Schrödinger dynamics of undulatory locomotion

Alexander E. Cohen,[1,2] Alasdair D. Hastewell,[1] Sreeparna Pradhan,[3] Steven W. Flavell,[3] and Jörn Dunkel[1,*]

[1] *Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139*

[2] *Department of Chemical Engineering, Massachusetts Institute of Technology, 25 Ames Street, Cambridge, MA 02142*

[3] *Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 43 Vassar Street, Cambridge, MA 02139*

(Dated: June 3, 2022)

Spectral mode representations play an essential role in various areas of physics, from quantum mechanics to fluid turbulence, but they are not yet extensively used to characterize and describe the behavioral dynamics of living systems. Here, we show that mode-based linear models inferred from experimental live-imaging data can provide an accurate low-dimensional description of undulatory locomotion in worms, robots, and snakes. By incorporating physical symmetries and known biological constraints into the dynamical model, we find that the shape dynamics are generically governed by Schrödinger equations in mode space. Similar to quantum systems, the eigenstates of the biophysical Hamiltonians enable the efficient classification and differentiation of locomotion behaviors in natural, simulated, and robotic organisms. While our analysis focuses on a widely studied class of biophysical locomotion phenomena, the underlying approach generalizes to other physical or living systems that permit a mode representation subject to geometric shape constraints.

Undulatory propulsion is the natural locomotion strategy [1, 2] of many aquatic and terrestrial animals, from worms [3–7] and fish [8, 9] to lizards [10] and snakes [11, 12]. The mechanical wave patterns that drive undulatory motion reflect an animal's behavioral state [13], providing a macroscopic physical readout of the underlying biochemical and neuronal excitations. Recent advances in automated live-imaging [14, 15] enable simultaneous observations of macroscopic locomotion dynamics and microscopic cellular activity [16–21], producing rapidly growing multi-scale data sets [22] that have to be tracked [23–25] and translated into predictive and interpretable models. Despite recent major progress in the experimental characterization [16–20] and biophysical description of specific organisms [4, 10, 26–32], a quantitative model inference framework for comparing experimentally observed undulatory dynamics within and across species has yet to be developed. In addition to providing unifying biophysical insights spanning different animal kingdoms, such a framework would also allow for a direct comparison of living systems with computational models [33, 34] and biomimetic robotic devices [35, 36].

Here, we use spectral mode representations to identify symmetry-constrained dynamical models that can capture and distinguish the undulatory locomotion of worms (*Caenorhabditis elegans*) [13], neuro-mechanical worm models [33], Mojave shovel-nosed snakes (*Chionactis occipitalis*) [37], and mechanical snakes. Compared with traditional continuum descriptions of undulatory shape-deformations in position space, formulating locomotion models in mode space [38–41] offers several theoretical and practical advantages: (*i*) high-dimensional experimental data can be efficiently compressed to obtain an interpretable low-dimensional representation; (*ii*) the mode dynamics reduces to a system of linear ordinary differential equations (ODEs); (*iii*) physical symmetries and biological constraints can be efficiently encoded through the structure of the dynamical matrix; (*iv*) all model parameters can be directly inferred from experimental data using ODE sensitivity methods [42, 43] that exploit the imposed matrix structure [44]. In particular, for undulatory locomotion, we find that translational invariance, rotational invariance, and length constraints generically lead to a Schrödinger equation [45] in mode space. Analogous to the characterization of quantum systems in terms of their spectra and eigenstates [46], the eigenspaces of the effective Hamiltonians enable an efficient classification of the locomotion dynamics of worms, snakes, robots, and computational models. While our discussion focuses on an important subclass of biophysical dynamics, the underlying approach generalizes to other physical or living systems that permit a mode representation while being subject to exact or approximate geometric constraints.

The planar undulatory locomotion of an elongated worm-like object can be described by its centerline $(x(s,t), y(s,t))$, where s is the arc length and t denotes time [Fig. 1(a) and (b)]. While tens to hundreds of points are typically required for an accurate depiction of an organism's shape in position space [16], interpretable lower-dimensional representations can often be obtained by projecting on suitable polynomial, trigonometric, or other basis functions [26, 47]. Although system-specific representations, such as PCA-based eigenworms [26, 31], yield near-optimal compression, general orthogonal basis systems enable direct comparisons across systems, which is the main objective of our study. Here, we use Chebyshev polynomials [48] of the first kind, $T_k(s)$, which are known to have advantageous analytical and computational properties; in principle other bases can

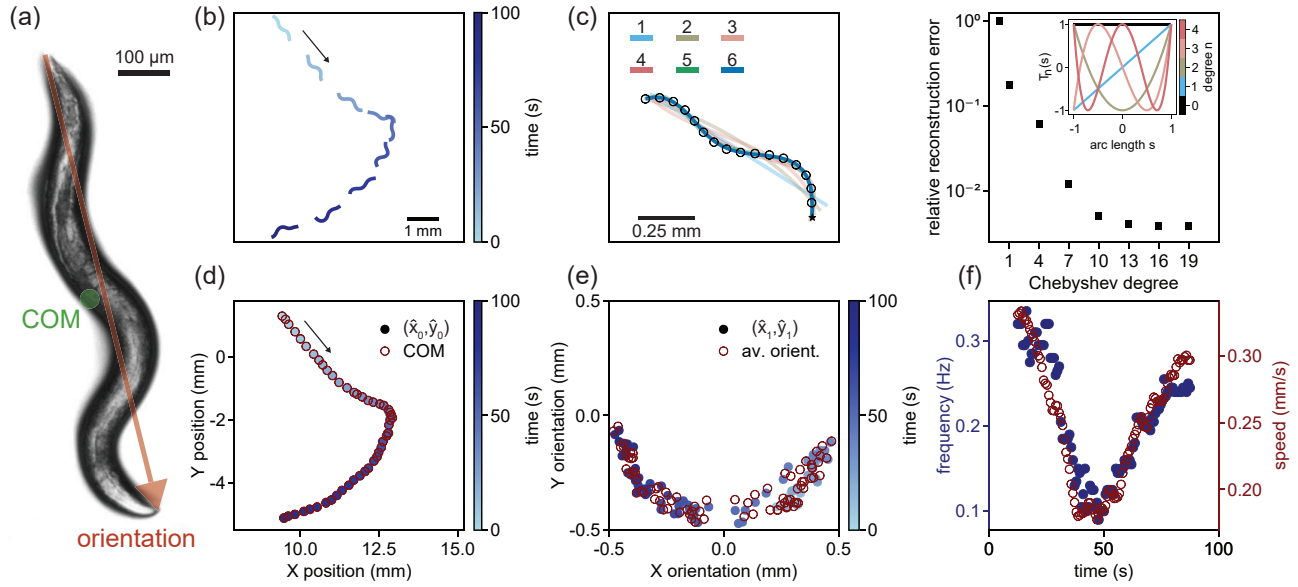

FIG. 1. Chebyshev mode representation enables an efficient and interpretable low-dimensional description of undulatory locomotion across species and model systems. (a) Experimental image of *C. elegans* worm with center of mass (COM) and mean orientation overlayed. (b) Tracked centerline of worm over 100 seconds. Arrow indicates direction of motion. (c) A small number of Chebyshev polynomials suffices to accurately reconstruct the worm shape (left). Reconstruction error (right) decays rapidly as the Chebyshev degree n increases. (d) The zeroth-order Chebyshev coefficients follow closely the worm's geometric COM, illustrating the physical interpretability of Chebyshev mode representation. (e) Similarly, the first-order Chebyshev coefficients represent the tail-to-head worm orientation. (f) The mode-averaged dominant frequency of Chebyshev mode oscillations correlates closely with the locomotion speed of worm.

be used as well. The dynamics of the two scalar fields $(x(s,t), y(s,t))$ can then be represented in terms of their leading Chebyshev coefficients, $\hat{x}_k(t)$ and $\hat{y}_k(t)$, up to degree n defined by

$$\begin{bmatrix} x(s,t) \\ y(s,t) \end{bmatrix} = \sum_{k=0}^{n} T_k(s) \begin{bmatrix} \hat{x}_k(t) \\ \hat{y}_k(t) \end{bmatrix}. \tag{1}$$

For the experimental imaging data analyzed below, $n+1 = 10$ modes suffice for achieving reconstruction errors less than 1% [Fig. 1(c); SI 2]. To calculate the kth coefficient, we take the inner product with respect to the Chebyshev weight function $w(s) = 1/\sqrt{1-s^2}$,

$$\begin{bmatrix} \hat{x}_k(t) \\ \hat{y}_k(t) \end{bmatrix} = \frac{\gamma_n}{\pi} \int_{-1}^{1} \mathrm{d}s \; w(s)\, T_k(s) \begin{bmatrix} x(s,t) \\ y(s,t) \end{bmatrix} \tag{2}$$

where $\gamma_0 = 1$ and $\gamma_n = 2$ for $n > 0$. We illustrate the physical meaning of the Chebyshev modes using recent tracking microscopy video data [16] for *C. elegans* [Fig. 1(a) and (b)], a widely studied model organism with 95 body wall muscle cells, 302 neurons, and a rich set of behavioral states and corresponding locomotion patterns [13]. The degree-0 coefficients $[\hat{x}_0(t), \hat{y}_0(t)]$, obtained from Eq. (2) with $T_0(s) = 1$, describe the w-weighted Chebyshev center of mass (CCOM) of the moving worm, which follows closely the geometric center of mass [Fig. 1(d)]. The degree-1 coefficient vector with $T_1(s) = s$ represents the mean orientation of the worm [Fig. 1(e), SI]. Similarly, the Chebyshev vectors $[\hat{x}_k, \hat{y}_k]$ with $k \geq 2$ encode curvature and higher deformation modes [Fig. 1(c), inset]. The average dominant frequency across the mode oscillations closely matches the speed of the worm in real space [Fig. 1(f); SI 1].

Equipped with this representation, we seek to formulate a dynamical model for undulatory motion in mode space. Defining a combined mode vector $\boldsymbol{\Psi} = [\hat{x}_0, \ldots, \hat{x}_n, \hat{y}_0, \ldots, \hat{y}_n]$, the most general coupled linear first-order dynamics is $\dot{\boldsymbol{\Psi}} = M\boldsymbol{\Psi}$. Incorporating symmetries and invariances into the model imposes additional structure on M. Rotational invariance constrains M to have a block form with equal diagonal blocks and opposite sign off-diagonal blocks (SI). Translational invariance requires the CCOM $\boldsymbol{\psi}_0 = [\hat{x}_0, \hat{y}_0]$ to decouple from the higher degree coefficients $\hat{\boldsymbol{\psi}} = [\hat{\boldsymbol{\psi}}_{\boldsymbol{x}}, \hat{\boldsymbol{\psi}}_{\boldsymbol{y}}] = [\hat{x}_1, \ldots, \hat{x}_n, \hat{y}_1, \ldots, \hat{y}_n]$ that describe the orientation and shape. Abbreviating $x_s = \partial x/\partial s$, an additional biophysical constraint for undulatory motion is that the length of the centerline $\ell(t) = \int_{-1}^{1} \mathrm{d}s \sqrt{(x_s(s,t))^2 + (y_s(s,t))^2}$, remains approximately constant (SI). In mode space, length variations can be bounded by conserving the convex quadratic functional

$$\tilde{\ell}^2 = \int_{-1}^{1} \mathrm{d}s \; \left[x_s(s,t)^2 + y_s(s,t)^2\right] = \hat{\boldsymbol{\psi}}^{\dagger} \begin{bmatrix} W & 0 \\ 0 & W \end{bmatrix} \hat{\boldsymbol{\psi}} \tag{3}$$

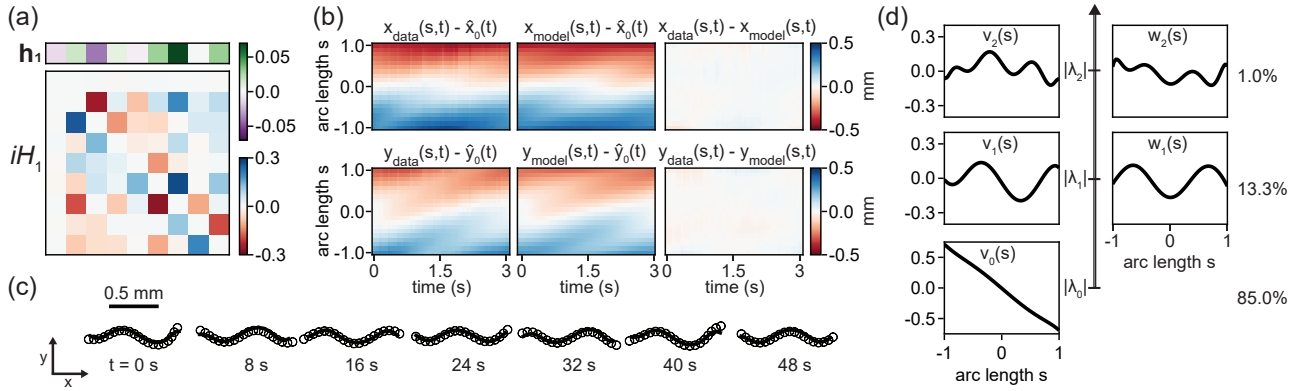

FIG. 2. Inferred Schrödinger dynamics replicate stereotypical *C. elegans* locomotion. (a) Representative propulsion vector $\boldsymbol{h}_1$ and Hamiltonian sub-matrix iH_1, see Eq. (4). (b) Kymopgraphs of $x(s,t)$ and $y(s,t)$ coordinate fields for observed data (left) and model prediction (middle) show little deviation (right), confirming that the Schrödinger model [Eq. (4)] can accurately capture undulatory shape dynamics of *C. elegans*. (c) Real-space dynamics predicted by the Schrödinger model (line) is consistent with the observed worm dynamics (circles); see Movie S1. Model has been fitted on a single period ($\tau = 3.05$s); experimental data has been periodically extended for visualization to avoid overlapping body segments. (d) Real-space shape functions [Eq. (5)] corresponding to the three smallest magnitude eigenvalues, $\lambda_k^{\pm} = \pm k\lambda$ for $k = 0, 1, 2$, account for $> 98\%$ of the shape dynamics, enabling a generalizable low-rank description.

where W is a symmetric matrix with elements $W_{k,m} = \int_{-1}^{1} \mathrm{d}s\, \partial_s T_k(s)\, \partial_s T_m(s)$. In particular, W is positive definite and can thus be interpreted as a basis-specific metric. Taylor expanding the curve length ℓ around the space- and time-average of x_s^2 and y_s^2, denoted by $\langle\cdot\rangle$, shows that ℓ is approximately proportional to $\tilde{\ell}^2/\sqrt{\langle x_s^2 + y_s^2\rangle}$; additionally, the Cauchy-Schwarz inequality implies $\ell^2 \leq 2\tilde{\ell}^2$ (SI). Therefore, demanding constant $\tilde{\ell}$ ensures ℓ remains approximately constant and bounded, as desired. Keeping Eq. (3) constant forces the shape-modes $\hat{\boldsymbol{\psi}}$ onto a hyperellipsoid, with axes determined by W. Using the Cholesky factorization $W = LL^\dagger$, this hyperellipsoid can be transformed to a unit hypersphere by defining the rescaled mode vectors $\boldsymbol{\psi}_x = (L^\dagger/\tilde{\ell})\hat{\boldsymbol{\psi}}_x$, $\boldsymbol{\psi}_y = (L^\dagger/\tilde{\ell})\hat{\boldsymbol{\psi}}_y$, and $\boldsymbol{\psi} = [\boldsymbol{\psi}_x, \boldsymbol{\psi}_y]$. Under this transformation, the length constraint (3) becomes a normalization condition

$$\boldsymbol{\psi}^\dagger \boldsymbol{\psi} = 1. \tag{4a}$$

Combined with rotational and translational invariance, the normalization restricts the class of permissible linear models to the form (SI)

$$\dot{\boldsymbol{\psi}}_0 = H_0 \boldsymbol{\psi} \tag{4b}$$

$$i\dot{\boldsymbol{\psi}} = H\boldsymbol{\psi}, \tag{4c}$$

where H_0 and H have the structure

$$H_0 = \begin{bmatrix} \boldsymbol{h}_1^\dagger & \boldsymbol{h}_2^\dagger \\ -\boldsymbol{h}_2^\dagger & \boldsymbol{h}_1^\dagger \end{bmatrix}, \quad H = \begin{bmatrix} H_1 & H_2 \\ -H_2 & H_1 \end{bmatrix} \tag{4d}$$

and iH is real skew-symmetric so H is Hermitian and has real eigenvalues. Equation (4b) describes how the CCOM dynamics couples to the body oscillations through H_0. Equation (4c), which governs the shape dynamics, is mathematically equivalent to a Schrödinger equation with Hamiltonian H [45]. For straight motion, one does not expect the x- and y-modes to couple significantly, so we may set $\boldsymbol{h}_2 = 0$ and $H_2 = 0$ in this case. To confirm that Eqs. (4) can indeed describe and distinguish the undulatory dynamics of *C. elegans* worms [16] and other systems, we implemented a inference framework for estimating the propulsion vector $\boldsymbol{h}_1$ and the shape Hamiltonian H_1 from experimental data (Fig. 2 and 3).

Both $\boldsymbol{h}_1$ and H_1 can be efficiently determined from tracked centerlines via a physics-informed dynamic mode decomposition [49, 50] that exploits matrix structure [44]. Since iH_1 is real skew-symmetric, it permits the spectral decomposition $iH_1 = Q\Sigma Q^\dagger$, where Q is a real orthogonal matrix and Σ is a real block diagonal matrix. This leaves $\lfloor n^2/2 \rfloor$ parameters in Σ and Q plus n in $\boldsymbol{h}_1$ to be determined from data. The number of parameters can be reduced further by constraining the spectrum of H_1 to take integer multiples of the base frequency $\lambda = 2\pi/\tau$, where τ is the dominant oscillation period. We found that estimating τ directly from the Fourier spectrum (SI) yields periodic oscillatory dynamics consistent with experimentally observed straight motion. To infer the remaining parameters $(Q, \boldsymbol{h}_1)$ and to avoid numerical differentiation of noisy data, our inference scheme compares numerically integrated predictions from Eqs. (4) directly to the experimental data (SI). The underlying algorithm sequentially optimizes Q and $\boldsymbol{h}_1$ by minimizing deviations from both real space worm shapes and mode space trajectories, to balance shape matching with model generalizability, and to prevent overfitting (SI). Minimization is performed using gradient-based optimization [51–53] with forward mode automatic differentiation through the ODE solver [41–43]. Overall, this scheme makes it possi-

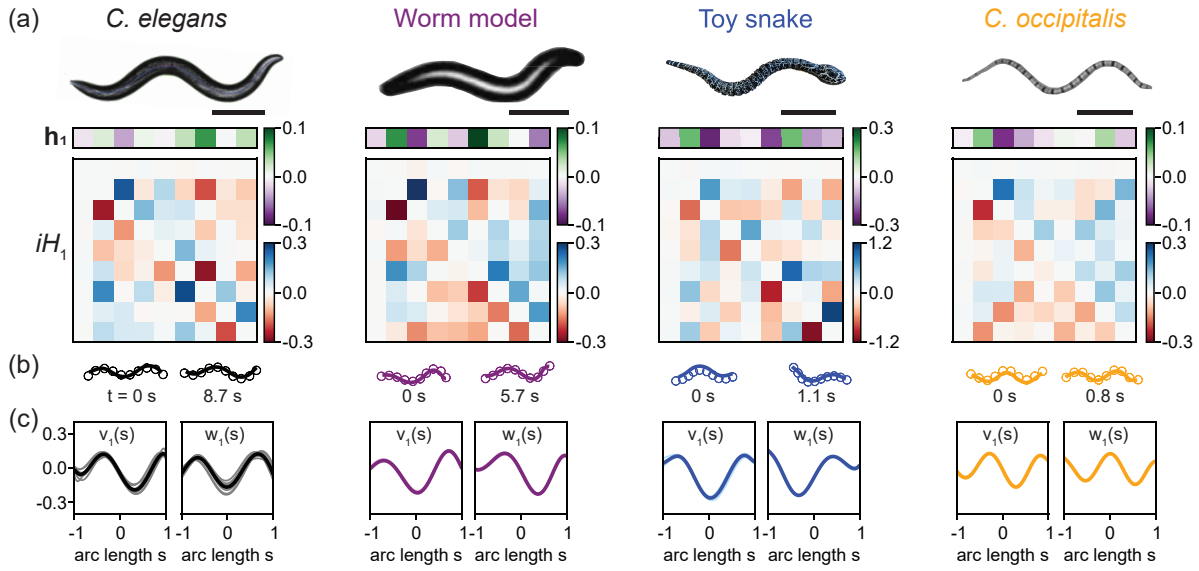

FIG. 3. Mode-space Hamiltonians provide a compact dynamical description of undulatory motion across different species and model systems. (a) Living and nonliving systems [33, 37] analyzed here and representative instantaneous Hamiltonians inferred from a single oscillation period. The eigenspaces of the Hamiltonians enable the comparison and classification of undulation dynamics, see Fig. 4. Scale bars are 0.25 mm (worm), 0.25 mm (worm model), 10 cm (toy snake), and 10 cm (snake). (b) Inferred Schrödinger model dynamics (line) provide an accurate description of the observed dynamics (circles). Models were fitted on a single period $\tau = 3.1$ s (worm), 2.2 s (worm model), 0.45 s (toy snake), and 0.33 s (snake); see also Movie S1. Experimental data has been periodically extended for visualization to avoid overlapping body segments. (c) The dominant shape eigenvectors $v_1(s)$ and $w_1(s)$ are consistent within each species and capture differences between species.

ble to infer the instantaneous shape Hamiltonian H_1 and the propulsion vector $\boldsymbol{h}_1$ from just a single oscillation period. For *C. elegans* (Fig. 2) as well as for previously proposed neuro-mechanical worm models [33], *C. occipitalis* snakes [37] and snake robots (Fig. 3), the best-fit models based on Eqs. (4) accurately capture the undulatory dynamics (Movie S1).

Since the shape dynamics are encoded by the Hamiltonian H, we can use its eigenstates to compare and classify undulatory motion across species and systems [46]. Indeed, for straight motions, it suffices to study the eigenstates of H_1. Considering $n = 9$ as before, H_1 has one zero eigenvalue $\lambda_0 = 0$ corresponding to the zero-mode eigenvector $\boldsymbol{\phi}_0$, and 4 distinct pairs of opposite sign eigenvalues $\lambda^{\pm}_{k\geq 1} = \pm k\lambda$ with a set of complex conjugate eigenvectors $\boldsymbol{\phi}^{\pm}_k$, where $\boldsymbol{\phi}^{+}_k = (\boldsymbol{\phi}^{-}_k)^*$. We define two real orthogonal mode space vectors $\boldsymbol{v}_k = \Re(\boldsymbol{\phi}^{+}_k)$ and $\boldsymbol{w}_k = \Im(\boldsymbol{\phi}^{+}_k)$ that span the eigenspace of $\boldsymbol{\phi}^{\pm}_k$. The real space shape functions corresponding to the real mode space vectors are

$$v_k(s) = \ell(L^{-1}\mathbf{T}(s))^{\dagger}\boldsymbol{v}_k, \quad w_k(s) = \ell(L^{-1}\mathbf{T}(s))^{\dagger}\boldsymbol{w}_k, \tag{5}$$

where $\mathbf{T}(s) = [T_1(s), T_2(s), \ldots, T_n(s)]$ is a vector of Chebyshev functions. Time varying linear combinations of $v_k(s)$ and $w_k(s)$ give the instantaneous centerline reconstruction (SI). We find that the zero-function $v_0(s)$ is close to the best fit straight line through the motion, accounting for 85% of the time-averaged centerline reconstruction while most of the oscillations are accounted for by the first excited-states $v_1(s)$ and $w_1(s)$ corresponding to the smallest magnitude non-zero eigenvalues (13.3%). Since most ($> 98\%$) of the dynamics is captured by the zero-state and first excited states, one can in fact further reduce the complexity of the Schrödinger model, by approximating H_1 through its projection $\hat{H}_1$ on the

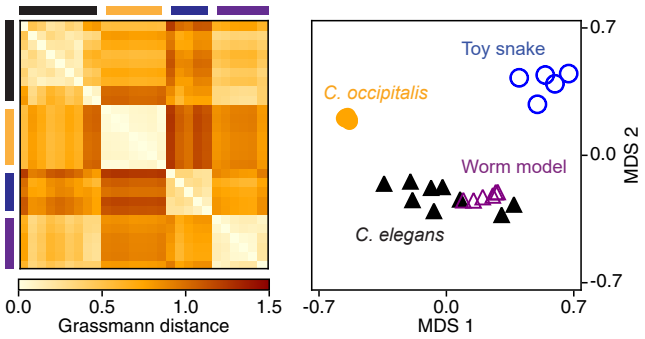

FIG. 4. Pairwise Grassmann distances between subspaces spanned by first excited eigenstates of the Hamiltonians and a corresponding phase diagram obtained by a multidimensional scaling (MDS) embedding capture the similarities and differences between undulatory locomotion in organisms, model simulations, and robots.

eigenspaces corresponding to the first two distinct eigenvalues. This additional low-rank approximation also further reduces the risk of overfitting and hence improves model generalizability, similar to sparsity promotion in other dynamical inference methods [54].

The compact low-rank characterization of the undulatory shape dynamics makes it possible to compare the locomotion behaviors of *C. elegans*, previously proposed neuro-mechanical worm models [33], *C. occipitalis* snakes [37], and robotic toy snakes, by measuring the Grassmann distance [55] between the dominant eigenspaces of $\hat{H}_1$. As most of the variation of the oscillatory dynamics is contained in the first excited-states $\boldsymbol{v}_1$ and $\boldsymbol{w}_1$, we determined the pairwise Grassmann distances between the eigenspaces spanned by $\boldsymbol{v}_1$ and $\boldsymbol{w}_1$ for the various systems (SI). Both the distance matrix and a corresponding 2D phase diagram constructed by multidimensional scaling reveal that the neuro-mechanical worm model [33] succeeds in reproducing key dynamical aspects of *C. elegans* locomotion, whereas the robotic toy snake used in our experiments is equally far from real snake or worm locomotion (Fig. 4).

From a practical perspective, the above results show how symmetry-constrained mode representations can facilitate a low-dimensional description and efficient classification of biophysical dynamics. The underlying inference framework is directly applicable to diagnose and quantify the effects of genetic or chemical perturbations on animal locomotion within and across species. From a theoretical perspective, the fact that translational and rotational invariance combined with a quadratic integral constraint generically lead to a Schrödinger equation [45] in mode space, promises advances in the quantitative understanding of biological systems, as the comprehensive toolbox of quantum physics [56, 57] now becomes available to characterize and predict behavioral dynamics.

ACKNOWLEDGEMENTS

A.E.C. and A.D.H. contributed equally to this work. This work was supported by a MathWorks Science Fellowship (A.D.H.), the JPB Foundation (S.W.F.), a Sloan Research Fellowship (S.W.F.), a McKnight Foundation Scholar Award (S.W.F.), NSF Award 1845663 (S.W.F.), Sloan Foundation Grant G-2021-16758 (J.D.), NSF Award DMS-1952706 (J.D.), and the Robert E. Collins Distinguished Scholarship Fund (J.D.).

* dunkel@mit.edu

[1] J. Gray, Undulatory propulsion, J. Cell Sci. **3**, 551 (1953).
[2] N. Cohen and J. H. Boyle, Swimming at low reynolds number: a beginners guide to undulatory locomotion, Contemp. Phys. **51**, 103 (2010).
[3] E. Niebur and P. Erdös, Theory of the locomotion of nematodes: dynamics of undulatory progression on a surface, Biophys. J. **60**, 1132 (1991).
[4] T. Majmudar, E. E. Keaveny, J. Zhang, and M. J. Shelley, Experiments and theory of undulatory locomotion in a simple structured medium, J. R. Soc. Interface **9**, 1809 (2012).
[5] A. Kudrolli and B. Ramirez, Burrowing dynamics of aquatic worms in soft sediments, Proc. Natl. Acad. Sci. U.S.A. **116**, 25569 (2019).
[6] J. Sznitman, P. K. Purohit, P. Krajacic, T. Lamitina, and P. E. Arratia, Material properties of Caenorhabditis elegans swimming at low reynolds number, Biophys. J. **98**, 617 (2010).
[7] X. N. Shen and P. E. Arratia, Undulatory swimming in viscoelastic fluids, Phys. Rev. Lett. **106**, 208101 (2011).
[8] A. J. Smits, Undulatory and oscillatory swimming, J. Fluid Mech. **874**, P1 (2019).
[9] J. G. Thewissen and F. Fish, Locomotor evolution in the earliest cetaceans: functional model, modern analogues, and paleontological evidence, Paleobiology **23**, 482 (1997).
[10] R. D. Maladen, Y. Ding, P. B. Umbanhowar, A. Kamor, and D. I. Goldman, Mechanical models of sandfish locomotion reveal principles of high performance subsurface sand-swimming, J. R. Soc. Interface **8**, 1332 (2011).
[11] Z. Guo and L. Mahadevan, Limbless undulatory propulsion on land, Proc. Natl. Acad. Sci. U.S.A. **105**, 3179 (2008).
[12] D. L. Hu, J. Nirody, T. Scott, and M. J. Shelley, The mechanics of slithering locomotion, Proc. Natl. Acad. Sci. U.S.A. **106**, 10081 (2009).
[13] S. W. Flavell, D. M. Raizen, and Y.-J. You, Behavioral states, Genetics **216**, 315 (2020).
[14] W. Hong, A. Kennedy, X. P. Burgos-Artizzu, M. Zelikowsky, S. G. Navonne, P. Perona, and D. J. Anderson, Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning, Proc. Natl. Acad. Sci. U.S.A. **112**, E5351 (2015).
[15] N. Pokala and S. W. Flavell, Recording and quantifying C. elegans behavior, in *C. elegans* (Springer, 2022) pp. 357–373.
[16] N. Cermak, K. Y. Stephanie, R. Clark, Y.-C. Huang, S. N. Baskoylu, and S. W. Flavell, Whole-organism behavioral profiling reveals a role for dopamine in state-dependent motor program coupling in C. elegans, Elife **9**, e57093 (2020).
[17] F. B. Shipley, C. M. Clark, M. J. Alkema, and A. M. Leifer, Simultaneous optogenetic manipulation and calcium imaging in freely moving C. elegans, Front. Neural Circuits **8**, 28 (2014).
[18] A. M. Leifer, C. Fang-Yen, M. Gershow, M. J. Alkema, and A. D. Samuel, Optogenetic manipulation of neural activity in freely moving Caenorhabditis elegans, Nat. methods **8**, 147 (2011).
[19] R. Prevedel, Y.-G. Yoon, M. Hoffmann, N. Pak, G. Wetzstein, S. Kato, T. Schrödel, R. Raskar, M. Zimmer, E. S. Boyden, *et al.*, Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy, Nat. Methods **11**, 727 (2014).
[20] K. Bozek, L. Hebert, Y. Portugal, A. S. Mikheyev, and G. J. Stephens, Markerless tracking of an entire honey bee colony, Nat. Commun. **12**, 1733 (2021).
[21] M. Liu, A. K. Sharma, J. W. Shaevitz, and A. M.

Leifer, Temporal processing and context dependency in Caenorhabditis elegans response to mechanosensation, Elife **7**, e36419 (2018).
[22] L. Hebert, T. Ahamed, A. C. Costa, L. O'Shaughnessy, and G. J. Stephens, Wormpose: Image synthesis and convolutional networks for pose estimation in C. elegans, PLoS Comput. Biol. **17**, e1008914 (2021).
[23] M. W. Mathis and A. Mathis, Deep learning tools for the measurement of animal behavior in neuroscience, Curr. Opin. Neurobiol. **60**, 1 (2020).
[24] T. D. Pereira, D. E. Aldarondo, L. Willmore, M. Kislin, S. S.-H. Wang, M. Murthy, and J. W. Shaevitz, Fast animal pose estimation using deep neural networks, Nat. Methods **16**, 117 (2019).
[25] A. Mathis, P. Mamidanna, K. M. Cury, T. Abe, V. N. Murthy, M. W. Mathis, and M. Bethge, Deeplabcut: markerless pose estimation of user-defined body parts with deep learning, Nat. Neurosci. **21**, 1281 (2018).
[26] G. J. Stephens, B. Johnson-Kerner, W. Bialek, and W. S. Ryu, Dimensionality and dynamics in the behavior of C. elegans, PLoS Comput. Biol. **4**, e1000028 (2008).
[27] A. C. Costa, T. Ahamed, and G. J. Stephens, Adaptive, locally linear models of complex dynamics, Proc. Natl. Acad. Sci. U.S.A. **116**, 1501 (2019).
[28] T. Ahamed, A. C. Costa, and G. J. Stephens, Capturing the continuous complexity of behaviour in Caenorhabditis elegans, Nat. Phys. **17**, 275 (2021).
[29] A. Hosoi and D. I. Goldman, Beneath our feet: strategies for locomotion in granular media, Annu. Rev. Fluid Mech. **47**, 431 (2015).
[30] Y. Ozkan-Aydin, D. I. Goldman, and M. S. Bhamla, Collective dynamics in entangled worm and robot blobs, Proc. Natl. Acad. Sci. U.S.A. **118**, e2010542118 (2021).
[31] G. J. Berman, D. M. Choi, W. Bialek, and J. W. Shaevitz, Mapping the stereotyped behaviour of freely moving fruit flies, J. R. Soc. Interface **11**, 20140672 (2014).
[32] E. Lauga and T. R. Powers, The hydrodynamics of swimming microorganisms, Rep. Prog. Phys. **72**, 096601 (2009).
[33] J. H. Boyle, S. Berri, and N. Cohen, Gait modulation in C. elegans: an integrated neuromechanical model, Front. Comput. Neurosci. **6**, 10 (2012).
[34] G. P. Sarma, C. W. Lee, T. Portegys, V. Ghayoomie, T. Jacobs, B. Alicea, M. Cantarelli, M. Currie, R. C. Gerkin, S. Gingell, *et al.*, Openworm: overview and recent advances in integrative biological simulation of Caenorhabditis elegans, Philos. Trans. R. Soc. Lond., B, Biol. Sci. **373**, 20170382 (2018).
[35] H. Marvi, C. Gong, N. Gravish, H. Astley, M. Travers, R. L. Hatton, J. R. Mendelson III, H. Choset, D. L. Hu, and D. I. Goldman, Sidewinding with minimal slip: Snake and robot ascent of sandy slopes, Science **346**, 224 (2014).
[36] J. Aguilar, T. Zhang, F. Qian, M. Kingsbury, B. McInroe, N. Mazouchova, C. Li, R. Maladen, C. Gong, M. Travers, *et al.*, A review on locomotion robophysics: the study of movement at the intersection of robotics, soft matter and dynamical systems, Rep. Prog. Phys. **79**, 110001 (2016).
[37] P. E. Schiebel, J. M. Rieser, A. M. Hubbard, L. Chen, D. Z. Rocklin, and D. I. Goldman, Mechanical diffraction reveals the role of passive dynamics in a slithering snake, Proc. Natl. Acad. Sci. U.S.A. **116**, 4798 (2019).
[38] A. Goldschmidt, E. Kaiser, J. L. Dubois, S. L. Brunton, and J. N. Kutz, Bilinear dynamic mode decomposition for quantum control, New J. Phys. **23**, 033035 (2021).
[39] P. J. Schmid, Dynamic mode decomposition of numerical and experimental data, J. Fluid Mech. **656**, 5 (2010).
[40] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, On dynamic mode decomposition: Theory and applications, J. Comput. Dyn. **1**, 391 (2014).
[41] N. Romeo, A. Hastewell, A. Mietke, and J. Dunkel, Learning developmental mode dynamics from single-cell trajectories, Elife **10**, e68679 (2021).
[42] Y. Ma, V. Dixit, M. J. Innes, X. Guo, and C. Rackauckas, A comparison of automatic differentiation and continuous sensitivity analysis for derivatives of differential equation solutions, in *2021 IEEE High Performance Extreme Computing Conference (HPEC)* (IEEE, 2021) pp. 1–9.
[43] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, A. Ramadhan, and A. Edelman, Universal differential equations for scientific machine learning, arXiv:2001.04385 [cs.LG] (2020).
[44] L. Sorber, M. Van Barel, and L. De Lathauwer, Structured data fusion, IEEE J Sel Top Signal Process **9**, 586 (2015).
[45] E. Schrödinger, An undulatory theory of the mechanics of atoms and molecules, Phys. Rev. **28**, 1049 (1926).
[46] D. J. Griffiths and D. F. Schroeter, *Introduction to quantum mechanics* (Cambridge university press, 2018).
[47] V. Kantsler and R. E. Goldstein, Fluctuations, dynamics, and the stretch-coil transition of single actin filaments in extensional flows, Phys. Rev. Lett. **108**, 038103 (2012).
[48] J. P. Boyd, *Chebyshev and Fourier spectral methods* (Courier Corporation, 2001).
[49] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic mode decomposition: data-driven modeling of complex systems* (SIAM, 2016).
[50] P. J. Baddoo, B. Herrmann, B. J. McKeon, J. N. Kutz, and S. L. Brunton, Physics-informed dynamic mode decomposition (pidmd), arXiv:2112.04307 (2021).
[51] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980 (2014).
[52] S. Wright, J. Nocedal, *et al.*, Numerical optimization, Springer **35**, 7 (1999).
[53] J. Zhuang, T. Tang, Y. Ding, S. C. Tatikonda, N. Dvornek, X. Papademetris, and J. Duncan, Adabelief optimizer: Adapting stepsizes by the belief in observed gradients, Adv. Neural Inf. Process. Syst. **33**, 18795 (2020).
[54] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. U.S.A. **113**, 3932 (2016).
[55] K. Ye and L.-H. Lim, Schubert varieties and distances between subspaces of different dimensions, SIAM J. Matrix Anal. Appl. **37**, 1176 (2016).
[56] R. E. Goldstein and D. M. Petrich, Solitons, Euler's equation, and vortex patch dynamics, Phys. Rev. Lett. **69**, 555 (1992).
[57] H. Hasimoto, A soliton on a vortex filament, J. Fluid Mech. **51**, 477 (1972).

Supplemental Materials

CONTENTS

DATA PREPROCESSING

Raw data from the experiments consist of video files of the animal/simulation/robot motion. The analysis process for this data follows closely to that described in [S16]. Each image frame is thresholded to obtain a binary image. These binary images are then dilated and eroded in order fill in holes while preserving object shape. The resulting image is then thinned to obtain a centerline. Finally, a set number of equally spaced points along the centerline is generated.

To obtain the Chebyshev mode coefficients, we obtain (x, y) positions at the Chebyshev points by linearly interpolating between the centerline points. We perform the transformation to a degree-19 Chebyshev polynomial, of which we use the mode coefficients up to degree-9 to represent the shape of the body.

In order to determine the dominant mode frequencies used in Figure 1c, we take the Fourier transform of modes $\hat{x}_k$ and $\hat{y}_k$ for $k = 1, ..., n$. We extract the frequency which has the largest power in the spectra, which we call the dominant frequency. We then average the dominant frequency for all modes and plot this average in Figure 1f.

TOY SNAKE EXPERIMENTS

The toy snake was manufactured by Top Race and purchased from Amazon. The toy snake was placed on green construction paper on a hardwood floor and was remotely controlled to move in a straight line. Videos were filmed on an iPhone 12 Pro resting on a surface 1 meter above the floor.

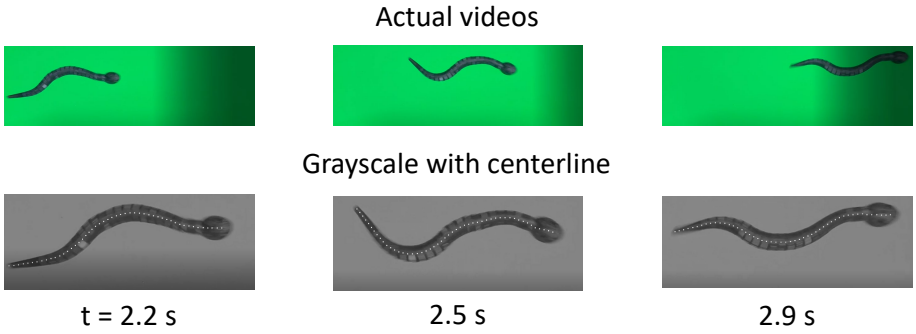

FIG. S1. (top) Images from the toy snake video used in the paper. (bottom) Corresponding grayscale images with points along the centerline overlayed.

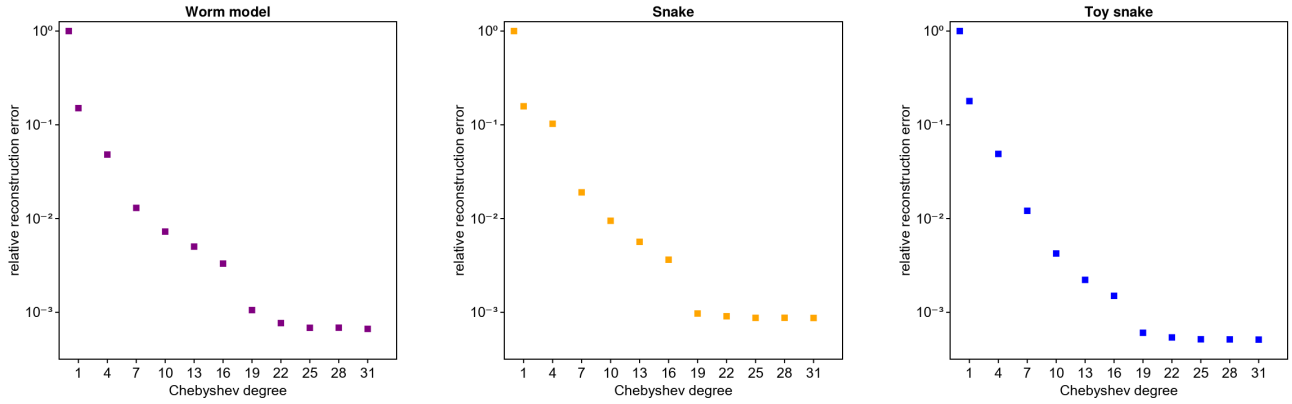

FIG. S2. Reconstruction error plots for the worm simulation, snake, and toy snake. The same plot for the worm is shown in Main Text Figure 1(c).

RECONSTRUCTION ERROR

The reconstruction error used in Figure 1c and Figure S2 is calculated using

$$\text{relative reconstruction error} = \frac{\sum_{t,s}\sqrt{(x_c - x_d)^2 + (y_c - y_d)^2}}{\sum_{t,s}\sqrt{(\hat{x}_0 - x_d)^2 + (\hat{y}_0 - y_d)^2}}, \tag{S1}$$

where x_c and y_c are the x and y points calculated from the Chebyshev approximation and x_d and y_d are the x and y points from the experimental data. This can be interpreted as a relative mean absolute error, where we scale relative to the distances between all worm points from the worm center of mass. Therefore, the error is a measure of the distance deviations from the worm CCOM (a close approximation to the COM) accounted for by the Chebyshev polynomial approximation.

INTERPRETATION OF COEFFICIENTS

Starting from the definition of the Chebyshev modes Eq. (1) we can express the $n = 1$ coefficients as,

$$\begin{bmatrix}\hat{x}_1(t)\\ \hat{y}_1(t)\end{bmatrix} = \frac{2}{\pi}\int_{-1}^{1} \mathrm{d}s\; w(s)\; s \begin{bmatrix}x(s,t)\\ y(s,t)\end{bmatrix} = \frac{2}{\pi}\int_{-1}^{1} \mathrm{d}s\; \frac{1}{w(s)} \begin{bmatrix}x_s(s,t)\\ y_s(s,t)\end{bmatrix} \tag{S2}$$

where we use the fact that $T_1(s) = s$, $\int \mathrm{d}s\; sw(s) = -(1-s^2)^{1/2} = -1/w(s)$ and integration by parts. The resulting expression is the $1/w$-weighted Chebyshev orientation.

CONSTRAINED LINEAR MODELS IN MODE SPACE

Starting from a vector field representation of a centerline, $\mathbf{r} = (x(t,s), y(t,s))$, where s is the arclength along the centerline and $\mathbf{r}$ is the Cartesian coordinate of the centerline at position s, we show how physical constraints impose structure on a linear model in mode space.

Rotational invariance

We expect that our model should be invariant under a rotation of the coordinate system given by

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Applying R to (1) tells us how the coefficients transform as well,

$$\begin{bmatrix} x'(s,t) \\ y'(s,t) \end{bmatrix} = \sum_{k=0}^{n} T_k(s) \begin{bmatrix} \hat{x}'_k(t) \\ \hat{y}'_k(t) \end{bmatrix}$$
$$R \begin{bmatrix} x(s,t) \\ y(s,t) \end{bmatrix} = \sum_{k=0}^{n} T_k(s)\, R \begin{bmatrix} \hat{x}_k(t) \\ \hat{y}_k(t) \end{bmatrix}$$

which implies that,

$$\begin{aligned} \begin{bmatrix} \hat{x}'_k(t) \\ \hat{y}'_k(t) \end{bmatrix} &= R \begin{bmatrix} \hat{x}_k(t) \\ \hat{y}_k(t) \end{bmatrix} \\ &= \begin{bmatrix} \hat{x}_k(t)\cos\theta - \hat{y}_k(t)\sin\theta \\ \hat{x}_k(t)\sin\theta + \hat{y}_k(t)\cos\theta \end{bmatrix}. \end{aligned}$$

Hence, the coefficient state vector $\mathbf{\Psi}$ transforms as $\mathbf{\Psi}' = (R \otimes \mathbb{I}_{n+1})\mathbf{\Psi}$ where $\otimes$ denotes the Kronecker product of two matrices and $\mathbb{I}_n$ is the $n \times n$ identity matrix. Applying this to a general linear dynamics for $\mathbf{\Psi}$, yields,

$$\dot{\mathbf{\Psi}}' = M\mathbf{\Psi}' \tag{S4a}$$
$$(R \otimes \mathbb{I}_{n+1})\dot{\mathbf{\Psi}} = M(R \otimes \mathbb{I}_{n+1})\mathbf{\Psi} \tag{S4b}$$

then equating M in (S4a) to (S4b) then yields the condition on M,

$$(R \otimes \mathbb{I}_{n+1})M = M(R \otimes \mathbb{I}_{n+1}) \tag{S4c}$$

Writing (S4c) out in block matrix form

$$\begin{bmatrix} \mathbb{I}_{n+1}\cos\theta & -\mathbb{I}_{n+1}\sin\theta \\ \mathbb{I}_{n+1}\sin\theta & \mathbb{I}_{n+1}\cos\theta \end{bmatrix} \begin{bmatrix} M_{xx} & M_{xy} \\ M_{yx} & M_{yy} \end{bmatrix} = \begin{bmatrix} M_{xx} & M_{xy} \\ M_{yx} & M_{yy} \end{bmatrix} \begin{bmatrix} \mathbb{I}_{n+1}\cos\theta & -\mathbb{I}_{n+1}\sin\theta \\ \mathbb{I}_{n+1}\sin\theta & \mathbb{I}_{n+1}\cos\theta \end{bmatrix}$$

and comparing the left and right hand sides, yields the following two constraints, $-M_{yx} = M_{xy} = \tilde{M}_2$ and $M_{xx} = \tilde{M}_{yy} = M_1$. Rotational invariance therefore enforces that M has the following block structure,

$$M = \begin{bmatrix} \tilde{M}_1 & \tilde{M}_2 \\ -\tilde{M}_2 & \tilde{M}_1 \end{bmatrix}. \tag{S5}$$

Translational invariance

We further expect that the model we learn should not depend on the origin of the coordinate system. The dynamics, therefore, should be invariant under $x' = x + c_x$ and $y' = y + c_y$. The coefficients in mode space transform

as, $\hat{x}_0' = \hat{x}_0 + c_x$, $\hat{y}_0' = \hat{y}_0 + c_y$ and the other coefficients are unchanged. As a result, the coefficient vector $\mathbf{\Psi}$ transforms as,

$$\mathbf{\Psi}' = \mathbf{\Psi} + \begin{bmatrix} c_x \boldsymbol{e}_1 \\ c_y \boldsymbol{e}_1 \end{bmatrix}, \tag{S6}$$

where $\boldsymbol{e}_1$ is the standard $n+1$ dimensional unit vector along the first dimension. The transformed dynamical equation using M from (S5) can be written as,

$$\begin{aligned} \dot{\mathbf{\Psi}}' &= M\mathbf{\Psi}' \\ \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \mathbf{\Psi}_x + c_x\boldsymbol{e}_1 \\ \mathbf{\Psi}_y + c_y\boldsymbol{e}_1 \end{bmatrix} &= \begin{bmatrix} \tilde{M}_1 & \tilde{M}_2 \\ -\tilde{M}_2 & \tilde{M}_1 \end{bmatrix}\begin{bmatrix} \mathbf{\Psi}_x + c_x\boldsymbol{e}_1 \\ \mathbf{\Psi}_y + c_y\boldsymbol{e}_1 \end{bmatrix} \\ \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \mathbf{\Psi}_x \\ \mathbf{\Psi}_y \end{bmatrix} &= \begin{bmatrix} \tilde{M}_1 & \tilde{M}_2 \\ -\tilde{M}_2 & \tilde{M}_1 \end{bmatrix}\begin{bmatrix} \mathbf{\Psi}_x \\ \mathbf{\Psi}_y \end{bmatrix} + \begin{bmatrix} \tilde{M}_1 & \tilde{M}_2 \\ -\tilde{M}_2 & \tilde{M}_1 \end{bmatrix}\begin{bmatrix} c_x\boldsymbol{e}_1 \\ c_y\boldsymbol{e}_1 \end{bmatrix}. \end{aligned}$$

For the dynamics to be invariant under arbitrary shifts $[c_x, c_y]$ we need $\tilde{M}_1\boldsymbol{e}_1 = \mathbf{0}$ and $\tilde{M}_2\boldsymbol{e}_1 = \mathbf{0}$, which implies that that the first columns of $\tilde{M}_1$ and $\tilde{M}_2$ must be all 0. This decouples the dynamics of the $k = 0$ modes from the rest of the modes. Splitting $\mathbf{\Psi}$ into a 0 mode vector $\boldsymbol{\psi}_0 = [\hat{x}_0, \hat{y}_0]$ and a higher mode vector $\hat{\boldsymbol{\psi}} = [\hat{x}_1, \ldots, \hat{x}_n, \hat{y}_1, \ldots, \hat{y}_n]$, the dynamics becomes,

$$\dot{\boldsymbol{\psi}}_0 = \begin{bmatrix} \hat{\boldsymbol{m}}_1^\dagger & \hat{\boldsymbol{m}}_2^\dagger \\ -\hat{\boldsymbol{m}}_2^\dagger & \hat{\boldsymbol{m}}_1^\dagger \end{bmatrix}\hat{\boldsymbol{\psi}} \tag{S7a}$$

$$\dot{\hat{\boldsymbol{\psi}}} = \begin{bmatrix} \hat{M}_1 & \hat{M}_2 \\ -\hat{M}_2 & \hat{M}_1 \end{bmatrix}\hat{\boldsymbol{\psi}} \tag{S7b}$$

with the same structure as before but new block elements.

Length constraint

For undulatory motion we have one additional constraint: the length of the centerline is approximately constant. The length of the centerline in terms of the real fields $x(s,t)$ and $y(s,t)$ is,

$$\ell(t) = \int_{-1}^{1} \mathrm{d}s\, \sqrt{x_s(s,t)^2 + y_s(s,t)^2}. \tag{S8}$$

To allow for a convenient representation of an approximate length constraint in mode space we consider ℓ^2. Using the Cauchy-Schwarz inequality, $\langle f, g\rangle^2 \leq \langle f, f\rangle\langle g, g\rangle$, with metric $\langle f, g\rangle = \int_{-1}^{1} \mathrm{d}s\, f \cdot g$, we can derive a convex upper-bound on the square length using $f = 1$ and $g = \sqrt{x_s(s,t)^2 + y_s(s,t)^2}$,

$$\begin{aligned} \ell^2 &= \left(\int_{-1}^{1} \mathrm{d}s\, 1 \cdot \sqrt{x_s(s,t)^2 + y_s(s,t)^2}\right)^2 \\ &\leq \left(\int_{-1}^{1} \mathrm{d}s\, 1^2\right)\left(\int_{-1}^{1} \mathrm{d}s\, [x_s(s,t)^2 + y_s(s,t)^2]\right) \\ \ell^2 &\leq 2\tilde{\ell}^2 \end{aligned} \tag{S9}$$

where we define a convex approximate square length $\tilde{\ell}^2$,

$$\tilde{\ell}^2 = \int_{-1}^{1} \mathrm{d}s\, [x_s(s,t)^2 + y_s(s,t)^2]. \tag{S10}$$

We can find an approximation for ℓ in terms of $\tilde{\ell}$ by considering the Taylor expansion of $f(a,b) = \sqrt{a+b}$ around a_0 and b_0. Any $k = n + m$th order derivative of $f(a,b)$, is given by,

$$\frac{\partial^n}{\partial a^n}\frac{\partial^m}{\partial b^m} f(a,b) = -\frac{(-1)^{n+m}}{2\sqrt{\pi}}\Gamma(n+m-1/2)\frac{1}{(a+b)^{n+m-1/2}}$$

The Taylor series then becomes,

$$\begin{aligned}\sqrt{a+b} &= \sqrt{a_0+b_0} + \frac{1}{2\sqrt{a_0+b_0}}((a+b)-(a_0+b_0)) - \sum_{k=2}^{\infty}\frac{(-1)^k\Gamma(k-1/2)}{2\sqrt{\pi(a_0+b_0)}k!}\left(\frac{a+b}{a_0+b_0}-1\right)^k \\ &= \frac{1}{2}\sqrt{a_0+b_0} + \frac{a+b}{2\sqrt{a_0+b_0}} - \sum_{k=2}^{\infty}\frac{(-1)^k\Gamma(k-1/2)}{2\sqrt{\pi(a_0+b_0)}k!}\left(\frac{a+b}{a_0+b_0}-1\right)^k. \end{aligned} \tag{S11}$$

Setting $a = x_s^2$ and $b = y_s^2$ we can expand around $a_0 = \langle x_s^2\rangle$ and $b_0 = \langle y_s^2\rangle$ in (S11), where we use $\langle\cdot\rangle$ to represent the average value over s and t

$$\langle f\rangle = \frac{1}{2T}\int_0^T \mathrm{d}t \int_{-1}^1 \mathrm{d}s\, f(s,t),$$

to get an expansion of the square-root term in (S8),

$$\sqrt{x_s^2+y_s^2} = \frac{1}{2}\sqrt{\langle x_s^2\rangle+\langle y_s^2\rangle} + \frac{x_s^2+y_s^2}{2\sqrt{\langle x_s^2\rangle+\langle y_s^2\rangle}} - \sum_{k=2}^{\infty}\frac{(-1)^k\Gamma(k-1/2)}{2\sqrt{\pi(\langle x_s^2\rangle+\langle y_s^2\rangle)}k!}\left(\frac{x_s^2+y_s^2}{\langle x_s^2\rangle+\langle y_s^2\rangle}-1\right)^k \tag{S12}$$

Integrating (S12) over s, we get the following relationship between ℓ and $\tilde{\ell}$,

$$\begin{aligned}\ell = \int_{-1}^1 \mathrm{d}s\,\sqrt{x_s^2+y_s^2} &= \sqrt{\langle x_s^2\rangle+\langle y_s^2\rangle} + \frac{1}{2\sqrt{\langle x_s^2\rangle+\langle y_s^2\rangle}}\int_{-1}^1 \mathrm{d}s\,\left(x_s^2+y_s^2\right) + R(\Delta) \\ \ell \approx \ell_a &= \sqrt{\langle x_s^2\rangle+\langle y_s^2\rangle} + \frac{1}{2\sqrt{\langle x_s^2\rangle+\langle y_s^2\rangle}}\tilde{\ell}^2 \end{aligned} \tag{S13}$$

where we define

$$\Delta = \frac{x_s^2+y_s^2}{\langle x_s^2\rangle+\langle y_s^2\rangle} - 1$$

a measure of how much the deviations vary from their average. We can get a bound for the magnitude of $R(\Delta)$ by evaluating the remaining summation and utilizing the triangle inequality,

$$\begin{aligned}|R(\Delta)| &= \left|\int_{-1}^1 \mathrm{d}s \sum_{k=2}^{\infty}\frac{(-1)^k\Gamma(k-1/2)}{2\sqrt{\pi(\langle x_s^2\rangle+\langle y_s^2\rangle)}k!}\Delta^k\right| \\ &\le \int_{-1}^1 \mathrm{d}s \sum_{k=2}^{\infty}\frac{\Gamma(k-1/2)}{2\sqrt{\pi(\langle x_s^2\rangle+\langle y_s^2\rangle)}k!}|\Delta|^k \\ &\le \int_{-1}^1 \mathrm{d}s \sum_{k=2}^{\infty}\frac{\Gamma(k-1/2)}{2\sqrt{\pi(\langle x_s^2\rangle+\langle y_s^2\rangle)}k!}\Delta_M^k \\ &= \frac{1}{\sqrt{\pi(\langle x_s^2\rangle+\langle y_s^2\rangle)}}\sum_{k=2}^{\infty}\frac{\Gamma(k-1/2)}{k!}\Delta_M^k \\ &= \frac{2-\Delta_M-2\sqrt{1-\Delta_M}}{\sqrt{\langle x_s^2\rangle+\langle y_s^2\rangle}} \end{aligned} \tag{S14}$$

provided that $\Delta_M \le 1$, where $\Delta_M = \max_{s,t}|\Delta|$. For example, for the worm data considered here, $\Delta_M = 0.48$ and $\sqrt{\langle x_s^2\rangle+\langle y_s^2\rangle} = 0.50$, which give a value of $|R(\Delta)| < 0.15$, resulting in very close agreement between the true value of ℓ and the approximation calculated using (S13) (Fig. S3). In practice the maximum error is much lower: the maximum deviation between the approximation and true value of ℓ for the worm data is 0.0084 with a corresponding maximum relative error of 0.0073. Unitless values of Δ_M, $\sqrt{\langle x_s^2\rangle+\langle y_s^2\rangle}$, $|R(\Delta)|$ bounds and the maximum calculated errors are shown in Table I, further highlighting the validity of this approximation across of all the experimental systems studied here.

Since ℓ_a is a function solely of $\tilde{\ell}^2$, keeping ℓ_a constant is the same as keeping $\tilde{\ell}^2$ constant. We, therefore, continue working under the assumption that $\tilde{\ell}^2$ is constant for undulatory motion.

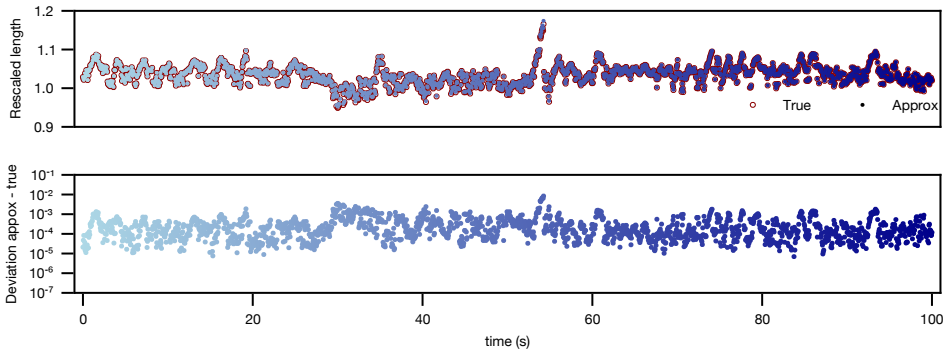

FIG. S3. (top) True length ℓ and approximate length ℓ_a calculated using (S13) show close agreement. (bottom) Deviation between $\ell_a - \ell$ shows that the true deviation is much lower than the bound 0.13. Also the deviation is always positive which means ℓ_a provides a close upper bound on ℓ.

Length constraint in mode-space

One of the benefits of working with $\tilde{\ell}^2$ rather than ℓ is that x and y appear quadratically, making it convenient to represent $\tilde{\ell}^2$ in mode space. We can express, $x_s(s,t)$ and $y_s(s,t)$ in mode space by differentiating (1),

$$\begin{bmatrix} x_s(s,t) \\ y_s(s,t) \end{bmatrix} = \sum_{k=0}^{n} \frac{\mathrm{d}}{\mathrm{d}s} T_k(s) \begin{bmatrix} \hat{x}_k(t) \\ \hat{y}_k(t) \end{bmatrix}. \tag{S15}$$

Substituting (S15) into the expression for $\tilde{\ell}^2$ (3) gives

$$\begin{aligned} \tilde{\ell}^2 &= \int_{-1}^{1} \mathrm{d}s \sum_{k,m=1}^{n} \left[\hat{x}_k(t)\hat{x}_m(t) + \hat{y}_k(t)\hat{y}_m(t)\right] \frac{\mathrm{d}T_k}{\mathrm{d}s}(s) \frac{\mathrm{d}T_m}{\mathrm{d}s}(s) \\ &= \sum_{k,m=1}^{n} \left[\hat{x}_k(t)\hat{x}_m(t) + \hat{y}_k(t)\hat{y}_m(t)\right] \int_{-1}^{1} \mathrm{d}s \, \frac{\mathrm{d}T_k}{\mathrm{d}s}(s) \frac{\mathrm{d}T_m}{\mathrm{d}s}(s) \\ &= \sum_{k,m=1}^{n} \left[\hat{x}_k(t)\hat{x}_m(t) + \hat{y}_k(t)\hat{y}_m(t)\right] W_{k,m} \\ &= \hat{\boldsymbol{\psi}}^{\dagger} \begin{bmatrix} W & 0 \\ 0 & W \end{bmatrix} \hat{\boldsymbol{\psi}}^{\dagger} \end{aligned} \tag{S16a}$$

where we define the symmetric matrix W with elements given by,

$$W_{k,m} = \int_{-1}^{1} \mathrm{d}s \, \frac{\mathrm{d}T_k}{\mathrm{d}s}(s) \frac{\mathrm{d}T_m}{\mathrm{d}s}(s). \tag{S16b}$$

System	Δ_M	$\sqrt{\langle x_s^2 \rangle + \langle y_s^2 \rangle}/\ell$	$\lvert R(\Delta) \rvert \times \ell$ bound	Maximum relative error
C. elegans	0.48	0.50	0.15	0.0073
Neuro-mechanical worm	0.42	0.50	0.12	0.0064
C. occipitalis	0.40	0.50	0.10	0.0066
Toy snake	0.41	0.50	0.11	0.026

TABLE I. Length approximation parameters for the systems studied here. The low relative errors highlight the validity of our relaxed length constraint.

Note that the values of W are fixed by the choice of basis. Therefore, W is a basis dependent constant. The matrix W has several advantageous properties: it is symmetric positive definite,

$$\mathbf{a}^\dagger W \mathbf{a} = \sum_{k,m=1}^{n} a_k W_{k,m} a_m = \int_{-1}^{1} \mathrm{d}s \left(\sum_{k=1}^{n} a_k \frac{\mathrm{d}T_k}{\mathrm{d}s}(s) \right) \left(\sum_{m=1}^{n} a_m \frac{\mathrm{d}T_m}{\mathrm{d}s}(s) \right) = \int_{-1}^{1} \mathrm{d}s \left(\sum_{k=1}^{n} a_k \frac{\mathrm{d}T_k}{\mathrm{d}s}(s) \right)^2 \geq 0, \tag{S17}$$

since we can interpret the summation in the parentheses as the derivative of a Chebyshev series representation of some function $f(s) = \sum_{k=0}^{n} a_k T_k(s)$. For the expression above to be 0, requires then that $f'(s) = 0$ for all s and, therefore, $f(s) = a_0 T_0(s)$, immediately implying that $\mathbf{a} = \mathbf{0}$ for equality in the above expression.

Making use of the following four Chebyshev polynomial identities,

$$\frac{\mathrm{d}}{\mathrm{d}s} T_0(s) = 0 \tag{S18a}$$

$$\frac{\mathrm{d}}{\mathrm{d}s} T_k(s) = k U_{k-1}(s) \tag{S18b}$$

$$U_n(s) U_m(s) = \sum_{k=0}^{2\min(n,m)} U_{|n-m|+2k}(s) \tag{S18c}$$

$$\int_{-1}^{1} \mathrm{d}s\, U_n(s) = \left. \frac{T_{n+1}}{n+1} \right|_{-1}^{1} = \frac{1}{n+1} - \frac{(-1)^{n+1}}{n+1} = \frac{1+(-1)^n}{n+1} \tag{S18d}$$

we can derive the values of $W_{m,n}$,

$$\begin{aligned}
W_{m,n} &= \int_{-1}^{1} \mathrm{d}s\, \frac{\mathrm{d}T_n}{\mathrm{d}s}(s) \frac{\mathrm{d}T_m}{\mathrm{d}s}(s) \\
&= nm \int_{-1}^{1} \mathrm{d}s\, U_{n-1}(s) U_{m-1}(s) \\
&= nm \int_{-1}^{1} \mathrm{d}s \sum_{k=0}^{\min(n-1,m-1)} U_{|n-m|+2k}(s) \\
&= nm \sum_{k=0}^{\min(n-1,m-1)} \int_{-1}^{1} \mathrm{d}s\, U_{|n-m|+2k}(s) \\
&= nm \sum_{k=0}^{\min(n-1,m-1)} \frac{1+(-1)^{|n-m|+2k}}{|n-m|+2k+1} \\
&= nm(1+(-1)^{|n-m|}) \sum_{k=0}^{\min(n,m)-1} \frac{1}{|n-m|+2k+1} \\
&= \begin{cases} 0 & \text{if } m-n \text{ even} \\ 2nm \sum_{k=|n-m|+1}^{n+m-1} \frac{1}{k} & \text{if } m-n \text{ even} \end{cases}
\end{aligned} \tag{S19}$$

which shows that W has a checkerboard pattern and is diagonally dominant (Fig. S4).

The basis dependent matrix W defines a hyperellipsoid that the mode vector $\hat{\boldsymbol{\psi}}$ lies on. Since W is symmetric, positive-definite it has a Cholesky factorization $W = LL^\dagger$, which we can use to define a new mode vector

$$\boldsymbol{\psi} = \frac{1}{\tilde{\ell}} \begin{bmatrix} L^\dagger & 0 \\ 0 & L^\dagger \end{bmatrix} \hat{\boldsymbol{\psi}}$$

that lies on the unit hypersphere. Since the matrix we apply is block-diagonal applying this transformation to S7 does not change the structure and we get the transformed equations,

$$\dot{\boldsymbol{\psi}}_0 = \begin{bmatrix} \boldsymbol{m}_1^\dagger & \boldsymbol{m}_2^\dagger \\ -\boldsymbol{m}_2^\dagger & \boldsymbol{m}_1^\dagger \end{bmatrix} \boldsymbol{\psi} \tag{S20a}$$

$$\dot{\boldsymbol{\psi}} = \begin{bmatrix} M_1 & M_2 \\ -M_2 & M_1 \end{bmatrix} \boldsymbol{\psi}. \tag{S20b}$$

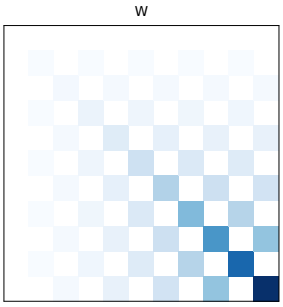

FIG. S4. W matrix for $n = 10$

Under this transformation the length constraint becomes,

$$\boldsymbol{\psi}^\dagger \boldsymbol{\psi} = 1. \tag{S20c}$$

Differentiating the new constraint with respect to time we get a dynamical constraint for length conservation,

$$\dot{\boldsymbol{\psi}}^\dagger \boldsymbol{\psi} + \boldsymbol{\psi}^\dagger \dot{\boldsymbol{\psi}} = 0$$
$$\boldsymbol{\psi}^\dagger \begin{bmatrix} M_1^\dagger + M_1 & M_2 - M_2^\dagger \\ M_2^\dagger - M_2 & M_1 + M_1^\dagger \end{bmatrix} \boldsymbol{\psi} = 0.$$

For this to be true for any $\boldsymbol{\psi}$, we must further constrain the real block-diagonal $M_1 = -M_1^\dagger = A$ to be skew-symmetric and the real block-off-diagonal $M_2 = M_2^\dagger = S$ to be symmetric. The final real form of our constrained dynamics is then given by,

$$\dot{\boldsymbol{\psi}}_0 = \begin{bmatrix} \boldsymbol{m}_1^\dagger & \boldsymbol{m}_2^\dagger \\ -\boldsymbol{m}_2^\dagger & \boldsymbol{m}_1^\dagger \end{bmatrix} \boldsymbol{\psi} \tag{S21a}$$

$$\dot{\boldsymbol{\psi}} = \begin{bmatrix} A & S \\ -S & A \end{bmatrix} \boldsymbol{\psi} \tag{S21b}$$

The full dynamical matrix has now been constrained to skew-symmetric and must, therefore, have purley imaginery eigenvalues. Therefore multiplying S21b by the imaginary unit i, results in our final complex Schrödinger type equations, that take the form of Schroödinger's equations in quantum mechanics for the shape dynamics,

$$\dot{\boldsymbol{\psi}}_0 = \begin{bmatrix} \boldsymbol{h}_1^\dagger & \boldsymbol{h}_2^\dagger \\ -\boldsymbol{h}_2^\dagger & \boldsymbol{h}_1^\dagger \end{bmatrix} \boldsymbol{\psi} \tag{S22a}$$

$$i\dot{\boldsymbol{\psi}} = \begin{bmatrix} H_1 & H_2 \\ -H_2 & H_1 \end{bmatrix} \boldsymbol{\psi} = H\boldsymbol{\psi} \tag{S22b}$$

when $H_1 = iA$ is a Hermitian matrix and $H_2 = iS$ is a skew-Hermitian matrix.

Straight motion

For straight motion we expect that the modes associated with x and the modes associated with y do not interact significantly leading to the further simplification that $H_2 = S = 0$ in the final constrained dynamics.

MODEL INFERENCE

Model inference: theory

The problem of learning an equation of the form (S21) from data for $\boldsymbol{\psi}$ and $\boldsymbol{\psi}_0$ at discrete time points $\{t_n\}_{n=0}^M$ can be formulated as a physics-informed dynamic mode decomposition (PI-DMD) optimization problem. In continuous

time the problem becomes,

$$\min_A \sum_{n=0}^{T} \|\dot{\boldsymbol{\psi}}(t_n) - A\boldsymbol{\psi}(t_n)\|_2^2 = \min_A \|\dot{P} - AP\|_F^2 \tag{S23}$$

where $P = [\boldsymbol{\psi}(t_0)\, \boldsymbol{\psi}(t_1)\, \boldsymbol{\psi}(t_2) \cdots \boldsymbol{\psi}(t_T)]$ is the matrix whose columns consists of the discrete time samples of $\boldsymbol{\psi}$. The minimization problem in (S23) has an analytical minimum in terms of the singular value decomposition (SVD) of P but requires numerically differentiating noisy data to calculate $\dot{P}$ an ill-posed and challenging problem. We, therefore, formulate the problem in discrete time. The general solution of a linear ODE of the form, $\dot{\mathbf{x}} = M\mathbf{x}$, is $\mathbf{x}(t) = \exp(At)\mathbf{x}(0)$. If the data are seperated by a constant time step Δt we can reformulate (S23) in the form,

$$\min_A \sum_{n=0}^{T-1} \left\|\boldsymbol{\psi}(t_{n+1}) - e^{A\Delta t}\boldsymbol{\psi}(t_n)\right\|_2^2 = \min_A \left\|P_{2:T} - e^{A\Delta t}P_{1:T-1}\right\|_F^2 \tag{S24}$$

where $P_{1:T-1}$ consists of the first $T-1$ columns of P and $P_{2:T}$ consists of the last $T-1$ columns. The skew-symmetric structure of the continuous time problem does not transfer to the discrete time problem, instead $\exp(A\Delta t)$ is an orthogonal matrix with a fixed form of its eigenvalues. Additionally, there is no guarantee that the matrix A will produce reintegrated trajectories close to the original input data. Exploiting the matrix exponential solution of linear ODEs we can modify (S24)

$$\min_A \sum_{m=1}^{T}\sum_{n=0}^{N} w_n \left[\boldsymbol{\psi}(t_m) - e^{At_m}\boldsymbol{\psi}(t_0)\right]_n^2 \tag{S25}$$

where we introduce the possibility of a weighting function w_n on the nth mode to account for the magnitude variations across the modes. The spectral theorem for skew-symmetric matrices tells us that a real $N \times N$ skew-symmetric matrix can be written in the form,

$$A = Q\Sigma Q^\dagger = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \boldsymbol{v}_1 & \boldsymbol{w}_1 & \boldsymbol{v}_2 & \boldsymbol{w}_2 & \cdots & \boldsymbol{v}_r & \boldsymbol{w}_r & \boldsymbol{v}_0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \Lambda_1 & & & & \\ & \Lambda_2 & & & \\ & & \ddots & & \\ & & & \Lambda_r & \\ & & & & 0 \end{bmatrix} \begin{bmatrix} \cdots & \boldsymbol{v}_1^\dagger & \cdots \\ \cdots & \boldsymbol{w}_1^\dagger & \cdots \\ \cdots & \boldsymbol{v}_2^\dagger & \cdots \\ \cdots & \boldsymbol{w}_2^\dagger & \cdots \\ & \vdots & \\ \cdots & \boldsymbol{v}_r^\dagger & \cdots \\ \cdots & \boldsymbol{w}_r^\dagger & \cdots \\ \cdots & \boldsymbol{v}_0^\dagger & \cdots \end{bmatrix}$$

where Q is a real orthogonal matrix $Q^\dagger Q = \mathbb{I}$, the Λ_i are 2×2 blocks matrices

$$\Lambda_i = \begin{bmatrix} 0 & \lambda_i \\ -\lambda_i & 0 \end{bmatrix}$$

with $r \leq \lfloor N/2 \rfloor$ the number of distinct complex conjugate pairs of eigenvalues and the remaining block is a $N - 2r \times N - 2r$ zero matrix. When N is odd there must always be a 0 row and column in Σ. The matrix exponential then has the simple form, $\exp(At) = Q\exp(\Sigma t)Q^\dagger$, where

$$\exp(\Sigma t) = \begin{bmatrix} \exp(\Lambda_1 t) & & & & \\ & \exp(\Lambda_2 t) & & & \\ & & \ddots & & \\ & & & \exp(\Lambda_r t) & \\ & & & & \mathbb{I} \end{bmatrix}$$

and

$$\exp(\Lambda_i t) = \begin{bmatrix} \cos\lambda_i t & \sin\lambda_i t \\ -\sin\lambda_i t & \cos\lambda_i t \end{bmatrix}.$$

The optimization loss function (S25) can then be written as,

$$\min_{Q,\{\lambda_i\}_{i=1}^r} \sum_{m=1}^{T} \sum_{n=0}^{N} w_n \left[\boldsymbol{\psi}(t_m) - Q\exp(\Sigma t_m) Q^\dagger \boldsymbol{\psi}(t_0)\right]_n^2 . \tag{S26}$$

Writing the formula in this way enables us to optimize or constrain the eigenvectors and eigenvalues of A separately. To optimize Q we follow the procedure in [S44], parameterizing the orthogonal matrix as the product of Householder matrices, $Q = H_1 H_2 \cdots H_N$, where each Householder matrix has the from

$$H_n = \begin{bmatrix} \mathbb{I}_{n-1} & 0 \\ 0 & \mathbb{I}_{N-n+1} - \frac{\mathbf{p}_n^\dagger \mathbf{p}_n}{|\mathbf{p}_n|^2} \end{bmatrix}$$

and can be further parameterized by a vector $\mathbf{p}_n$ of length $N - n + 1$. Fast in-place matrix vector multiplication algorithms exist for both Householder matrices and $\exp(\Sigma t)$ enabling us to efficiently compute the loss function from the combined parameter vector $\mathbf{p} = [\mathbf{p}_1, \cdots \mathbf{p}_N, \lambda_1, \cdots, \lambda_r]$. To minimize the loss function we calculate gradients using automatic differentiation and perform gradient descent using both the AdaBelief algorithm followed by BFGS.

Model inference: practice

The input data are the (x, y) positions along the centerline of an object undergoing undulatory motion. We begin by performing a Chebyshev transformation at each time step to transform the real space data as a function of time and arc length to mode data as a function of time, $\hat{\boldsymbol{\psi}}$. Next, we calculate the rescaled mode vectors $\boldsymbol{\psi}$ from the scaled mode vectors $\hat{\boldsymbol{\psi}}$.

Since we focus on learning dynamics across a timescale of one oscillation of the propulsive body wave, the next step is to determine the length of one oscillation from the starting time point. We first compute the FFT for a range of data starting at the specified time and find the maximum frequency amplitude. We then search near the period corresponding to the maximum frequency to locate the most similar shape to the initial shape, computed by finding the L2 norm of the difference between the shape modes. The minimum difference is then considered to be the end of the oscillation. We set the eigenvalues of the Hamiltonian that we will infer to be integer multiples of this oscillation frequency.

The model inference is performed on Equations (4). We first optimize the Hamiltonian with the eigenvalues constrained using the AdaBelief gradient-based optimization algorithm with forward-mode automatic differentation for gradient calculations. After this process, the BFGS algorithm is used to further optimize the Hamiltonian. After fitting the dynamics for the shape modes, the center of mass dynamics, described by $\mathbf{h}$, is optimized with the same optimization procedure used for the Hamiltonian.

Loss function weighting

As described in the paper, we utilize a loss function that consists of a combined mode space and real space loss. The real space loss $L_{\mathrm{RS}}(\mathbf{p}_H)$ calculates the mean square deviation between the CCOM subtracted field data $(x(s,t) - \hat{x}_0(t), y(s,t) - \hat{y}_0(t))$ and their prediction reconstructed from $\boldsymbol{\psi}_P$ calculated by integrating (4c) with $H_1(\mathbf{p}_{H_1})$. The mode space loss $L_{\mathrm{MS}}(\mathbf{p}_H)$ calculates the mean square deviation between the $\boldsymbol{\psi}_D$ calculated from the data and $\boldsymbol{\psi}_P$ normalized by the maximum standard deviations of all modes $\boldsymbol{\psi}_D$. The effect of the relative weighting of these losses is shown in Figure S5. As we increase the real space loss (move to the left on the x-axis), the centerline reconstruction error decreases (blue triangles). However, if we only have real space loss, the generalization error is very large (black circle in the top left corner), and the generalization error takes a minumum value for similar weighting of the two losses. Therefore, increasing the mode space loss promotes smaller generalization errors. In addition, while the low rank fit error is slightly larger than the full rank fit error (open versus closed triangles), the low rank generalization error is smaller than the full rank generalization error (open versus closed circles). These trends reveal that the mode space promotes generalizability while the real space loss promotes a good fit and the low rank model generalizes better than the full rank model. In the paper, we use a weight = 0.5, which is between the minimum generalization error weights for the full and low rank models.

ANALYTICAL SOLUTION TO THE MODEL

We can use the eigendecomposition of the final H_1 matrix to derive the analytical solution for $\boldsymbol{\psi}$ and $\mathbf{r}(s,t)-\mathbf{r}_0(t)$, where $\mathbf{r}_0 = \boldsymbol{\psi}_0 T_0(s) = \boldsymbol{\psi}_0$ represents the dynamics coming from the constant Chebyshev polynomial $T_0 = 1$ (CCOM). For a $(2K+1)\times(2K+1)$ dimensional H_1 matrix we will have 1 zero eigenvalue and a corresponding real eigenvector $\boldsymbol{\phi}_0$ and K opposite sign pairs of real eigenvalues and their corresponding complex conjugate eigenvectors, $\lambda_1^{\pm}, \lambda_2^{\pm}, \cdots, \lambda_K^{\pm}$ and $\boldsymbol{\phi}_1^{\pm}, \boldsymbol{\phi}_2^{\pm}, \cdots, \boldsymbol{\phi}_K^{\pm}$. The complex eigenvector can be decomposed into a real and imaginary part, $\boldsymbol{\phi}_k = \boldsymbol{v}_k + i\boldsymbol{w}_k$.

Using this formulation, the analytical solution corresponding to a single eigenvalue and eigenvector pair can be written as

$$\boldsymbol{s}_k^{\pm} = e^{i\lambda_k^{\pm} t}\boldsymbol{\phi}_k^{\pm}, \quad \boldsymbol{s}_0 = \boldsymbol{\phi}_0 = \boldsymbol{v}_0.$$

Using Euler's formula we get,

$$\boldsymbol{s}_k^{\pm} = (\cos\lambda_k^{\pm} t + i\sin\lambda_k^{\pm} t)\boldsymbol{\phi}_k^{\pm} = (\cos\lambda_k^{+} t \pm i\sin\lambda_k^{+} t)(\boldsymbol{v}_k \pm i\boldsymbol{w}_k),$$

which separates into real and imaginary parts

$$\boldsymbol{s}_k^{\pm} = (\cos\lambda_k^{+} t\boldsymbol{v}_k - \sin\lambda_k^{+} t\boldsymbol{w}_k) \pm i(\cos\lambda_k^{+} t\boldsymbol{w}_k + \sin\lambda_k^{+} t\boldsymbol{v}_k).$$

Defining $\tilde{\boldsymbol{v}}_k(t) = (\cos\lambda_k t\boldsymbol{v}_k - \sin\lambda_k t\boldsymbol{w}_k)$ and $\tilde{\boldsymbol{w}}_k(t) = (\cos\lambda_k t\boldsymbol{w}_k + \sin\lambda_k t\boldsymbol{v}_k)$ gives

$$\boldsymbol{s}_k^{\pm} = \tilde{\boldsymbol{v}}_k(t) \pm i\tilde{\boldsymbol{w}}_k(t).$$

Each pair of $\tilde{\boldsymbol{v}}_k(t)$ and $\tilde{\boldsymbol{w}}_k(t)$ are two linearly independent solutions so we can write the general solution as

$$\boldsymbol{s} = c_{0,1}\boldsymbol{v}_0 + \sum_{k=1}^{K} c_{k,1}\tilde{\boldsymbol{v}}_k(t) + c_{k,2}\tilde{\boldsymbol{w}}_k(t).$$

Since both $\boldsymbol{\psi}_x$ and $\boldsymbol{\psi}_y$ satisfy the same dynamical equation we get the general solution for $\boldsymbol{\psi}(t)$,

$$\boldsymbol{\psi}(t) = \begin{bmatrix}\boldsymbol{\psi}_x \\ \boldsymbol{\psi}_y\end{bmatrix} = \begin{bmatrix}c_{0,1}\boldsymbol{v}_0(t) \\ d_{0,1}\boldsymbol{v}_0(t)\end{bmatrix} + \sum_{k=1}^{K}\begin{bmatrix}c_{k,1}\tilde{\boldsymbol{v}}_k(t) + c_{k,2}\tilde{\boldsymbol{w}}_k(t) \\ d_{k,1}\tilde{\boldsymbol{v}}_k(t) + d_{k,2}\tilde{\boldsymbol{w}}_k(t)\end{bmatrix}. \tag{S27}$$

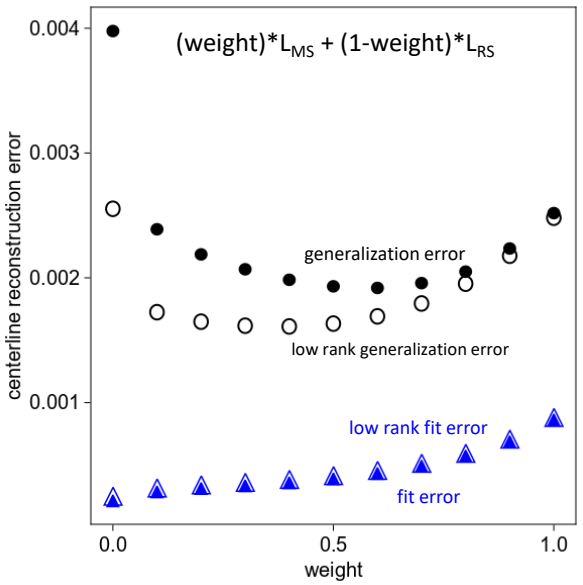

FIG. S5. Centerline reconstruction error for models trained with different weightings of the real space and mode space loss in the combined loss function. Generalization error is the centerline reconstruction error after simulating the model on an initial condition different from the initial condition used in fitting. Fit error is the centerline reconstruction error when simulating the model on the initial condition from fitting.

We now convert Eq. (S27) to real space. First, we undo the L scaling to go from the hypersphere back to the hyperellipsoid,

$$\hat{\boldsymbol{\psi}}(t) = \begin{bmatrix}\hat{\boldsymbol{\psi}}_x \\ \hat{\boldsymbol{\psi}}_y\end{bmatrix} = \ell \begin{bmatrix} c_{0,1}(L^{-1})^\dagger \boldsymbol{v}_0(t) \\ d_{0,1}(L^{-1})^\dagger \boldsymbol{v}_0(t)\end{bmatrix} + \ell \sum_{k=1}^{K} \begin{bmatrix} c_{k,1}(L^{-1})^\dagger \tilde{\boldsymbol{v}}_k(t) + c_{k,2}(L^{-1})^\dagger \tilde{\boldsymbol{w}}_k(t) \\ d_{k,1}(L^{-1})^\dagger \tilde{\boldsymbol{v}}_k(t) + d_{k,2}(L^{-1})^\dagger \tilde{\boldsymbol{w}}_k(t)\end{bmatrix} \tag{S28}$$

then the real space solution can be written for $x(s,t)$ as,

$$x(s,t) = \sum_{n=0}^{N} \hat{x}_n(t) T_n(s) = \hat{x}_0(t) + \mathbf{T}(s)^\dagger \hat{\boldsymbol{\psi}}_x(t)$$

where we use the fact that $T_0(s) = 1$ and define the Chebyshev vector $\mathbf{T}(s) = [T_1(s),\, T_2(s),\, \cdots,\, T_N(s)]$. Substituting in the solution (S28) we get,

$$x(s,t) - \hat{x}_0(t) = \ell \boldsymbol{T}(s)^\dagger (L^{-1})^\dagger \left[c_{0,1}\boldsymbol{v}_0 + \sum_{k=1}^{K} c_{k,1}\tilde{\boldsymbol{v}}_k + c_{k,2}\tilde{\boldsymbol{w}}_k \right]$$

using the definition of $\tilde{\boldsymbol{v}}_k$ and $\tilde{\boldsymbol{w}}_k$ the equation becomes,

$$= \ell \left[L^{-1}\mathbf{T}(s)\right]^\dagger \left[c_{0,1}\boldsymbol{v}_0 + \sum_{k=1}^{K} c_{k,1}(\cos(\lambda_k^+ t)\boldsymbol{v}_k - \sin(\lambda_k^+ t)\boldsymbol{w}_k) + c_{k,2}(\cos(\lambda_k^+ t)\boldsymbol{w}_k + \sin(\lambda_k^+ t)\boldsymbol{v}_k) \right]$$

$$x(s,t) - \hat{x}_0(t) = \ell \left[L^{-1}\mathbf{T}(s)\right]^\dagger \left[c_{0,1}\boldsymbol{v}_0 + \sum_{k=1}^{K} \cos(\lambda_k^+ t)(c_{k,1}\boldsymbol{v}_k + c_{k,2}\boldsymbol{w}_k) + \sin(\lambda_k^+ t)(c_{k,2}\boldsymbol{v}_k - c_{k,1}\boldsymbol{w}_k) \right],$$

finally we can write this in terms of the real space eigenfunctions $v_k(s) = \mathbf{T}^\dagger(s)(L^{-1})^\dagger \boldsymbol{v}_k$ and $w_k(s) = \mathbf{T}^\dagger(s)(L^{-1})^\dagger \boldsymbol{w}_k$

$$x(s,t) - \hat{x}_0(t) = \ell \left[c_{0,1}v_0(s) + \sum_{k=1}^{K} \cos(\lambda_k^+ t)(c_{k,1}v_k(s) + c_{k,2}w_k(s)) + \sin(\lambda_k^+ t)(c_{k,2}v_k(s) - c_{k,1}w_k(s)) \right], \tag{S29a}$$

a similar calculation for $y(s,t)$ yields,

$$y(s,t) - \hat{y}_0(t) = \ell \left[d_{0,1}v_0(s) + \sum_{k=1}^{K} \cos(\lambda_k t)(d_{k,1}v_k(s) + d_{k,2}w_k(s)) + \sin(\lambda_k t)(d_{k,2}v_k(s) - d_{k,1}w_k(s)) \right]. \tag{S29b}$$

Now consider the state where we enforce that the $\lambda_k = k\lambda$ are integer multiples of some base frequency λ

$$x(s,t) - \hat{x}_0(t) = \ell \left[c_{0,1}v_0(s) + \sum_{k=1}^{K} \cos(k\lambda t)(c_{k,1}v_k(s) + c_{k,2}w_k(s)) + \sin(k\lambda t)(c_{k,2}v_k(s) - c_{k,1}w_k(s)) \right]. \tag{S30}$$

We define the time-average square deviation from $\hat{x}_0(t)$ as

$$\left\langle (x(s,t) - \hat{x}_0(t))^2 \right\rangle_t (s) = \frac{\lambda}{2\pi} \int_0^{\frac{2\pi}{\lambda}} \mathrm{d}t\, (x(s,t) - \hat{x}_0(t))^2 .$$

Noting the orthogonality of sin and cos means that the only trigonometric functions that have non-zero time average are of the form $\cos(l\lambda t)^2$ and $\sin(l\lambda t)^2$, the can substitute (S30) into the definition for the squared deviation and only keep non-zero terms,

$$\begin{aligned}\left\langle (x(s,t) - \hat{x}_0(t))^2 \right\rangle_t (s) = \frac{\ell^2\lambda}{2\pi} \int_0^{\frac{2\pi}{\lambda}} \mathrm{d}t\, c_{0,1}^2 v_0(s)^2 &+ \sum_{k=1}^{K} \left[(c_{k,1}v_k(s) + c_{k,2}w_k(s))\right]^2 \cos(k\lambda t)^2 \\ &+ \sum_{k=1}^{K} \left[c_{k,2}v_k(s) - c_{k,1}w_k(s)\right]^2 \sin(k\lambda t)^2 \\ = \ell^2 c_{0,1}^2 v_0(s)^2 + \frac{\ell^2}{2} \sum_{k=1}^{K} \left(c_{k,1}^2 + c_{k,2}^2\right) &\left(v_k(s)^2 + w_k(s)^2\right). \end{aligned} \tag{S31a}$$

Similarly we have for $y(s,t)$ we have,

$$\left\langle (y(s,t)-\hat{y}_0(t))^2 \right\rangle_t (s) = \ell^2 d_{0,1}^2 v_0(s)^2 + \frac{\ell^2}{2}\sum_{k=1}^{K}\left(d_{k,1}^2+d_{k,2}^2\right)\left(v_k(s)^2+w_k(s)^2\right). \tag{S31b}$$

This means that the contribution to the mean square deviation corresponding to each eigenvalue takes the form of a spatial density $\rho_k(s) = v_k(s)^2 + w_k(s)^2$ with corresponding weight $c_{0,1}^2 + d_{0,1}^2$ for $k=0$ and $c_{k,1}^2 + c_{k,2}^2 + d_{k,1}^2 + d_{k,2}^2$ for $k \geq 1$.

We can readily extend this time average to a square deviation from some time-constant line $l(s)$,

$$\left\langle (x(s,t)-\hat{x}_0(t)-l(s))^2 \right\rangle_t (s) = \left(\ell c_{0,1} v_0(s) - l(s)\right)^2 + \frac{\ell^2}{2}\sum_{k=1}^{K}\left(c_{k,1}^2+c_{k,2}^2\right)\left(v_k(s)^2+w_k(s)^2\right). \tag{S31c}$$

For example, if $l(s) = \ell c_{0,1} v(s)$, an approximation to the center line of the worm, the first term vanishes meaning the dominant term in the deviation is,

$$\left\langle (x(s,t)-\hat{x}_0(t)-\ell c_{0,1} v_0(s))^2 \right\rangle_t (s) \approx \frac{\ell^2}{2}\left(c_{1,1}^2+c_{1,2}^2\right)\left(v_1(s)^2+w_1(s)^2\right)$$

Note since $\boldsymbol{v}_k$ and $\boldsymbol{w}_k$ are the columns or our orthogonal matrix Q in the spectral decomposition $\boldsymbol{v}_k^\dagger \boldsymbol{v}_l = \delta_{l,k}$, $\boldsymbol{w}_k^\dagger \boldsymbol{w}_l = \delta_{l,k}$ and $\boldsymbol{v}_k^\dagger \boldsymbol{w}_l = 0$. This implies that,

$$\begin{aligned}\int_{-1}^{1} w(s) v_l(s) v_k(s) &= \sum_{n=1}^{N}\sum_{m=1}^{N}\left[(L^{-1})^\dagger \boldsymbol{v}_k\right]_n \left[(L^{-1})^\dagger \boldsymbol{v}_l\right]_m \int_{-1}^{1} \mathrm{d}s\, w(s) T_n(s) T_m(s) \\ &= \frac{\pi}{2}\sum_{n=1}^{N}\left[(L^{-1})^\dagger \boldsymbol{v}_k\right]_n \left[(L^{-1})^\dagger \boldsymbol{v}_l\right]_n \end{aligned} \tag{S32}$$

GRASSMANN DISTANCE

To calculate the distances between the subspaces spanned by the real and imaginary parts of the eigenvector corresponding to the smallest nonzero eigenvalues (Fig. 4), we use the Grassmann distance. The Grassmann distance between two subspaces can be calculated by

$$d_G(A,B) = \sqrt{\sum_i \theta_i^2}, \tag{S33}$$

where A and B are two matrices whose columns are an orthonormal basis of their respective subspaces and θ_i are the principal angles between A and B [S55]. The principal angles can be calculated through an SVD, where the singular values of $A^\top B$ are $\sigma_i = \cos(\theta_i)$.