general arguments from the theory stochastic processes [35], the functional $\mathcal{A}$ is manifestly invariant under both parity and time reversal of the shearing protocol, i.e.,

$$\mathcal{A}[\omega] = \mathcal{A}[\mathcal{P}\omega] = \mathcal{A}[\mathcal{T}\omega] = \mathcal{A}[\mathcal{P}\mathcal{T}\omega],$$

where $\mathcal{P}\omega(t) \equiv -\omega(t)$ and $\mathcal{T}\omega(t) \equiv \omega(T-t)$ with $t \in [0, T]$. This implies that, in the absence of additional constraints breaking these symmetries explicitly, global optima ω^* maximizing $\mathcal{A}$ must be *at least* fourfold degenerate, if $\mathcal{T}\omega^* \neq \omega^* \wedge \mathcal{P}\mathcal{T}\omega^* \neq \omega^*$, or twofold degenerate, if $\mathcal{T}\omega^* = \omega^* \vee \mathcal{P}\mathcal{T}\omega^* = \omega^*$. The numerical finding in Ref. [15] that globally optimal protocols tend to be $\mathcal{T}$-symmetric is thus nontrivial. We will return to this point shortly.

We henceforth assume $m_{p,1}^2 + m_{p,2}^2 \geq 4m_r^2$, such that $\lambda_{\mathsf{M}} \geq 0$ is real. While some of our results may generalize straightforwardly to the case of imaginary λ_{M}, we argue based on Eq. (14) that such protocols would in general offer a comparatively poorer performance, due to the boundedness of the ensuing trigonometric functions, and are thus of lesser interest.

Optimal protocol under fixed shear—We now proceed to maximize $\mathcal{A}[\omega]$ as given in Eq. (14) with a linear constraint on total shear,

$$\Omega = \int_0^T dt \gamma_{\mathsf{M}} \omega(t), \tag{16}$$

while demanding $\omega(t) > 0$. In the notation of Eq. (9), the characteristic rate squared $\gamma_{\mathsf{M}}^2 = m_{p,1}^2 + m_{p,2}^2$ is clearly nonzero for all nontrivial $\mathsf{M} \neq 0$ excluding solid rotation. By discretization of the time integrals in Eq. (14) (see Appendix B), it can be shown that, for any M, the globally optimal protocol is given by a Dirac-delta impulse at $t = T/2$,

$$\omega_{\text{shear}}^*(t) = \frac{\Omega}{\gamma_{\mathsf{M}}} \delta\left(t - \frac{T}{2}\right). \tag{17}$$

Figure 2(a) shows the result along with the outcome of numerical optimization. Equation (17) is qualitatively consistent with the numerical finding of Ref. [15] that optimal protocols for fixed total shear in a Taylor-Couette cell resemble Dirac-delta impulses in the regime $\Omega \ll 1$. This may be rationalized by noticing that the bulk of a Taylor-Couette flow is well-approximated by homogeneous simple shear for short times and thin gaps.

In realistic settings, we may want to regularize the optimum (17) through the introduction of a total variation regulator penalizing large values of $|\omega'(t)|$; see Appendix C, where we solve the corresponding problem by means of the Euler-Lagrange equation.

Optimal protocol under fixed dissipation—In the second scenario, we determine the optimal protocol under a

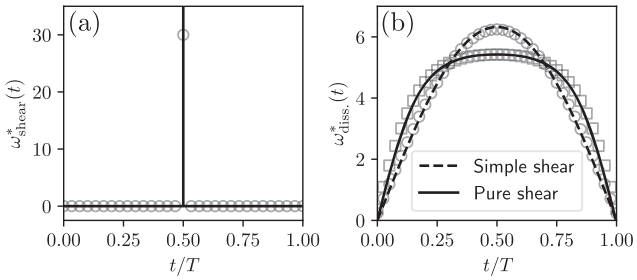

FIG. 2. Numerical validation of the optimal protocols for (a) fixed shear (fixing $\Omega = 1$) and (b) fixed dissipation (fixing $\sigma T/\eta = 5$), showing perfect agreement between theory (lines) and the results of numerical optimization of the discretized problem (markers).

nonlinear constraint on total dissipation per unit area,

$$\sigma = \eta \int_0^T dt \sum_{i,j} (\partial_{r_i} v_j + \partial_{r_j} v_i)^2 = \eta \gamma_{\mathsf{M}}^2 \int_0^T dt \omega^2(t), \quad (18)$$

where η denotes the dynamic viscosity. We now maximize the action $\mathcal{A}[\omega]$ by demanding that its functional derivative with respect to $\omega(t)$, along with a Lagrange multiplier to enforce the constraint in Eq. (18), is zero,

$$\frac{\delta \mathcal{A}[\omega]}{\delta \omega(t)} - 2\mu_\sigma \omega(t) = 0. \quad (19)$$

The derivative follows from Eq. (14) and reads

$$\frac{\delta \mathcal{A}[\omega]}{\delta \omega(t)} = 2\frac{\gamma_{\mathsf{M}}^2}{\lambda_{\mathsf{M}}^2} \int_0^t dt_1 \int_t^T dt_2 \sinh\left(2\lambda_{\mathsf{M}} \int_{t_1}^{t_2} d\tau \omega(\tau)\right). \quad (20)$$

It vanishes for $t = 0$ and $t = T$, supporting our physical intuition that shear contributes most to the mixing efficiency when performed at intermediate times. Differentiating (19) thrice with respect to t and using lower order derivatives to substitute the integrals in the resulting expression (see Appendix D for details), we eventually find that any dependence on the multiplier μ_σ drops out and we are left with the nonlinear ordinary differential equation

$$\omega''(t) = -c^2 \omega(t) + 2\lambda_{\mathsf{M}}^2 \omega^3(t) \quad (21)$$

with an unknown parameter $c^2 > 0$. The boundary conditions are natural, $\omega(0) = \omega(T) = 0$, as evinced from Eq. (19) by noticing that $\delta\mathcal{A}/\delta\omega$ vanishes at the end points [cf. Eq. (20)]. Remarkably, Eq. (21) has a Hamiltonian structure and it is indeed the equation of motion of the anharmonic Duffing oscillator [36]. Its exact solution is given by the Jacobi elliptic functions [37,38]. In particular,

$$\omega^*_{\mathrm{diss}}(t) = \frac{2K(m)\sqrt{m}}{\lambda_{\mathsf{M}} T} \mathrm{sn}\left(\frac{2K(m)t}{T}\middle| m\right), \quad (22)$$

where sn denotes the elliptic sine and where we chose the constant c^2 such that both boundary conditions are satisfied. After reparametrizing the dissipation $\sigma_{\mathsf{M}} \equiv \sigma/(\eta\gamma_{\mathsf{M}}^2)$, we obtain an additional implicit equation for the parameter $0 \le m < 1$ that determines the amplitude, namely $4K(m)[K(m) - E(m)] = \sigma_{\mathsf{M}} \lambda_{\mathsf{M}}^2 T$. Here, $K(m)$ and $E(m)$ denote the complete elliptic integral of the first and second kind [38], respectively. Since $\lim_{\sigma_{\mathsf{M}}\lambda_{\mathsf{M}}^2 T \to 0} m = 0$ and $\mathrm{sn}(u|0) = \sin(u)$, the optimal protocol (22) converges to a sine form, $\omega^*_{\mathrm{diss}}(t) = \sqrt{2\sigma_{\mathsf{M}}/T}\sin(\pi t/T)$, in the limit of small dissipation or for simple shear, $\lambda_{\mathsf{M}} = 0$. Both results agree with the results of numerical optimization [Fig. 2(b)].

Bounds on mixing efficiency—Having access to explicit expressions for the optimal protocols further allows for the derivation of tight bounds on the mixing efficiency. For the case of fixed shear, we substitute (17) into (14) and use the relation (15) to obtain

$$\Delta I(\Omega; \mathsf{M}) \le \frac{1}{2} \ln\left[1 + \frac{\gamma_{\mathsf{M}}^2}{8\lambda_{\mathsf{M}}^2} \sinh^2\left(\frac{\Omega\lambda_{\mathsf{M}}}{\gamma_{\mathsf{M}}}\right)\right]. \quad (23)$$

This bound possesses an M-independent quadratic asymptote at small shear, $\Delta I \le \Omega^2/16 + \mathcal{O}(\Omega^3)$, and is linear at large shear, $\Delta I \le \Omega\lambda_{\mathsf{M}}/\gamma_{\mathsf{M}} + \mathcal{O}(1)$, where it retains an explicit dependence on M.

For the case of fixed dissipation, we may combine Eqs. (14), (15), and (22) to obtain an integral expression for the σ-dependent bound on the mixing efficiency for a given M. When $\lambda_{\mathsf{M}} = 0$ (e.g., simple shear), this bound is determined by explicitly evaluating Eq. (14), expanded to lowest nontrivial order in λ_{M}, in the sinusoidal limit discussed above (cf. Appendix A). This produces the expression (Fig. 3, dashed line)

$$\Delta I(\sigma; \mathsf{M})|_{\lambda_{\mathsf{M}}=0} \le \frac{1}{2} \ln\left[1 + \frac{\sigma T}{2\eta\pi^2}\right]. \quad (24)$$

For $\lambda_{\mathsf{M}} > 0$, the action evaluated at the optimum does not have a closed-form expression and the bound has to be determined by numerical integration (Fig. 3, solid line).

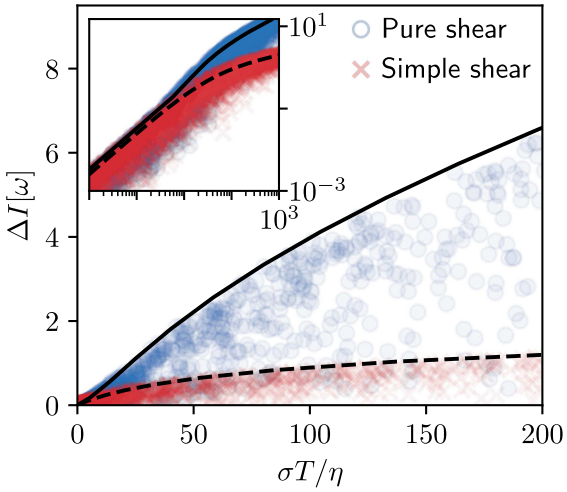

FIG. 3. Dissipation bounds on the mixing efficiency. The tight bound (solid line) is obtained by numerical evaluation of Eq. (14) at the optimum, as described in the main text. The dashed line corresponds to the result for simple shear, Eq. (24). Markers indicate the performance of random controls, sampled from independent realizations of a one dimensional Wiener process, $\omega(t) = sW_t$ with $\log_{10}(s) \in (-2, 4)$ sampled from a uniform distribution. Here, we fix $\lambda_{\mathsf{M}}/\gamma_{\mathsf{M}} = 1/2$ for pure shear, while for simple shear $\lambda_{\mathsf{M}}/\gamma_{\mathsf{M}} = 0$.

The mixing efficiency of the optimal pure shear protocol always surpasses that of simple shear at a given dissipation. For small σ, the two bounds approach each other (Fig. 3, inset), with corrections of order $\mathcal{O}(\lambda_{\mathsf{M}}^4\sigma_{\mathsf{M}}^2T^2)$. At larger σ, we may still construct an explicit, albeit looser, upper bound by evaluating Eq. (14) at constant $\omega(t) = \max_\tau \omega^*_{\mathrm{diss}}(\tau)$. This is discussed in Appendix E, where we additionally show that in the large dissipation asymptote $\sigma_{\mathsf{M}}\lambda_{\mathsf{M}}^2 T \gg 1$, the mixing efficiency may not grow faster than $\sqrt{\sigma}$.

Conclusion—We have derived exact expressions for the information-optimal mixing protocols associated with a generic planar shear flow with time-dependent shear rate under constraints of total shear [Eq. (17)] and total dissipation [Eq. (22)]. The generalization to more complex constraints is straightforward: for example, one could determine the optimal control at fixed dissipation while also requiring the total shear Ω to be zero, such that the fluid returns to the original position like in Taylor's experiment [3]. The functional form of the optimal protocols is found to be universal, i.e., independent of the shear matrix M, and to inherit the time-reversal symmetry of the mixing efficiency ($\omega^* = \mathcal{T}\omega^*$). These results allowed us to compute explicit bounds on the mixing efficiency, as given in Eq. (23) for fixed shear and Eq. (24) for fixed dissipation. The latter result is of particular significance, as it establishes a universal limit on the degradation or erasure of information in drift-diffusive systems, a principle for nonequilibrium physics that extends beyond the immediate context of low Reynolds number mixing.

Given the highly nontrivial form of the functional (14), which defines the optimal control problem, it is remarkable that such closed-form results could be obtained, pointing to mutual information as a promising avenue for further analytical studies of low-Reynolds mixing and enhancement of out-of-equilibrium relaxation more broadly. In particular, future work could explore how the optimal protocols and bounds presented here generalize to nontrivial 3D shear flows or linear flows where each matrix element M_{ij} varies independently under a set of protocols $\omega_{ij}(t)$. The study of finite and periodic systems (e.g., the Couette annulus from Ref. [15]) offers another interesting challenge. Finally, it is worth pointing out that, while the uniqueness of the optimal protocols studied here combined with the $\mathcal{T}$ invariance of the nonlinear action implies that the optima inherit this symmetry (see Ref. [19] for a related discussion), there may in principle be nonlinear control problems for which the $\mathcal{T}$ symmetry is spontaneously broken in some parameter regime. The search for instances of this phenomenon constitutes a fascinating direction for future work.

Acknowledgments—L. C. acknowledges support from the Alexander von Humboldt Foundation. A. V. acknowledges support from the Slovenian Research and Innovation Agency (Grant No. P1-0099).

Data availability—The data that support the findings of this Letter are openly available [39].

[1] E. M. Purcell, Life at low Reynolds number, Am. J. Phys. **45**, 3 (1977).
[2] J. Arrieta, J. H. Cartwright, E. Gouillart, N. Piro, O. Piro, and I. Tuval, Geometric mixing, Phil. Trans. R. Soc. A **378**, 20200168 (2020).
[3] National Committee for Fluid Mechanics Films, *Illustrated Experiments in Fluid Mechanics: The NCFMF Book of Film Notes* (MIT Press, Cambridge, MA, 1972).
[4] J. P. Heller, An unmixing demonstration, Am. J. Phys. **28**, 348 (1960).
[5] E. Villermaux, Mixing versus stirring, Annu. Rev. Fluid Mech. **51**, 245 (2019).
[6] E. Tang and R. Golestanian, Quantifying configurational information for a stochastic particle in a flow-field, New J. Phys. **22**, 083060 (2020).
[7] G. I. Taylor, Dispersion of soluble matter in solvent flowing slowly through a tube, Proc. R. Soc. A **219**, 186 (1953).
[8] A. D. Stroock, S. K. Dertinger, A. Ajdari, I. Mezic, H. A. Stone, and G. M. Whitesides, Chaotic mixer for microchannels, Science **295**, 647 (2002).
[9] C. J. Campbell and B. A. Grzybowski, Microfluidic mixers: From microfabricated to self-assembling devices, Phil. Trans. R. Soc. A **362**, 1069 (2004).
[10] R. O. Grigoriev, M. F. Schatz, and V. Sharma, Chaotic mixing in microdroplets, Lab Chip **6**, 1369 (2006).
[11] 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).
[12] H. Aref, J. R. Blake, M. Budišić, S. S. Cardoso, J. H. Cartwright, H. J. Clercx, K. El Omari, U. Feudel, R. Golestanian, E. Gouillart *et al.*, Frontiers of chaotic advection, Rev. Mod. Phys. **89**, 025007 (2017).
[13] P. Konopka, Analytical Gaussian solutions for anisotropic diffusion in a linear shear flow, J. Non-Equilib. Thermodyn. **20**, 78 (1995).
[14] 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).
[15] Y. Shi, R. Golestanian, and A. Vilfan, Mutual information as a measure of mixing efficiency in viscous fluids, Phys. Rev. Res. **6**, L022050 (2024).
[16] Z. Lin, J.-L. Thiffeault, and C. R. Doering, Optimal stirring strategies for passive scalar mixing, J. Fluid Mech. **675**, 465 (2011).
[17] O. Gubanov and L. Cortelezzi, Towards the design of an optimal mixer, J. Fluid Mech. **651**, 27 (2010).
[18] T. Schmiedl and U. Seifert, Optimal finite-time processes in stochastic thermodynamics, Phys. Rev. Lett. **98**, 108301 (2007).
[19] S. A. M. Loos, S. Monter, F. Ginot, and C. Bechinger, Universal symmetry of optimal control at the microscale, Phys. Rev. X **14**, 021032 (2024).

[20] K. Proesmans, J. Ehrich, and J. Bechhoefer, Optimal finite-time bit erasure under full control, Phys. Rev. E **102**, 032105 (2020).
[21] R. Garcia-Millan, J. Schüttler, M. E. Cates, and S. A. M. Loos, Optimal closed-loop control of active particles and a minimal information engine, arXiv:2407.18542.
[22] L. Cocconi, B. Mahault, and L. Piro, Dissipation-accuracy tradeoffs in autonomous control of smart active matter, New J. Phys. **27**, 013002 (2025).
[23] M. C. Engel, J. A. Smith, and M. P. Brenner, Optimal control of nonequilibrium systems through automatic differentiation, Phys. Rev. X **13**, 041032 (2023).
[24] L. K. Davis, K. Proesmans, and E. Fodor, Active matter under control: Insights from response theory, Phys. Rev. X **14**, 011012 (2024).
[25] K. Proesmans, Precision-dissipation trade-off for driven stochastic systems, Commun. Phys. **6**, 226 (2023).
[26] P. V. Danckwerts, The definition and measurement of some characteristics of mixtures, Appl. Sci. Res. **3**, 279 (1952).
[27] J.-L. Thiffeault, Using multiscale norms to quantify mixing and transport, Nonlinearity **25**, R1 (2012).
[28] M. Camesasca, M. Kaufman, and I. Manas-Zloczower, Quantifying fluid mixing with the Shannon entropy, Macromol. Theory Simul. **15**, 595 (2006).
[29] J.-L. Thiffeault, Nonuniform mixing, Phys. Rev. Fluids **6**, 090501 (2021).
[30] Y.-K. Tsang, T. M. Antonsen Jr, and E. Ott, Exponential decay of chaotically advected passive scalars in the zero diffusivity limit, Phys. Rev. E **71**, 066301 (2005).
[31] F. N. M. de Sousa Filho, V. Pereira de Sá, and E. Brigatti, Entropy estimation in bidimensional sequences, Phys. Rev. E **105**, 054116 (2022).
[32] N. G. van Kampen, The expansion of the master equation, Adv. Chem. Phys. **34**, 245 (1976).
[33] In this case, one of the three coordinates [say, $z(t)$] reduces to an independent Wiener process and the covariance matrix Σ acquires a block diagonal structure. Its determinant is given by the product of the nontrivial upper block, which accounts for the coupled dynamics in the (x, y) plane and is identical to the covariance for the 2D case and the $\Sigma_{zz}(t) = 2Dt$ element, which is protocol-independent.
[34] E. J. Putzer, Avoiding the Jordan canonical form in the discussion of linear systems with constant coefficients, Am. Math. Mon. **73**, 2 (1966).
[35] C. Dieball and A. Godec, Coarse graining empirical densities and currents in continuous-space steady states, Phys. Rev. Res. **4**, 033243 (2022).
[36] J. M. T. Thompson and H. B. Stewart, *Nonlinear Dynamics and Chaos* (John Wiley & Sons, New York, 2002).
[37] A. Salas, J. E. C. Hernández, and L. J. M. Hernández, The Duffing oscillator equation and its applications in physics, Math. Probl. Eng. **2021**, 9994967 (2021).
[38] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Applied Mathematics Series, (US Government Printing Office, Washington, 1968), Vol. 55.
[39] L. Cocconi, Y. Shi and A. Vilfan, Information-optimal mixing at low Reynolds number [data set] (2025), 10.5281/zenodo.15745302.

End Matter

Appendix A: Simple shear and the $\lambda_{\mathsf{M}} = 0$ case—Equation (11) indicates that the case $\lambda_{\mathsf{M}} = 0$ might require special attention. However, we anticipated that the general results derived in the main text apply to this limiting case upon taking the limit $\lambda_{\mathsf{M}} \to 0$. Indeed, expanding Eq. (14) to leading order in small λ_{M} we obtain the quadratic functional

$$\mathcal{A}[\omega] \simeq \gamma_{\mathsf{M}}^2 \int_0^T dt_1 \int_0^T dt_2 [T - \max(t_1, t_2)] \min(t_1, t_2) \omega_{t_1} \omega_{t_2}. \tag{A1}$$

Using simple shear as an example, we now show that Eq. (A1) is exact in the case $\lambda_{\mathsf{M}} = 0$.

Consider simple shear along the x coordinate axis, $m_{p,1} = 2m_r$ with $m_{p,2} = 0$ in Eq. (9). We thus have $\lambda_{\mathsf{M}} = 0$ and, without loss of generality, we rescale rates such that $2m_r = 1$, whereby $\gamma_{\mathsf{M}} = 1$. Since $y(t)$ reduces to a simple Wiener process with diffusivity D, one can directly evaluate the diagonal elements of the covariance matrix,

$$\begin{aligned}
\Sigma_{yy}(t) &= \int_0^t dt_1 \int_0^t dt_2 \langle \dot{y}(t_1) \dot{y}(t_2) \rangle = 2Dt \\
\Sigma_{xx}(t) &= \int_0^t dt_1 \int_0^t dt_2 \langle \dot{x}(t_1) \dot{x}(t_2) \rangle \\
&= 2D \left[t + \int_0^t dt_1 \int_0^t dt_2 \, \omega(t_1) \omega(t_2) \; \min(t_1, t_2) \right].
\end{aligned} \tag{A2}$$

Using Eq. (10) and dropping terms that are independent of the shearing protocol, we conclude that $\operatorname{argmin}_\omega \{ I[\mathbf{r}_t; \mathbf{r}_0] \} = \operatorname{argmax}_\omega \{ \mathcal{A}_{\text{simple}}[\omega] \}$ with a quadratic action $\mathcal{A}_{\text{simple}}[\omega]$ given exactly by the right-hand side of Eq. (A1).

The optimal protocol for fixed total shear Ω and $\omega > 0$ can be determined by identifying the upper bound $\mathcal{A}_{\text{simple}} \leq \max(V)\Omega^2$, where $V(t_1, t_2) \equiv [T - \max(t_1, t_2)] \min(t_1, t_2)$ and noticing that the Dirac-delta impulse $\omega^*(t) = \Omega \delta(t - T/2)$ saturates it, in agreement with Eq. (17). The bound itself follows from the possibility to interpret $\omega(t_1)\omega(t_2)/\Omega^2$ as a normalized measure.

The optimum under the constraint of total dissipation σ can be determined by augmenting the action $\mathcal{A}_{\text{simple}}$ through a suitable Lagrange multiplier enforcing

Eq. (18) and setting its functional derivative to zero to obtain the integral-operator eigenvalue equation

$$\int_0^T dt_1 \min(t_1,t)[T-\max(t_1,t)]\omega^*(t_1) = \mu_\sigma \omega^*(t). \quad \text{(A3)}$$

It follows from substituting (A3) into (A1) that $\mathcal{A}_{\text{simple}} = \mu_\sigma \sigma/\eta$. We are thus specifically after the largest eigenvalue $\mu_\sigma^{(\max)}$ of the operator appearing in (A3). Noticing that (A3) also entails natural boundary conditions at $t=0$ and $t=T$, the eigenfunctions are $f_n(t) = \sin(n\pi t/T)$ with $n \in \mathbb{N}$ and associated eigenvalues $\lambda_n = T^2/(\pi^2 n^2)$. The optimal protocol is thus given by $\omega^*(t) = \sqrt{2\sigma/\eta T}\sin(\pi t/T)$ and it is equal to to the low λ_{M} limit of the optimal protocol for a generic shear M derived in the main text. This finding is validated numerically in Fig. 2.

Appendix B: Proof of Eq. (17)—Let $\Gamma(t)$ denote the antiderivative of the rescaled shear rate, whereby $\int_{t_1}^{t_2} dt \lambda_{\mathsf{M}} \omega(t) = \Gamma(t_2) - \Gamma(t_1)$. Applying the constraint of total shear and demanding that $\omega(t) > 0$, we have $\Gamma(0) = 0$, $\Gamma(T) = \Omega\lambda_{\mathsf{M}}/\gamma_{\mathsf{M}}$ and that $\Gamma(t)$ is a monotonically increasing function of t. The right-hand side of Eq. (14) may then be rewritten as

$$\mathcal{A}[\omega] = \text{const} + \frac{\gamma_{\mathsf{M}}^2}{2\lambda_{\mathsf{M}}^2}\int_0^T dt_1 \frac{1}{\mathcal{F}(t_1)}\int_0^T dt_2 \mathcal{F}(t_2), \quad \text{(B1)}$$

with $\mathcal{F}(t) \equiv e^{2\Gamma(t)}$. The discretized form of the integral in Eq. (B1) reads

$$\Delta t^2\left(\mathcal{F}_i + \sum_{m_1\neq i}\mathcal{F}_{m_1}\right)\left(\frac{1}{\mathcal{F}_i} + \sum_{m_2\neq i}\frac{1}{\mathcal{F}_{m_2}}\right) = \Delta t^2\left(1 + a\mathcal{F}_i + b\frac{1}{\mathcal{F}_i} + ab\right), \quad \text{(B2)}$$

with a, b two positive coefficients independent of $\mathcal{F}_i$. For any $i = 1,\ldots,N = T/\Delta t$, Eq. (B2) is maximized with respect to $\mathcal{F}_i$ when $\mathcal{F}_i = \mathcal{F}(0) = 1$ or $\mathcal{F}_i = \mathcal{F}(T) = e^{2\Omega\lambda_{\mathsf{M}}/\gamma_{\mathsf{M}}}$. Combined with the fact that $\mathcal{F}$ is monotonically increasing, we conclude that $\Gamma(t)$ is a Heaviside step function and thus $\omega^*(t) \propto \delta(t - t_{\text{pulse}})$. Extremizing with respect to t_{pulse} finally gives Eq. (17).

Appendix C: Total variation regularization—We can regularize the optimal protocol under the constraint of total shear by augmenting the action (14) with a total variation regulator to penalize sharp changes in the shear rate,

$$\mathcal{A}_{\text{reg}}[\omega] = \mathcal{A}[\omega] - 2\mu\int_0^T dt[\omega'(t)]^2, \qquad \mu > 0. \quad \text{(C1)}$$

For the sake of brevity, we only consider the limiting case of small λ_{M}, for which the action is quadratic [Eq. (A1)]. Denoting by $G(t)$ the second antiderivative of the shear rate $\omega(t)$, $G(t) = \int_0^t dt' \int_0^{t'} dt'' \omega(t'')$, we obtain, after some integrations by parts and using $G(0) = 0$,

$$\mathcal{A}_{\text{reg}}[G] = 2\int_0^T dt(T-t)G''(t)(tG'(t) - G(t)) - \mu[G'''(t)]^2. \quad \text{(C2)}$$

We can now derive the associated Euler-Lagrange equation: $TG^{(2)}(t) - 2\,\mu G^{(6)}(t) = 0$. It is solved by the (smooth) function

$$\omega^*(t) = c_1\cos(\alpha t) + c_2\sin(\alpha t) + c_3 e^{\alpha t} + c_4 e^{-\alpha t}, \quad \text{(C3)}$$

with $\alpha = (T/2\mu)^{\frac{1}{4}}$ and $c_i \in \mathbb{R}$. Rather than deriving particular expressions for $c_i(\Omega,\mu)$, we limit ourselves to pointing out that, assuming symmetric boundary conditions $\omega^*(0) = \omega^*(T)$, the optimum will be time-reversal symmetric for all Ω and μ. This follows from substituting (C3) back into the action (C1) and noticing that the resulting expression is a quadratic function of the integration coefficients. Since both the boundary conditions and the constraint of total rotation amount to linear relations among the c_i, the global maximum (obtained by extremizing the action with respect to the one remaining free parameter) must be unique and thus symmetric. Note also that in the regularized problem we do not need to break the $\mathcal{P}$ symmetry explicitly for the optimum to be well-defined, since singular behavior at the boundaries is sufficiently penalized by the regulator.

Appendix D: Derivation of Eq. (21)—Consider Eq. (19) with the functional derivative $\delta\mathcal{A}/\delta\omega(t)$ as given in Eq. (20). Again, let $\Gamma(t)$ denote the antiderivative of the rescaled shear rate, whereby $\int_{t_1}^{t_2} dt\lambda_{\mathsf{M}}\omega(t) = \Gamma(t_2) - \Gamma(t_1)$. Differentiating once, twice, and three times with respect to t we obtain, respectively,

$$\bar{\mu}_\sigma \omega'(t) = \int_0^T dt_1 \sinh\{2[\Gamma(t_1) - \Gamma(t)]\} \quad \text{(D1)}$$

$$\bar{\mu}_\sigma \omega''(t) = -2\lambda_{\mathsf{M}}\omega(t)\int_0^T dt_1 \cosh\{2[\Gamma(t_1) - \Gamma(t)]\} \quad \text{(D2)}$$

$$\bar{\mu}_\sigma \omega'''(t) = 4\lambda_{\mathsf{M}}^2\omega^2(t)\int_0^T dt_1 \sinh\{2[\Gamma(t_1) - \Gamma(t)]\} - 2\lambda_{\mathsf{M}}\omega'(t)\int_0^T dt_2 \cosh\{2[\Gamma(t_1) - \Gamma(t)]\}, \quad \text{(D3)}$$

with $\bar{\mu}_\sigma \equiv \mu_\sigma \lambda_{\mathsf{M}}^2/\gamma_{\mathsf{M}}^2$. Substituting Eqs. (D1) and (D2) into the right-hand side of Eq. (D3), we can eliminate the integrals and obtain the $\bar{\mu}_\sigma$-independent ordinary differential

equation

$$\frac{\omega'''}{\omega} - \frac{\omega'\omega''}{\omega^2} - 4\lambda_{\mathsf{M}}^2\omega\omega' = \frac{d}{dt}\left(\frac{\omega''}{\omega} - 2\lambda_{\mathsf{M}}^2\omega^2\right) = 0. \tag{D4}$$

Equation (21) follows by direct integration, with c^2 the integration constant.

Appendix E: Looser bound on mixing efficiency at fixed dissipation—Except for the case $\lambda_{\mathsf{M}} = 0$, for which the closed-form expression (24) was obtained, we have seen in the main text that evaluating the tight bound on the mixing efficiency at fixed dissipation requires numerical integration as a final step. Nevertheless, the action $\mathcal{A}$ being a monotonically increasing function of $\omega(t)$ for any t, we may still construct an explicit, albeit looser, upper bound for $\lambda_{\mathsf{M}} > 0$ by evaluating Eq. (14) at constant $\omega(t) = \omega_{\max}$ with $\omega_{\max} \equiv \max_\tau \omega^*_{\mathrm{diss}}(\tau)$. From inspection of Eq. (22) we find that

$$\omega_{\max} = \frac{2K(m)\sqrt{m}}{\lambda_{\mathsf{M}} T}, \tag{E1}$$

where we recall that the parameter $0 \le m < 1$ is fixed by the implicit equation $4K(m)[K(m) - E(m)] = \sigma_{\mathsf{M}}\lambda_{\mathsf{M}}^2 T$. The following bound on the mixing efficiency follows by direct integration of the action functional

$$\Delta I(\sigma; M)|_{\lambda_{\mathsf{M}}>0} \le \frac{1}{2}\ln\left[1 + \frac{\gamma_{\mathsf{M}}^2}{4\xi_{\mathsf{M}}^2}\left(\sinh^2(\xi_M) - \xi_M^2\right)\right], \tag{E2}$$

where $\xi_{\mathsf{M}} \equiv 2\lambda_{\mathsf{M}}\omega_{\max} T$. Since $\lim_{\sigma\lambda_{\mathsf{M}}^2 T\to\infty} mK^2(m) = \sigma_{\mathsf{M}}\lambda_{\mathsf{M}}^2 T/4$, which follows from the implicit equation for m, the right-hand side of Eq. (E2) has a square root asymptote in the limit of large dissipation, specifically

$$\Delta I(\sigma; M)|_{\lambda_{\mathsf{M}}>0} \le \sqrt{\frac{4\sigma\lambda_{\mathsf{M}}^2 T}{\eta\gamma_{\mathsf{M}}^2}}, \tag{E3}$$

for $\sigma_{\mathsf{M}}\lambda_{\mathsf{M}}^2 T \gg 1$. We conclude that, for any M with $\lambda_{\mathsf{M}} > 0$, the tight bound (24) cannot grow faster than $\sqrt{\sigma}$ in the large dissipation regime. Indeed, numerical evidence indicates that this is also the asymptotic scaling of the tight bound for pure shear.