

Capturing the Continuous Complexity of Behavior in *C. elegans*

Tosif Ahamed^a, Antonio C. Costa^b, Greg J. Stephens^{a,b}

^a*Biological Physics Theory Unit,
OIST Graduate University, Okinawa 904-0495, Japan*

^b*Department of Physics and Astronomy,
Vrije Universiteit Amsterdam,
1081HV Amsterdam, The Netherlands*

Animal behavior is often quantified through subjective, incomplete variables that may mask essential dynamics. Here, we develop a behavioral state space in which the full instantaneous state is smoothly unfolded as a combination of short-time posture dynamics. Our technique is tailored to multivariate observations and extends previous reconstructions through the use of maximal prediction. Applied to high-resolution video recordings of the roundworm *C. elegans*, we discover a low-dimensional state space dominated by three sets of cyclic trajectories corresponding to the worm’s basic stereotyped motifs: forward, backward, and turning locomotion. In contrast to this broad stereotypy, we find variability in the presence of locally-unstable dynamics, and this unpredictability shows signatures of deterministic chaos: a collection of unstable periodic orbits together with a positive maximal Lyapunov exponent. The full Lyapunov spectrum is symmetric with positive, chaotic exponents driving variability balanced by negative, dissipative exponents driving stereotypy. The symmetry is indicative of damped, driven Hamiltonian dynamics underlying the worm’s movement control.

INTRODUCTION

Animals move in a wide variety of ways; the complex posture dynamics generating these behaviors span multiple spatiotemporal scales, and exhibit both regularity and variability [1, 2]. At large scales, behavior is structured, organized into stereotyped motifs such as walking or running, but the dynamics within each motif can be highly irregular [3, 4]. This complexity is apparent in spontaneous behaviors [5, 6], but also in highly stereotyped sequences such as an “escape response” [7], which must also be unpredictable for successful avoidance from motile predators [8]. Despite the importance of behavior in fields ranging from neuroscience [9, 10], ethology [11, 12], control theory [13], robotics and artificial intelligence [14] to the physics of living systems [15], the complexity of movement presents unique challenges in quantification, analysis, and understanding.

Technological advances, including recent progress in machine vision [16, 17], now make it possible to gather high-resolution movement data, even in complex, naturalistic settings and for animals with intricate body plans [18–20]. But how do we map high-resolution recordings of animal behavior into a compressed set of interpretable numbers while retaining maximal information about the dynamics? Indeed, among biological signals, behavior exhibits a remarkable divergence of descriptions, from representations based on pixels and wavelets (see e.g. [21]), to postures (see e.g. [22]) to more abstract states (see e.g. [23, 24]). Certainly, a good representation should capture the difference between distinct movement patterns. An ideal representation will also allow near-future predictions and be interpretable so as to provide insight into movement control principles. Finally, we seek to

reveal rather than impose on the structure of the behavioral signal, letting the representation and analysis guide important characteristics such as continuous vs. discrete, variable vs. stereotyped and spontaneous vs. controlled.

We detail the construction and application of a *behavioral state space* inspired by the similar approach of dynamical systems (also known as a phase space [25, 26], not to be confused with “state-space models” in statistics [27]). A point in our generally multidimensional behavioral state space represents the complete, near-instantaneous movements of an animal: posture and short-time posture changes. As time evolves, the state-space point follows a smooth trajectory, thus providing a geometrical encoding of behavior. Combining dynamical systems theory with high-resolution posture time series of the nematode *C. elegans*, we exploit the detailed structure of these trajectory encodings to seek a new quantitative perspective of ethological analysis [28].

STATE SPACE RECONSTRUCTION BY MAXIMIZING PREDICTABILITY

We consider a d -dimensional time series $\vec{y}(t)$ of duration T collected in a $T \times d$ matrix Y , which represents noisy, incomplete measurements of an underlying dynamical system, Fig. 1. With a state space reconstruction, we seek a coordinate transformation Ψ that maps Y into a space X that is topologically equivalent [29] to the state space of the underlying dynamical system, a process known as time series embedding [30, 31]. Dynamical embeddings have been used to model complex phenomena such as ecological and neural dynamics [32, 33], and to characterize the stability and symmetry of their re-

constructed attractors [34]. Although early approaches primarily used single-variable measurements, multivariate embeddings provide better reconstructions [35] and can improve prediction [36].

In our approach, we first lift the d -dimensional measurements into a Kd dimensional space of K contiguous delays and then project to a smaller m -dimensional subspace. Formally we decompose the embedding $\Psi = P_m \circ \Phi_K$ into a delay map Φ_K in which we iteratively stack $(K-1)$ -delayed copies of Y into a $(T-K+1) \times Kd$ matrix \bar{Y}_K , followed by a dimensionality reduction transformation P_m which projects \bar{Y}_K onto an $m < Kd$ dimensional space. P_m can in principle be any transformation and examples include numerical derivatives [37], delay coordinates [38] and random projections [33]. Here, we use singular value decomposition (SVD) [31, 35] followed by independent components analysis (ICA) [39], which results in a state space with independent components spanning the dimensions of the first m singular vectors. In matrix notation $X_m = \bar{Y}_K \Gamma_m$, where Γ_m is the $Kd \times m$ matrix of basis vectors spanning the m dimensional state space, while X_m contains the state space trajectories. This space of transformations allows for both derivative and more general linear filters [40] and the resulting coordinates reflect the most significant linear modes of the dynamics [31, 40].

The reconstruction is parameterized by the window length K and the state space dimension m and we describe a new, principled procedure for determining (K, m) based on optimal prediction. Notably, embedding parameters have often been chosen heuristically (see e.g. [31, 40]). To predict future observations we use N_b nearest neighbors in the reconstructed state space, Fig. 2A(left). To compute $\vec{y}_{\text{est}}(t + \tau)$, the τ -step prediction of $\vec{y}(t)$, we average the future of the nearest neighbors of the corresponding state space point so that $\vec{x}_{\text{est}}(t + \tau) = \langle \vec{x}(t + \tau) \rangle_{N_b}$ and then apply Φ_K^{-1} to pull \vec{x}_{est} back to observation space. This is known as the nearest neighbor predictor and also as Lorenz’s “method of analogs” [41]. The nearest neighbor predictor provides a lower bound to the predictability of a state space reconstruction as it is equivalent to a zeroth order Taylor approximation of the dynamics in a local neighborhood.

We quantify the prediction quality after τ steps using the error

$$E(\tau) = \sum_{j=1}^d \langle (y_j(t' + \tau) - (y_{\text{est}})_j(t' + \tau))^2 \rangle_{t'}^{1/2} \quad (1)$$

as shown in Fig. 2A(middle). Although $E(\tau)$ is a function we seek a single scalar that captures overall predictability. For a completely predictable system $E(\tau)$ is constant with a value corresponding to the noise level in the observations. On the other hand, for systems where predictions get worse over time, $E(\tau)$ grows according to a non-trivial process, possibly involving multiple timescales [41–43], shown schematically in Fig. 2A(right).

As long as the system is stationary, the error is bounded by the maximum distance within the state space, denoted e_s ; $E(\tau)$ grows until it saturates to e_s as $\tau \rightarrow \infty$, at which time the predictions are as good as choosing randomly from the sampled state space. We use the cumulative difference between the early-time and asymptotic error to define T_{pred} as a new measure of predictability,

$$T_{\text{pred}} = \frac{1}{e_s} \int_0^\infty (e_s - E(\tau)) d\tau = \frac{\Delta}{e_s} \quad (2)$$

where Δ is the area between the curve $E(\tau)$ and the asymptote e_s . A state space reconstruction with a large value of T_{pred} is good in the sense that it allows us to predict future observations for as long as possible. Although several previous studies about state space reconstruction are based on prediction as a guiding principle, they have used the predictive error in a more *ad hoc* manner, either by setting τ to a specific value [30, 32, 44, 45], or by integrating $E(\tau)$ to a chosen time τ_0 [46].

The average prediction error for an arbitrary time τ' is $\langle E(\tau') \rangle = \frac{1}{\tau'} \int_0^{\tau'} E(\tau) d\tau$. At large enough τ' , the error $E(\tau)$ approaches e_s and we can write $\langle E(\tau') \rangle = e_s - \Delta/\tau'$. Thus, the average prediction error is reduced from its asymptotic limit by an amount given by Δ/τ' . T_{pred} is also the characteristic timescale for a ball of points to randomize according to the state space density.

We demonstrate our embedding approach on a noisy measurement of a single coordinate of the Lorenz system (Methods: Lorenz System) and display the results in Fig. 2B-E. We find that T_{pred} increases with K for $K < 25$ frames after which it decreases gradually and we choose $K^* = 25$. We project \bar{Y}_{K^*} on the first m singular vectors and find that T_{pred} decreases after $m^* = 3$.

THE LOW-DIMENSIONAL STATE SPACE OF *C. ELEGANS* LOCOMOTION

We leverage our state space reconstruction to elucidate the behavior of the nematode *C. elegans* freely-foraging on a flat agar plate [47, 48]. In 2D, worms move by making dorsoventral sinusoidal bends along their body [49, 50], which can be captured through high resolution tracking microscopy to give a multidimensional time-series of posture changes [51]. Despite the variety of visible postures, most of the shape variation is captured by a linear combination of a small number of primitive shape dimensions (eigenworms) [22, 48], Fig. 3A.

Projections along the eigenworm dimensions describe the worm’s instantaneous shape and are not a direct indication of behavior, which arises from posture changes. Dynamical representations based on derivatives [22, 47, 52], and on sequences of postures [9, 53, 54] have been used to quantitatively explore the worm’s behavior. Importantly, the low-dimensionality of the worm’s shape space doesn’t imply simplicity and low-dimensionality of

the behavioral dynamics, and there are several signs of complexity in *C. elegans* behavior, such as heavy-tailed distributions [53], hierarchical structure in posture sequences [55, 56], indications of dynamical criticality in local linear approximation of the dynamics [52], as well as simultaneous presence of stereotypy and variability in posture sequences [22, 47, 53].

To reconstruct the state space of the worm’s posture dynamics, we start with a $T \times 5$ measurement matrix Y consisting of 5 eigenworm coefficients for a recording of duration $T = 33600$ frames (sampled at 16 Hz), Fig. 3A. We stack $(K - 1)$ time-shifted copies of Y to give the $(T - K + 1) \times 5K$ state matrix \bar{Y}_K . To estimate the optimal window size, we compute T_{pred} for each choice of K , as shown in Fig. 3B for a single representative worm, and choose $K^* = 12$. Within this window, we find that predictability saturates with $m = 7$ singular vectors, Fig. 3C. Analysis of each worm in the foraging dataset reveals a similar simplicity, Fig. S1. Despite its observed complexity, worm behavior is characterized by a low-dimensional state space.

We increase the interpretability of the worm’s state space reconstruction through a final transformation to independent components. We use the FastICA algorithm [39] on the $m = 7$ projections of the delay matrix \bar{Y}_{K^*} to obtain independent coordinate directions and we denote these coordinates behavioral modes $\vec{\Gamma}$. We show the seven behavioral modes in Fig. 3D as curvature kymographs and note that they come in three groups, broadly corresponding to the three coarse categories of worm movement: forward, backward and turning locomotion. Specifically, Γ_{f1} and Γ_{f2} modes correspond to the ventrally and dorsally initiated anterior-posterior body waves that worms make during forward locomotion. The reversal modes Γ_{r1} and Γ_{r2} capture the posterior-anterior body waves worms make during backward locomotion. Finally, $\{\Gamma_{t1}, \Gamma_{t2}, \Gamma_{t3}\}$ correspond to longer-ranged body bends. Large projections along Γ_{t1} and Γ_{t2} correspond to bends directed towards the ventral or dorsal direction respectively during a delta-turn like bend [48], while Γ_{t3} corresponds to an Omega-turn like bend. In this representation, worm locomotion is approximated by linearly combining these modes with time-varying amplitudes. We find similar modes for different choices of m^* (Fig. S2) and also for an ensemble embedding constructed by concatenating all $N = 12$ foraging organisms, Figs. S3-S4. We note that the behavioral modes emerge in an unsupervised manner, with no prior information of the worm’s movement.

The topology and geometry of trajectories in the behavioral state space contain important qualitative and quantitative information about worm behavior. A 10-min trajectory is visualized in Fig. 3E as projections onto the three mode combinations described above. In the (X_{f1}, X_{f2}) , and (X_{r1}, X_{r2}) planes, trajectories are colored by the centroid velocity of the worm, negative for backward locomotion and positive for forward loco-

tion. Trajectories in the (X_{t1}, X_{t2}, X_{t3}) space are colored by the mean body curvature. Large excitations in each of the three projections correspond to forward, backward and turning locomotion respectively. Specifically, trajectories in the (X_{f1}, X_{f2}) plane form a prominent circular band indicating nearly constant amplitude body waves during forward locomotion. Reversals emerge as trajectories spiraling from the center to a maximum radius in (X_{r1}, X_{r2}) plane, and then spiraling in as a reversal ends. Finally, deep body bends are represented as large transient orbits, with ventral turns and dorsal turns on opposite sides. Wild type worms have a ventral bias in their deep body bends, which is visible in the state space as a greater density of orbits on one side of 3D projection.

The state space also captures relationships between different body wave patterns. For example, we find that most reversals transition to forward by way of a deep ventral bend (Fig. S5), an observation that was previously reported in the context of the escape response and pirouette reorientation sequence [57, 58]. To quantify the relative activity of each set of body waves and the phase relationships between them, we define normalized mode amplitudes, $A_i = \frac{\bar{X}_i \cdot \bar{X}_i}{\bar{X} \cdot \bar{X}}$, where $i \in \{f, r, t\}$. The A_i range from 0 to 1 and measure the relative activity of different body wave patterns. We use these amplitudes to examine the behavior of $N = 92$ on-food worms where a brief laser impulse is applied to the head, resulting in a localized thermal stimulus provoking an escape response [48, 59], shown schematically in Fig. 3F. We project the posture dynamics of each stimulated worm onto the ensemble modes (Fig. S3) and show the normalized mode amplitudes averaged across all worms, Fig. 3G. The amplitudes capture the timescales and phase relationships between different body wave patterns during an escape sequence. In particular, the turning modes are strongly suppressed after the initiation of the reversal, increasing gradually as the reversal ends and worms transition into a turn. The turning amplitude then decreases, while forward amplitude increases as worms resume forward movement in the opposite direction.

UNSTABLE PERIODIC ORBITS AND DETERMINISTIC BEHAVIORAL VARIABILITY

The state space of worm locomotion is organized such that neighboring points correspond to similar behavioral sequences of length K . However, these neighboring sequences diverge with time, resulting in longer-time unpredictability, shown as an example in Fig. 4A. To understand this variability we note the strong cyclic appearance of trajectories within the projections, Fig. 3E, suggesting that cycles play an important role. We search for periodic orbits by identifying the first recurrence times in a neighborhood [60–62]. Briefly, given a point $\vec{x}(i)$ in state space, we find the smallest $k > i$ such that $\vec{x}(k)$ is in the neighborhood of $\vec{x}(i)$. The sequence

$[\vec{x}(i), \vec{x}(i+1), \dots, \vec{x}(k)]$ is then detected as a periodic orbit of period $p = k - i$ (Methods: Periodic Orbits). Across all foraging worms, the distribution of the number of periodic orbits exhibits peaks at approximately integer multiples of a minimum period p_{min} corresponding to the frequency of each worm’s body wave during forward locomotion, Fig. 4B (inset). We quantify the stability of each periodic trajectory by computing its maximal Floquet exponent (Methods: Floquet Exponents). The distribution of Floquet exponents is largely positive, indicating that the worm’s periodic orbits are mostly unstable, Fig. 4B. The unstable periodic orbits (UPOs) of worm behavior provide a longer timescale description of the movement and also a quantitative characterization of the trajectory divergence in Fig. 4A. We estimate the maximal Lyapunov exponent λ_{max} by a weighted average of the Floquet exponents of periodic orbits of increasing length, weighted by $e^{-\mu_1 p}$, where μ_1 is the maximal Floquet exponent of the orbit, and p is its period [63]. Including orbits of duration up to $p = 8$, Fig. 4C (blue), provides an approximation of λ_{max} , which agrees with direct trajectory divergence estimates averaged across all worms (gray bar, see also Fig. S6 and Methods: Maximal Lyapunov Exponent). The average across random segments of the same length converges more slowly, Fig. 4C (red).

The detected periodic orbits are interpretable in terms of commonly observed *C. elegans* behaviors. Orbits with the minimum period p_{min} correspond to forward and backward crawling including orbits with a dorsal or ventral bias, Fig. S7(B-C). More surprisingly, longer periodic orbits are composites, corresponding to longer time reorientation behaviors of the worm’s navigation and escape strategies [57, 58, 64]. In Fig. 4D (blue) we show state space trajectories of one such period-4 orbit. This orbit is composed of a reversal followed by a deep body bend, and subsequent forward movement; a posture sequence previously reported in pirouette reorientation and escape behaviors [57, 58]. Though this periodic orbit is several body waves long, it is repeated almost exactly at different times during the recording. We show one such recurrence Fig. 4D (orange), along with the corresponding posture sequences. The presence of such UPOs suggests an intriguing view of the worm’s foraging dynamics as following a complex landscape composed of unstable orbits, a picture that is rigorously correct for chaotic systems [65, 66]. Periodic orbits have also been investigated in a number of biological systems, including neuronal activities [67], human electroencephalograms [67], crayfish photoreceptors [68, 69], as well as cardiac arrhythmias and seizures [70–72].

SYMMETRIC LYAPUNOV SPECTRUM AND DAMPED-DRIVEN HAMILTONIAN DYNAMICS

While the behavior of *C. elegans* is simpler than most animals, the quantitative dynamics of worm posture defy

a straightforward interpretation or even, as yet, a model (see e.g. [73, 74] for reviews). There is rough stereotypy in the orbits corresponding to each behavior, but also large cycle-to-cycle variation. Such variability is linked to a positive maximal Lyapunov exponent and unstable periodic orbits, Fig. 4(B,C), Fig. S6, so that even within a “single” behavior such as forward crawling, each cycle is deterministically different. To more fully illuminate this variability, we examine the dynamics along all dimensions within the state space.

In an m -dimensional state space, local neighborhoods are sheared by the flow and are simultaneously stretched and squeezed along different directions, dynamics which are invariantly characterized the Lyapunov exponents, $\lambda_{i=1\dots m}$. Formally, such stretching and squeezing is described by the Jacobian $J_{\vec{x}(t)}$, which maps an m -dimensional spherical neighborhood to an m -dimensional ellipsoid. The spectrum of Lyapunov exponents is given by the infinite time average of the logarithms of the principle axes of the ellipsoid, as illustrated in Fig. 5A. Positive Lyapunov exponents reflect directions along which trajectory bundles expand, while negative exponents reflect shrinking directions.

The Lyapunov exponents reveal important information about the dynamics of a system (see e.g. [75]). The sum of the exponents is the average dissipation rate: zero for conservative systems and negative for those with dissipation. The sum of the positive exponents bounds the metric or Kolmogorov-Sinai (KS) entropy rate [76, 77], providing a principled measure of the unpredictability. In addition, the spectrum of Lyapunov exponents can reveal underlying symmetries and conservation laws. For example, continuous dynamical systems exhibit at least one zero exponent corresponding to time-translation invariance along the direction of the flow.

We compute the Lyapunov spectrum for the state space of *C. elegans* (Methods: Lyapunov Spectrum and Jacobian Estimation), and show bootstrapped density estimates of the $m = 7$ exponents across different worms, Fig. 5B. We find two positive exponents, $\lambda_1 = 0.66 (0.62, 0.69) s^{-1}$, $\lambda_2 = 0.29 (0.26, 0.32) s^{-1}$, and a third, near-zero exponent $\lambda_3 = 0.056 (-0.02, 0.11) s^{-1}$. The KS entropy rate is thus bounded by the sum of positive exponents as $h_{KS} \leq 1 (0.93, 1.09) \text{ nats/s}$ (note that we have restored the units of nats for ease of comparison with other entropy measures). The sum of all of the Lyapunov exponents is negative, indicating that the system is dissipative with a dissipation rate of, $\sum_i \lambda_i = -0.94 (-1.15, -0.78) s^{-1}$. Although trajectory bundles expand locally, dissipation causes them to contract as a whole and relax to an attracting manifold. We estimate the dimension of the attractor as the Kaplan-Yorke dimension $D_{KY} = 5.93 (5.75, 6.08)$ [78]. The combination of local expansion generating variability and local contraction generating stereotypy is an essential complexity of the worm’s posture dynamics.

The Lyapunov spectrum also exhibits a striking sym-

metry; exponents come in conjugate pairs that sum to the same number $\alpha = -0.27 (-0.3, -0.24) s^{-1}$, Fig. 5B (inset). The entire spectrum is thus symmetric about $\frac{\alpha}{2}$ (dotted line). The symmetry is also present in 6- and 8-dimensional embeddings, Fig. S8. Symmetric Lyapunov spectra have been previously observed in at least two kinds of damped-driven Hamiltonian systems: coupled oscillators with viscous damping where α is the dissipation per degree of freedom [79], and thermostatted molecular dynamic simulations where α is a feedback friction force that acts to maintain a dynamic equilibrium by either keeping the the kinetic energy of the particles constant [80–83]. Interestingly, in a biomechanical model of larval *Drosophila* locomotion, damped-driven Hamiltonian chaotic dynamics were sufficient to generate realistic forward and backward crawling, as well as more complex reorientation behaviors [84].

DISCUSSION

We use sequences of multidimensional data to reconstruct a maximally predictive state space (Fig. 1, Fig. 2). Conceptually, our approach is a timescale separation; short-time sequences define the reconstructed state variables while longer-time dynamics are encoded as state space trajectories. Our reconstruction *explicitly* seeks the full state information available in short-time dynamics, analogous to discovering the additional variable of velocity from the displacement time series of a simple oscillator. Such information is often added implicitly, for example through the choice of derivative filters in neural imaging [85, 86]. Both the resulting state variables and the geometry and topology of their trajectories offer important, coordinate-invariant understanding of the processes generating the dynamics.

We applied our approach to the posture time series of the locomotor behavior of the roundworm *C. elegans* and found that the state space is spanned by a 7D basis of interpretable modes Γ (Fig. 3), and their coefficients X , which are qualitatively similar for all worms in our foraging dataset. The basis is divided into three groups closely corresponding to forward, backward and turning locomotion. Biologically, these behaviors are linked to three classes of motor neurons: A and B ventral cord neurons which drive and backward and forward locomotion respectively, and sublateral motor neurons such as SMB and SMD which control deep body bends [87]. Furthermore, excitatory classes of ventral cord motor neurons were recently reported to be capable of spontaneous rhythm generation and proposed to be central pattern generators for forward and backward locomotion [88, 89]. We expect worms defective in different motor neurons to display a smaller projection on the behavioral modes in an interpretable manner. Although we have described the results for a 7D dimensional embedding, similar re-

sults are also found in 6 and 8 dimensions (Fig. S2-S3,S4). Indeed, it is likely that higher modes capturing head movement and other subtle motions exist but carry little predictive structure in our analyzed conditions. Mutant worms defective in motor control, different sensorimotor contexts, or even faster sampling rates could reveal the presence of subtle, additional dynamics.

In our embedding, the state space trajectories retained significant variability, occupying much of the volume in the reconstructed space. A measure of this volume is the Kaplan-York dimension and we find $D_{KY} \sim 6$, not substantially smaller than the embedding dimension. One hypothesis for this variability is that worm behavior is stochastic and results from noise induced transitions between a finite number of elements such as stable limit cycles representing individual stereotyped motifs [21, 47, 90, 91]. However, the exponential divergence of nearby state-space trajectories (Fig. S6) and the consistency of this divergence with the spectrum of unstable periodic orbits (Fig. 4), as well as the symmetric Lyapunov spectrum (Fig. 5) provide evidence for important, deterministic variation. From the perspective of deterministic chaos, behavioral dynamics are an aperiodic wandering among an infinite number of unstable periodic orbits, allowing an animal to generate an infinite number of behavioral sequences. Indeed, this agrees with the finding that the number of novel sequences in *C. elegans* behavior grows with the observation time [53]. On the other hand, stereotyped trajectories can emerge naturally as orbits with low values of the maximal Floquet exponent. Such trajectories can also be generated by stabilizing periodic orbits with control, e.g. a simple linear controller of the form $K(\mathbf{g}(t) - \mathbf{x}(t))$, where $\mathbf{g}(t)$ are the desired goal dynamics, $\mathbf{x}(t)$ is the current state and K is a control gain matrix [92, 93].

The symmetric form of the Lyapunov spectrum suggests that the worm’s behavioral dynamics can be interpreted as normal modes of a system of coupled, damped and driven, Hamiltonian oscillators,

$$\begin{aligned}\dot{Q}_i &= \frac{\partial H}{\partial P_i} \\ \dot{P}_i &= -\frac{\partial H}{\partial Q_i} + C(Q_i, P_i, \psi(t)),\end{aligned}\tag{3}$$

where (Q_i, P_i) are the generalized position and momentum coordinates for the i^{th} normal mode. The Hamiltonian is a scalar function governing the time-independent dynamics resulting from the mechanics of the worm’s body, while $C(Q_i, P_i, \psi(t))$ encapsulates the time-dependent neuromuscular control forces due to interaction of worm’s body with the environment, proprioceptive feedback and neural processing of various sensory stimuli $\psi(t)$. Dynamics based on Hamiltonian structure are often associated with optimality and conservation laws and multiple efforts have reported quantities that remain roughly constant across a range external loads

during *C. elegans* locomotion, such as the normalized wave length of the body wave, angle of attack, bending power, and the phase relationship between the muscle activity and body curvature [94–98]. Following the example of thermostatted dynamics (designed to capture constant temperature dynamics, see e.g. [83, 99]), such emergent constants could be explained through feedback control arising from proprioceptive feedback, which is thought to underlie gait modulation in *C. elegans* [100, 101]. Our work also allows for connections between non-equilibrium thermodynamics and worm behavior. For example, worm dynamics breaks the Hamiltonian time-reversible symmetry in a continuous fashion via the dissipation rate α , which sets the characteristic time-scale at which dynamics can be considered time-reversible symmetric. In addition, the sum of Lyapunov exponents reported here is an estimate of the entropy production rate [102].

The dynamical invariants such as Lyapunov exponents, dimensions and entropies made accessible by our embedding approach provide important constraints and new understanding for short-time behavior consisting of neuromuscular control along with the biomechanics of the body and its environmental interaction. However, longer timescales are also present in the short periodic orbits, which are interpretable as forward/backward locomotion, and also longer time reorientation sequences such as pirouettes. Longer timescales can also be addressed through a systematic coarse-graining of the continuous state space dynamics which results in a transfer operator, see e.g. [103]. In this approach the partition itself subsumes much of the nonlinearity so that the eigenvalues of the transfer operator can provide a systematic and useful timescale separation. In contrast, linear measures like the power spectrum are often not informative on the original dynamics of complex systems [104].

While we expect a dynamical systems perspective to be generally useful in understanding natural behavior, the analysis here benefits from the relative simplicity of the worm’s foraging dynamics and the resulting interpretability of the modes. Though other settings and organisms may generate more complex embeddings, important dynamical information such as trajectory stability and dynamical invariants can still be extracted from the state space reconstruction. Embedding ideas have also been recently used to understand the global brain dynamics of *C. elegans* [105] and to identify metastable sets and slow order parameters from molecular dynamics simulations using Markov operators [106, 107].

Across wide areas of science there has been a remarkable increase in the availability of precise, multidimensional and dynamical data and new analysis ideas are emerging (see e.g. [52, 108–110]). Here, we improve on the prior work on state space reconstruction [38, 111–116], where much was in the context of either univariate measurements or known dynamical systems and included a heuristic search of reconstruction parameters. However, challenges associated with high-dimensionality,

data sampling and nonstationarity remain. For example, the one-step error for N samples from a D dimensional dynamical system is $E(1)/e_s \approx N^{-1/D}$ [117, 118]—larger dimensional systems require exponentially more data to keep $E(1)/e_s \ll 1$. A related difficulty is the Euclidean metric used to find nearest neighbor distances, which is invalid even in lower-dimensional spaces with large curvature fluctuations. In this setting, it might be possible to use metric learning algorithms [119] to recover a suitable metric from data. Finally, recent progress in leveraging artificial neural networks to recover dynamical invariants [120] and to seek state-space representations with parsimonious dynamics [121] offers promising directions for combining a principled dynamical perspective with high-dimensional, real-world systems.

ACKNOWLEDGMENTS

We thank David Jordan, Ian Ehteridge and Antonio Celani for comments. This work was supported by OIST Graduate University (TA, GJS), a program grant from the Netherlands Organization for Scientific Research (AC, GJS), and by Vrije Universiteit Amsterdam (GJS).

METHODS

Software: Code for all analysis reported here was written in MATLAB [122] and is publicly available: <https://bitbucket.org/tosifahamed/behavioral-state-space>.

Experimental Details: A brief description of the foraging and escape response datasets is given below. For more details please see the original manuscripts [47, 48].

Foraging Dataset: $N = 12$ L4-stage N2 worms were recorded at 32 Hz with high resolution tracking microscopy. For the analysis here the data was downsampled to 16 Hz. Worms were cultivated under standard conditions at 20° C [123]. Before the assay, worms were cleaned of *E.coli* bacteria by a 1-minute immersion in NGM buffer. Worms were then placed on a 9.1cm assay plate (Petri-Dish) with a 5cm radius copper ring pressed into the agar surface for confinement. The assay started 5 minutes after the transfer and lasted 35 minutes.

Escape Response Dataset: $N = 92$ mid to late L4 stage N2 worms were targeted on the head with a 100 ms, 75 mA IR laser pulse from a diode laser ($\lambda = 1440\text{nm}$), resulting in a localized temperature change of approximately 0.5° C. Images were recorded at 20 Hz for 30 s (10 s before stimulation and 20 s after stimulation). To prevent adaptation each worm was only assayed once. To match the sampling rate of the foraging dataset, the posture time series was interpolated and downsampled to 16 Hz using the MATLAB [122] `resample` command.

State Space Reconstruction for the Lorenz system: We simulated the Lorenz system [124],

$$\begin{aligned} \dot{s}_1 &= 10(s_2 - s_1) \\ \dot{s}_2 &= s_1(28 - s_3) - s_2 \\ \dot{s}_3 &= s_1s_2 - \frac{8}{3}s_3 \end{aligned}$$

using MATLAB’s ode45 Runge-Kutta ODE solver [122] with a time-step $dt = 0.01$ s and error tolerances of 10^{-8} . We take the variable s_1 as the observation time series $y(t)$. To simulate a noisy observation process we add to $y(t)$ a uniform white noise with standard deviation of 0.5% the standard deviation of s_1 .

Image Analysis and Posture Space Estimation:

The tracking and posture space estimation follows [48]. Briefly, we parameterize the shape of a worm by tangent angles calculated at 100 points along the body image skeleton. For a recording session of T frames, this results in a $T \times 100$ matrix Θ , containing the shape information for each uncrossed frame where the worm’s body does not intersect itself. Next, a 5-dimensional approximation of the 100 dimensional posture space is calculated by projecting the elements of Θ onto the basis given by the first 5 singular vectors (eigenworms) of Θ . For frames with a body crossing an inverse tracking algorithm is used to identify the eigenworm projections [48].

Worm State Space Reconstruction:

Given a d -dimensional time-series in $Y = [\mathbf{y}_1^{1:T}, \dots, \mathbf{y}_d^{1:T}]$, along with an estimate of the optimal embedding window K^* , and minimum embedding dimension m^* , the state space reconstruction proceeds as follows. First, we create the $L \times K^*d$ matrix \bar{Y}_{K^*} containing delayed copies of the mean subtracted measurements, $\bar{Y}_{K^*} = [\mathbf{y}_{1:d}^{1:L}, \mathbf{y}_{1:d}^{2:(L+1)}, \dots, \mathbf{y}_{1:d}^{K^*:T}]$, where $L = (T - K^* + 1)$. For the postures of *C. elegans*, the measurements are composed of $d = 5$ eigenworm coefficients. Next, we perform ICA on the space formed by the first m^* singular vectors of \bar{Y}_{K^*} using the FastICA algorithm [39] to obtain an m^* -dimensional state space spanned by the independent basis vectors, which we call behavioral modes and denote $\mathbf{\Gamma}$. Projections of \bar{Y}_{K^*} on the state space are contained in the $L \times m^*$ state space matrix X . Each row, $\vec{x}(t)$, of X is the behavioral state encoding the instantaneous behavior of the worm at time t , while the temporal sequence, $[\vec{x}(t), \dots, \vec{x}(t+\tau)]$, forms a continuous trajectory in state space which encodes the shape change dynamics of a behavioral sequence.

Choosing Reconstruction Parameters by Maximizing Predictability:

To choose the reconstruction parameters (K, m) we first vary K in the range $1 \leq K \leq K_{max}$ and estimate T_{pred} in the candidate state space \bar{Y}_K formed by the delayed observations. We set K^* as the minimum K where T_{pred} as a function of K begins to

decrease. In cases where T_{pred} saturates but doesn’t decrease, we choose K^* as the K at which T_{pred} saturates. As a guide, we choose K_{max} such that for any delay larger than K_{max} the autocorrelation function of the observations is close to zero. If K_{max} appears too short, then it can be increased step-wise until $T_{pred}(K)$ starts decreasing. For the Lorenz system we have $K_{max} = 100$ frames, while for the worm data we have $K_{max} = 30$ frames. Intuitively, K^* should allow the reconstruction to capture the fastest time scale of the system, which for chaotic systems is set by the period of the smallest UPO, p_{min} . Increasing K^* further filters across longer periods and in the limit $K \rightarrow \infty$, the SVD filter becomes a discrete Fourier transform [125]. On the other end, K^* should be large enough to embed the dynamics completely. Using the bound given by Takens embedding theorem [31, 38], we get $(2m^* + 1)/d \leq K^* < p_{min}$. Empirically, we find that $p_{min}/4 \leq K^* \leq p_{min}/2$.

Once the embedding window is set as K^* , we next perform the singular value decomposition $\bar{Y}_{K^*} = U\Sigma V^T$. The first m columns of U contain the normalized projections of \bar{Y}_{K^*} onto its first m singular vectors. To find the embedding dimension, we vary m and compute T_{pred} as above. We set the embedding dimension m^* as the minimum m where T_{pred} as a function of m saturates or begins to decrease.

Nearest Neighbor Prediction: We estimate the τ -step future of an observation $\vec{y}(t')$, denoted $\vec{y}_{est}(t' + \tau)$, from an average of the τ -step future of N_b nearest neighbors of the corresponding state space point $\vec{x}(t')$. Specifically, we find N_b nearest neighbors of $\vec{x}(t')$ in state space, denoted by $\vec{z}(t'; r)$ for the r^{th} nearest neighbor of $\vec{x}(t')$, and average their values after τ steps, $\vec{x}_{est}(t' + \tau) = \langle \vec{z}(t' + \tau; r) \rangle_r$, for all N_b neighbors. Finally, we project $\vec{x}_{est}(t' + \tau)$ back to the observation space to get $\vec{y}_{est}(t' + \tau)$. We take only the transverse nearest neighbors (i.e. neighbors that are not in succession). The transverse nearest neighbors of $\vec{x}(t')$ are identified by the local minima of $R_{t'}(t)$ estimated using the `findpeaks` function in MATLAB [122], where $R_{t'}(t)$ is the distance between $\vec{x}(t')$ and all other points in state space. We quantify the τ -step prediction accuracy by the root mean squared error

$$E(\tau) = \sum_{j=1}^d \langle (y_j(t' + \tau) - (y_{est})_j(t' + \tau))^2 \rangle_t^{1/2}$$

for $N = 10^4$ different test points $\vec{y}(t')$ in the measurement time series. The predictions are made to a maximum prediction time which is long enough so that $E(\tau)$ saturates to e_s . The root mean squared error is also a function of the N_b , the total number of nearest neighbors used for prediction. Making this dependence explicit, we write $E(\tau, N_b)$ when N_b is considered a variable. We set the number of neighbors by minimizing the bias and variance of the one-step prediction error.

Prediction Timescale T_{pred} : In cases where the error growth $E(\tau)$ is well approximated by a sigmoid

(conjectured by Lorenz for chaotic systems with a single Lyapunov exponent λ [41]), one can show $T_{pred} = \frac{1}{\lambda} \log\left(\frac{e_s}{e_1}\right)$, where e_1 is the one time step error $E(1)$. Based on information theoretic considerations Farmer derived the upper bound for the predictability time scale as $\frac{D_I}{h_{KS}} \log\left(\frac{e_s}{e_1}\right)$, where D_I is the information dimension and h_{KS} is the Kolmogorov-Sinai entropy rate, which is consistent with our calculation for the sigmoid assumption. Importantly, these estimates shed light on the asymptotic behavior of T_{pred} . For small values of K and m , the error is affected by some fraction of false nearest neighbors due to underembedding [44] leading to an overestimate of the local expansion rate and consequently the positive Lyapunov exponents. This causes a drop in T_{pred} via the $1/\lambda$ term. On the other hand, as we increase K , the average Euclidean distance between nearest neighbors, e_1 steadily increases leading to a decrease in T_{pred} for high dimensions. In the middle of these two extremes we find a range of suitable values for the embedding window K . The SVD coordinates are weighted by decreasing singular values, which correspond to the variance of the data projected along the different singular vectors. In the noiseless case, the singular values decay towards zero, while in the presence of noise they decay before saturating to the standard deviation of noise (termed noise floor in Ref. [31]), thus higher dimensions are generically dominated by noise in this case. Consequently, e_1 doesn't increase as a function of m in the noiseless case leading T_{pred} to saturate after successful embedding. However, in the presence of noise e_1 increases causing T_{pred} to go down.

Calculation of T_{pred} : We developed a fixed-point algorithm to estimate T_{pred} . We begin with an initial guess of e_s labeled e_s^0 and time τ_s^0 such that $E(\tau) > e_s^0$ for all $\tau \geq \tau_s^0$. Next, noting that for large times $\int E(\tau) d\tau = e_s \tau - \Delta$ we fit a line to a numerical estimate of $\int E(\tau) d\tau$ from τ_s^0 to τ_{max} . The slope of this line provides the next estimate of e_s labeled e_s^1 , and the intercept is the next estimate of the area Δ , labeled Δ^1 . We use e_s^1 to again estimate τ_s^1 and fit a line to $\int E(\tau) d\tau$ from τ_s^1 to τ_{max} , repeating the process until the estimates for e_s^j and Δ^j converge. Using the final estimates of Δ and e_s we can get a robust estimate $T_{pred} = \frac{\Delta}{e_s}$. A schematic of this iterative process is shown in Fig. S9. In our experience it only takes 3-4 iterations for the estimates to converge. To obtain the error bars we bootstrap across the prediction test points, generating 100 bootstrapped $E(\tau)$ curves along with T_{pred} estimates for each. These are then used to estimate the 95% confidence intervals of T_{pred} .

Periodic Orbits: To detect periodic orbits of length p we identify close recurrences in the state space after approximately p time steps [60–62]. Specifically, we start at point $\vec{x}(i)$ in state space, and find smallest $k > i$ such that

$\|\vec{x}(k) - \vec{x}(i)\| < \epsilon$. The sequence $[\vec{x}(i), \vec{x}(i+1), \dots, \vec{x}(k)]$ is then stored as a periodic orbit of period $p = k - i$. Again, we only consider transverse recurrences to avoid sequential points. If such a k cannot be found then it implies that no periodic orbits exist at the scale of ϵ . The distance scale of the recurrence ϵ is calculated through the function $\epsilon(r, t)$ as defined in Ref. [61], which gives the r^{th} smallest distance between state space points separated by time t . An example of this function is shown in Fig. S7A. Small values of $\epsilon(r, t)$, identified by local minima of $\epsilon(r, t)$, indicate close recurrences after times t_* , and in consequence reveal the existence of periodic orbits of length t_* . If $\epsilon(r, t)$ does not show any local minima, then periodic orbits cannot be detected from the data. In this manner $\epsilon(r, t_*)$ gives the minimum distance at which we must look to find a periodic orbit of length t_* . The smallest recurrence time corresponding to the first local minima of $\epsilon(r, t)$ equals the smallest period p_{min} detected in the data. In the Jacobian and maximal exponent calculations described below, ϵ_* is the distance corresponding to p_{min} . Finally, we set $r = m^*$, the dimension of the reconstructed state space.

Maximal Lyapunov Exponent λ_{max} : Our test for the exponential divergence of neighboring trajectories follows standard approaches [126]. Specifically, we consider a reference trajectory $\vec{x}(t' + \tau)$ and its nearest neighbors within a distance ϵ_* (see Methods: Periodic Orbits). We then track the average distance between the reference and neighboring trajectories over time to obtain the curve $\delta_{t'}(\tau)$. A significant linear region in the $\langle \log \delta_{t'}(\tau) \rangle_{t'}$ curve indicates an exponential divergence of neighboring trajectories, while the slope of the linear region provides an estimate of the maximal Lyapunov exponent λ_{max} . There is typically a transient before the exponential growth where the perturbation vector aligns itself with the Lyapunov vector corresponding to the maximal exponent. In the Lorenz system it is seen that this transient arises from the finite-size of the perturbation and vanishes in the infinitesimal limit. To avoid effects due to non-stationarities we perform this calculation on the final two minutes of the recording.

Jacobian Estimation: The Jacobian at the point $\vec{x}(i)$ in state space, denoted $\mathbf{J}_{\vec{x}(i)}$, is the derivative of the dynamics at $\vec{x}(i)$, forming the local linear approximation of the dynamics at that point. We use a modified version of the Jacobian estimation algorithm described in Ref. [127], which solves a weighted regression problem $\mathbf{B}_{\vec{x}(i)}^1 = \mathbf{B}_{\vec{x}(i)} \cdot \mathbf{J}_{\vec{x}(i)}$, where points are assigned weights according to their distance from $\vec{x}(i)$ as per the weighting function defined below. $\mathbf{B}_{\vec{x}(i)}$ is a $(T - K^*) \times (m + 1)$ matrix containing all weighted state space points concatenated with a column of ones, while $\mathbf{B}_{\vec{x}(i)}^1$ is a $(T - K^*) \times m$ matrix containing all weighted successors. Each row of $\mathbf{B}_{\vec{x}(i)}$ and $\mathbf{B}_{\vec{x}(i)}^1$ is weighted by $w(k) = \exp\left\{-\frac{\|\vec{x}(i) - \vec{x}(k)\|}{\epsilon_*}\right\}$. The estimated local Jacobian matrix is then given by $\mathbf{J}_{\vec{x}(i)} = \mathbf{B}_{\vec{x}(i)}^\dagger \cdot \mathbf{B}_{\vec{x}(i)}^1$ where $\mathbf{B}_{\vec{x}(i)}^\dagger$ is the pseudoinverse

of $\mathbf{B}_{\bar{x}(t)}$ which we compute using the `pinv` function in MATLAB [122]. Note that ϵ_* is the distance scale corresponding to the minimum period recurrence (see Methods: Periodic Orbits).

Floquet Exponents: The real parts of the Floquet exponents of a periodic orbit, which measure their stability, are equal to the Lyapunov exponents of the orbit [128]. To estimate the Floquet exponents of a periodic orbit, we estimate the maximal local Lyapunov exponent along the orbit using established algorithms [129, 130]. We use a recursive QR iteration for obtaining the eigenvalues of the product of Jacobian matrices along a periodic orbit. The average of the logarithms of the eigenvalues give the m local Lyapunov exponents of the orbit in an m dimensional state space. The maximal exponent is then the

Floquet exponent of the periodic orbit.

Lyapunov Exponents of Random Sequences: To calculate the exponents for short random sequences in Fig. 4C, we proceed as above but instead of the sequence being a periodic orbit, it is formed by starting at a random random point in state space and following it for the duration of a periodic orbit. As a result these sequences are not necessarily recurrent.

Lyapunov Spectrum: To estimate the full spectrum of Lyapunov exponents for a given state space, we use the recursive QR procedure as above [129, 130], but now calculate the product of Jacobians over the entire recording instead of a short UPO segment.

-
- [1] J. Gray, *How Animals Move*, Pelican books (University Press, 1953).
 - [2] N. Scafetta, D. Marchi, and B. J. West, *Understanding the complexity of human gait dynamics*, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **19**, 026108 (2009).
 - [3] P. Grobstein, *Variability in behavior and the nervous system*, *Encyclopedia of Human Behavior* **4**, 447 (1994).
 - [4] K. Newell and D. Corcos, *Variability and Motor Control* (Human Kinetics Publishers, 1993).
 - [5] B. J. Cole, *Fractal time in animal behaviour: the movement activity of drosophila*, *Animal Behaviour* **50**, 1317 (1995).
 - [6] A. Maye, C.-h. Hsieh, G. Sugihara, and B. Brembs, *Order in spontaneous behavior*, *PloS one* **2**, e443 (2007).
 - [7] G. M. Card, *Escape behaviors in insects*. Current opinion in neurobiology **22**, 180 (2012).
 - [8] T. Y. Moore, K. L. Cooper, A. A. Biewener, and R. Vasudevan, *Unpredictability of escape trajectory explains predator evasion ability and microhabitat preference of desert rodents*, *Nature communications* **8** (2017).
 - [9] A. Gomez-Marin, J. J. Paton, A. R. Kampff, R. M. Costa, and Z. F. Mainen, *Big behavioral data: psychology, ethology and the foundations of neuroscience*. *Nature neuroscience* **17**, 1455 (2014).
 - [10] J. W. Krakauer, A. A. Ghazanfar, A. Gomez-Marin, M. A. MacIver, and D. Poeppel, *Neuroscience needs behavior: Correcting a reductionist bias*. *Neuron* **93**, 480 (2017).
 - [11] A. E. X. Brown and B. de Bivort, *Ethology as a physical science*, *Nature Physics* **14**, 653 (2018).
 - [12] G. J. Berman, *Measuring behavior across scales*. *BMC biology* **16**, 23 (2018).
 - [13] N. J. Cowan, M. M. Ankarali, J. P. Dyhr, M. S. Madhav, E. Roth, S. Sefati, S. Sponberg, S. A. Stamper, E. S. Fortune, and T. L. Daniel, *Feedback control as a framework for understanding tradeoffs in biology*, *American Zoologist* **54**, 223 (2014).
 - [14] J. Aguilar, T. Zhang, F. Qian, M. Kingsbury, B. McInroe, N. Mazouchova, C. Li, R. Maladen, C. Gong, M. Travers, R. L. Hatton, H. Choset, P. B. Umbanhowar, and D. I. Goldman, *A review on locomotion robotics: the study of movement at the intersection of robotics, soft matter and dynamical systems*. Reports on progress in physics. Physical Society (Great Britain) **79**, 110001 (2016).
 - [15] S. Ramaswamy, *Active matter*, *Journal of Statistical Mechanics: Theory and Experiment* **2017**, 054002 (2017).
 - [16] Mathis, Alexander, Mamidanna, Pranav, Cury, Kevin M, Abe, Taiga, Murthy, Venkatesh N, Mathis, Mackenzie Weygandt, and Bethge, Matthias, *DeepLabCut: markerless pose estimation of user-defined body parts with deep learning*, *Nature Neuroscience* **21**, 1281 (2018).
 - [17] Pereira, Talmo D, Aldarondo, Diego E, Willmore, Lindsay, Kislin, Mikhail, Wang, Samuel S H, Murthy, Mala, and Shaevitz, Joshua W, *Fast animal pose estimation using deep neural networks*, *Nature Methods* **16**, 117 (2019).
 - [18] S. Han, E. Taralova, C. Dupre, and R. Yuste, *Comprehensive machine learning analysis of Hydra behavior reveals a stable basal behavioral repertoire*. *eLife* **7**, 413 (2018).
 - [19] A. B. Wiltschko, M. J. Johnson, G. Iurilli, R. E. Peterson, J. M. Katon, S. L. Pashkovski, V. E. Abraira, R. P. Adams, and S. R. Datta, *Mapping sub-second structure in mouse behavior*. *Neuron* **88**, 1121 (2015).
 - [20] J. M. Graving, D. Chae, H. Naik, L. Li, B. Koger, B. R. Costelloe, and I. D. Couzin, *DeepPoseKit: a software toolkit for fast and robust animal pose estimation using deep learning*, *bioRxiv* (2019), 10.1101/620245.
 - [21] G. J. Berman, D. M. Choi, W. Bialek, and J. W. Shaevitz, *Mapping the stereotyped behaviour of freely moving fruit flies*, *Journal of the Royal Society Interface* **11**, 20140672 (2014).
 - [22] G. J. Stephens, B. Johnson-Kerner, W. Bialek, and W. S. Ryu, *Dimensionality and Dynamics in the Behavior of C. elegans*, *PLoS Computational Biology* **4**, e1000028+ (2008).
 - [23] Flavell, Steven W, Pokala, Navin, Macosko, Evan Z, Albrecht, Dirk R, Larsch, Johannes, and Bargmann, Cor-

- nelia I., *Serotonin and the Neuropeptide PDF Initiate and Extend Opposing Behavioral States in C. elegans*, *Cell* **154**, 1023 (2013).
- [24] R. E. Johnson, S. Linderman, T. Panier, C. L. Wee, E. Song, K. J. Herrera, A. Miller, and F. Engert, *Probabilistic Models of Larval Zebrafish Behavior: Structure on Many Scales*, bioRxiv (2019).
- [25] S. H. Strogatz, *Nonlinear Dynamics and Chaos with Student Solutions Manual: With Applications to Physics, Biology, Chemistry, and Engineering* (CRC Press, 2018).
- [26] D. D. Nolte, *The tangled tale of phase space*, *Physics today* **63**, 33 (2010).
- [27] J. Durbin and S. J. Koopman, *Time series analysis by state space methods* (Oxford university press, 2012).
- [28] N. Tinbergen, *On aims and methods of ethology*, *Zeitschrift für Tierpsychologie* **20**, 410 (1963).
- [29] J. Palis and W. de Melo, *Geometric Theory of Dynamical Systems* (Springer US, 1982).
- [30] M. Casdagli, S. Eubank, J. D. Farmer, and J. Gibson, *State space reconstruction in the presence of noise*, *Physica D: Nonlinear Phenomena* **51**, 52 (1991).
- [31] D. S. Broomhead and G. P. King, *Extracting qualitative dynamics from experimental data*, *Physica D: Nonlinear Phenomena* **20**, 217 (1986).
- [32] G. Sugihara and R. M. May, *Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series*, *Nature* **344**, 734 (1990).
- [33] S. Tajima, T. Yanagawa, N. Fujii, and T. Toyozumi, *Untangling brain-wide dynamics in consciousness by cross-embedding*, *PLOS Computational Biology* **11**, e1004537 (2015).
- [34] H. Kantz and T. Schreiber, *Nonlinear time series analysis*, Vol. 7 (Cambridge university press, 2004).
- [35] P. Read, *Phase portrait reconstruction using multivariate singular systems analysis*, *Physica D: Nonlinear Phenomena* **69**, 353 (1993).
- [36] H. Ye and G. Sugihara, *Information leverage in interconnected ecosystems: Overcoming the curse of dimensionality*, *Science* **353**, 922 (2016).
- [37] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, *Geometry from a time series*, *Physical review letters* **45**, 712 (1980).
- [38] F. Takens, *Detecting strange attractors in turbulence*, *Dynamical Systems and Turbulence, Warwick 1980*, *Dynamical Systems and Turbulence Lecture Notes in Mathematics*, **898**, 366 (1981).
- [39] A. Hyvrinen, *Fast and robust fixed-point algorithms for independent component analysis*. *IEEE transactions on neural networks* **10**, 626 (1999).
- [40] J. F. Gibson, J. Doyne Farmer, M. Casdagli, and S. Eubank, *An analytic approach to practical state space reconstruction*, *Physica. D, Nonlinear phenomena* **57**, 1 (1992).
- [41] E. Lorenz, in *Seminar on Predictability, 4-8 September 1995*, Vol. 1, ECMWF (ECMWF, Shinfield Park, Reading, 1995) pp. 1–18.
- [42] E. Aurell, G. Boffetta, A. Crisanti, G. Paladin, and A. Vulpiani, *Predictability in the large: an extension of the concept of lyapunov exponent*, *Journal of Physics A: Mathematical and General* **30**, 1 (1997).
- [43] J. B. Gao, J. Hu, W. W. Tung, and Y. H. Cao, *Distinguishing chaos from noise by scale-dependent lyapunov exponent*, *Physical Review E* **74** (2006).
- [44] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, *Determining embedding dimension for phase-space reconstruction using a geometrical construction*, *Physical Review A* **45**, 3403 (1992).
- [45] L. Cao, A. Mees, and K. Judd, *Dynamics from multivariate time series*, *Physica D: Nonlinear Phenomena* **121**, 75 (1998).
- [46] K. Judd and A. Mees, *Embedding as a modeling problem*, *Physica D: Nonlinear Phenomena* **120**, 273 (1998).
- [47] G. J. Stephens, M. Bueno de Mesquita, W. S. Ryu, and W. Bialek, *Emergence of long timescales and stereotyped behaviors in Caenorhabditis elegans*, *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7286 (2011).
- [48] O. D. Broekmans, J. B. Rodgers, W. S. Ryu, and G. J. Stephens, *Resolving coiled shapes reveals new reorientation behaviors in C. elegans*, *eLife* **5**, e17227 (2016).
- [49] N. A. Croll, *Behavioural analysis of nematode movement*. *Advances in parasitology* **13**, 71 (1975).
- [50] N. A. Croll, *Components and patterns in the behaviour of the nematode Caenorhabditis elegans*, *Journal of zoology* **176**, 159 (1975).
- [51] J. Likitlersuang, G. Stephens, K. Palanski, and W. S. Ryu, *C. elegans tracking and behavioral measurement*, *J Vis Exp*, e4094 (2012).
- [52] A. C. Costa, T. Ahamed, and G. J. Stephens, *Adaptive, locally linear models of complex dynamics*. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 1501 (2019).
- [53] R. F. Schwarz, R. Branicky, L. J. Grundy, W. R. Schafer, and A. E. X. Brown, *Changes in postural syntax characterize sensory modulation and natural variation of C. elegans locomotion*. *PLoS computational biology* **11**, e1004322 (2015).
- [54] 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** (2018).
- [55] A. Gomez-Marin, G. J. Stephens, and A. E. X. Brown, *Hierarchical compression of Caenorhabditis elegans locomotion reveals phenotypic differences in the organization of behaviour*. *Journal of the Royal Society, Interface* **13** (2016).
- [56] S. Gupta and A. Gomez-Marin, *A context-free grammar for Caenorhabditis elegans behavior*, bioRxiv (2019), 10.1101/708891.
- [57] J. M. Gray, J. J. Hill, and C. I. Bargmann, *A circuit for navigation in Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3184 (2005).
- [58] J. L. Donnelly, C. M. Clark, A. M. Leifer, J. K. Pirri, M. Haburcak, M. M. Francis, A. D. T. Samuel, and M. J. Alkema, *Monoaminergic orchestration of motor programs in a complex C. elegans behavior*, *PLoS Biology* **11**, e1001529 (2013).
- [59] A. Mohammadi, J. B. Rodgers, I. Kotera, and W. S. Ryu, *Behavioral response of Caenorhabditis elegans to localized thermal stimuli*, *BMC neuroscience* **14**, 66 (2013).
- [60] D. P. Lathrop and E. J. Kostelich, *Characterization of an experimental strange attractor by periodic orbits*, *Physical Review A* **40**, 4028 (1989).

- [61] K. Pawelzik and H. Schuster, *Unstable periodic orbits and prediction*, Physical Review A **43**, 1808 (1991).
- [62] R. Badii, E. Brun, M. Finardi, L. Flepp, R. Holzner, J. Parisi, C. Reyl, and J. Simonet, *Progress in the analysis of experimental chaos through periodic orbits*, Reviews of Modern Physics **66**, 1389 (1994).
- [63] P. Cvitanović, *Invariant measurement of strange sets in terms of cycles*, Physical Review Letters **61**, 2729 (1988).
- [64] J. T. Pierce-Shimomura, T. M. Morse, and S. R. Lockery, *The fundamental role of pirouettes in Caenorhabditis elegans chemotaxis*. The Journal of neuroscience : the official journal of the Society for Neuroscience **19**, 9557 (1999).
- [65] J. Banks, J. Brooks, G. Cairns, G. Davis, and P. Stacey, *On devaney's definition of chaos*, The American mathematical monthly **99**, 332 (1992).
- [66] R. Devaney, *An introduction to chaotic dynamical systems* (Westview press, 2008).
- [67] P. So, J. T. Francis, T. I. Netoff, B. J. Gluckman, and S. J. Schiff, *Periodic orbits: a new language for neuronal dynamics*, Biophysical journal **74**, 2776 (1998).
- [68] D. Pierson and F. Moss, *Detecting periodic unstable points in noisy chaotic and limit cycle attractors with applications to biology*, Physical review letters **75**, 2124 (1995).
- [69] X. Pei and F. Moss, *Characterization of low-dimensional dynamics in the crayfish caudal photoreceptor*, Nature **379**, 618 (1996).
- [70] A. Garfinkel, M. L. Spano, W. L. Ditto, and J. N. Weiss, *Controlling cardiac chaos*, Science **257**, 1230 (1992).
- [71] S. J. Schiff, K. Jerger, D. H. Duong, T. Chang, M. L. Spano, and W. L. Ditto, *Controlling chaos in the brain*, Nature **370**, 615 (1994).
- [72] N. Mishra, M. Hasse, B. Biswal, and H. P. Singh, *Reliability of unstable periodic orbit based control strategies in biological systems*, Chaos: An Interdisciplinary Journal of Nonlinear Science **25**, 043104 (2015).
- [73] J. Gjorgjieva, D. Biron, and G. Haspel, *Neurobiology of Caenorhabditis elegans Locomotion: Where Do We Stand?* Bioscience **64**, 476 (2014).
- [74] N. Cohen and T. Sanders, *Nematode locomotion: dissecting the neuronal-environmental loop*, Current opinion in neurobiology **25**, 99 (2014).
- [75] A. Pikovsky and A. Politi, *Lyapunov exponents: a tool to explore complex dynamics* (Cambridge University Press, 2016).
- [76] Y. B. Pesin, *Dimension theory in dynamical systems: contemporary views and applications* (University of Chicago Press, 2008).
- [77] L.-S. Young, *Dimension, entropy and lyapunov exponents*, Ergodic theory and dynamical systems **2**, 109 (1982).
- [78] P. Frederickson, J. L. Kaplan, E. D. Yorke, and J. A. Yorke, *The liapunov dimension of strange attractors*, Journal of differential equations **49**, 185 (1983).
- [79] U. Dressler, *Symmetry property of the lyapunov spectra of a class of dissipative dynamical systems with viscous damping*, Physical Review A **38**, 2103 (1988).
- [80] D. Gupalo, A. Kaganovich, and E. Cohen, *Symmetry of lyapunov spectrum*, Journal of statistical physics **74**, 1145 (1994).
- [81] C. P. Dettmann and G. Morriss, *Proof of lyapunov exponent pairing for systems at constant kinetic energy*, Physical Review E **53**, R5545 (1996).
- [82] D. Ruelle, *Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics*, Journal of Statistical Physics **95**, 393 (1999).
- [83] J. N. Bright, D. J. Evans, and D. J. Searles, *New observations regarding deterministic, time-reversible thermostats and gauss's principle of least constraint*, The Journal of chemical physics **122**, 194106 (2005).
- [84] J. Loveless, K. Lagogiannis, and B. Webb, *Modelling the mechanics of exploration in larval drosophila*. PLoS computational biology **15**, e1006635 (2019).
- [85] S. Kato, H. S. Kaplan, T. Schrödel, S. Skora, T. H. Lindsay, E. Yemini, S. Lockery, and M. Zimmer, *Global Brain Dynamics Embed the Motor Command Sequence of Caenorhabditis elegans*, Cell **163**, 656 (2015).
- [86] X. Chen, F. Randi, A. M. Leifer, and W. Bialek, *Searching for collective behavior in a small brain*, Physical Review E **99**, 052418 (2019).
- [87] S. J. Cook, T. A. Jarrell, C. A. Brittin, Y. Wang, A. E. Bloniarz, M. A. Yakovlev, K. C. Q. Nguyen, L. T.-H. Tang, E. A. Bayer, J. S. Duerr, H. E. Blow, O. Hobert, D. H. Hall, and S. W. Emmons, *Whole-animal connectomes of both Caenorhabditis elegans sexes*. Nature **571**, 63 (2019).
- [88] S. Gao, S. A. Guan, A. D. Fouad, J. Meng, T. Kawano, Y.-C. Huang, Y. Li, S. Alcaire, W. Hung, Y. Lu, Y. B. Qi, Y. Jin, M. Alkema, C. Fang-Yen, and M. Zhen, *Excitatory motor neurons are local oscillators for backward locomotion*. eLife **7** (2018).
- [89] T. Xu, J. Huo, S. Shao, M. Po, T. Kawano, Y. Lu, M. Wu, M. Zhen, and Q. Wen, *Descending pathway facilitates undulatory wave propagation in Caenorhabditis elegans through gap junctions*, Proceedings of the National Academy of Sciences **115**, E4493 (2018).
- [90] S. Revzen and J. M. Guckenheimer, *Finding the dimension of slow dynamics in a rhythmic system*, Journal of The Royal Society Interface **9**, 957 (2011).
- [91] G. J. Berman, W. Bialek, and J. W. Shaevitz, *Predictability and hierarchy in drosophila behavior*, Proceedings of the National Academy of Sciences **113**, 11943 (2016).
- [92] K. Pyragas, *Continuous control of chaos by self-controlling feedback*, Physics letters A **170**, 421 (1992).
- [93] E. A. Jackson, *The opcl control method for entrainment, model-resonance, and migration actions on multiple-attractor systems*. Chaos: An Interdisciplinary Journal of Nonlinear Science **7**, 550 (1997).
- [94] C. Fang-Yen, M. Wyart, J. Xie, R. Kawai, T. Kodger, S. Chen, Q. Wen, and A. D. Samuel, *Biomechanical analysis of gait adaptation in the nematode Caenorhabditis elegans*, Proceedings of the National Academy of Sciences **107**, 20323 (2010).
- [95] J. Karbowski, C. J. Cronin, A. Seah, J. E. Mendel, D. Cleary, and P. W. Sternberg, *Conservation rules, their breakdown, and optimality in caenorhabditis sinusoidal locomotion*, Journal of theoretical biology **242**, 652 (2006).
- [96] M. Backholm, A. Kasper, R. Schulman, W. Ryu, and K. Dalnoki-Veress, *The effects of viscosity on the undulatory swimming dynamics of C. elegans*, Physics of Fluids **27**, 091901 (2015).
- [97] S. Berri, J. H. Boyle, M. Tassieri, I. A. Hope, and

- N. Cohen, *Forward locomotion of the nematode *C. elegans* is achieved through modulation of a single gait*, HFSP journal **3**, 186 (2009).
- [98] V. J. Butler, R. Branicky, E. Yemini, J. F. Liewald, A. Gottschalk, R. A. Kerr, D. B. Chklovskii, and W. R. Schafer, *A consistent muscle activation strategy underlies crawling and swimming in *Caenorhabditis elegans**, Journal of the Royal Society Interface **12**, 20140963 (2015).
- [99] D. J. Evans and W. Hoover, *Flows far from equilibrium via molecular dynamics*, Annual review of fluid mechanics **18**, 243 (1986).
- [100] J. H. Boyle, S. Berri, and N. Cohen, *Gait modulation in *C. elegans*: an integrated neuromechanical model*, Frontiers in computational neuroscience **6**, 10 (2012).
- [101] Q. Wen, M. D. Po, E. Hulme, S. Chen, X. Liu, S. W. Kwok, M. Gershow, A. M. Leifer, V. Butler, C. Fang-Yen, et al., *Proprioceptive coupling within motor neurons drives *C. elegans* forward locomotion*, Neuron **76**, 750 (2012).
- [102] D. Daems and G. Nicolis, *Entropy production and phase space volume contraction*, Physical Review E **59**, 4000 (1999).
- [103] E. M. Bollt and N. Santitissadeekorn, *Applied and computational measurable dynamics*, Vol. 18 (SIAM, 2013).
- [104] P. M. Riechers and J. P. Crutchfield, *Fraudulent White Noise: Flat power spectra belie arbitrarily complex processes*, arXiv e-prints, arXiv:1908.11405 (2019).
- [105] C. Brennan and A. Proekt, *A quantitative model of conserved macroscopic dynamics predicts future motor commands*, eLife **8** (2019).
- [106] G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé, *Identification of slow molecular order parameters for Markov model construction*, Journal of Chemical Physics **139**, 1 (2013).
- [107] C. R. Schwantes and V. S. Pande, *Improvements in Markov State Model construction reveal many non-native interactions in the folding of NTL9*, Journal of Chemical Theory and Computation **9**, 2000 (2013).
- [108] S. Linderman, A. Nichols, D. Blei, M. Zimmer, and L. Paninski, *Hierarchical recurrent state space models reveal discrete and continuous dynamics of neural activity in *C. elegans**, bioRxiv (2019), 10.1101/621540.
- [109] B. C. Daniels, W. S. Ryu, and I. Nemenman, *Automated, predictive, and interpretable inference of *Caenorhabditis elegans* escape dynamics*, Proceedings of the National Academy of Sciences **116**, 7226 (2019).
- [110] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control* (Cambridge University Press, 2019).
- [111] M. Casdagli, T. Sauer, and J. A. Yorke, *Embedology*, J. Stat. Phys **65**, 579 (1991).
- [112] J. Stark, D. S. Broomhead, M. Davies, and J. Huke, *Delay embeddings for forced systems. ii. stochastic forcing*, Journal of Nonlinear Science **13**, 519 (2003).
- [113] M. R. Muldoon, D. S. Broomhead, J. P. Huke, and R. Hegger, *Delay embedding in the presence of dynamical noise*, Dynamics and Stability of Systems **13** (1998).
- [114] J. Stark, *Delay embeddings for forced systems. i. deterministic forcing*, Journal of Nonlinear Science **9**, 255 (1999).
- [115] J. P. Huke and D. S. Broomhead, *Embedding theorems for non-uniformly sampled dynamical systems*, Nonlinearity **20**, 2205 (2007).
- [116] E. R. Deyle and G. Sugihara, *Generalized theorems for nonlinear state space reconstruction*, PLoS One **6**, e18295 (2011).
- [117] J. D. Farmer, *Information dimension and the probabilistic structure of chaos*, Zeitschrift für Naturforschung A **37**, 1304 (1982).
- [118] J. D. Farmer and J. J. Sidorowich, *Predicting chaotic time series*, Physical review letters **59**, 845 (1987).
- [119] E. P. Xing, M. I. Jordan, S. J. Russell, and A. Y. Ng, in *Advances in neural information processing systems* (2003) pp. 521–528.
- [120] J. Pathak, Z. Lu, B. R. Hunt, M. Girvan, and E. Ott, *Using machine learning to replicate chaotic attractors and calculate lyapunov exponents from data*, Chaos: An Interdisciplinary Journal of Nonlinear Science **27**, 121102 (2017).
- [121] K. Champion, P. Zheng, A. Y. Aravkin, S. L. Brunton, and J. N. Kutz, *A unified sparse optimization framework to learn parsimonious physics-informed models from data*, arXiv preprint arXiv:1906.10612 (2019).
- [122] MATLAB, *version 9.3.0 (R2017b)* (The MathWorks Inc., Natick, Massachusetts, 2017).
- [123] J. E. Sulston and S. Brenner, *The DNA of *C. elegans**, Genetics **77**, 95 (1974).
- [124] E. N. Lorenz, *Deterministic Nonperiodic Flow*. Journal of Atmospheric Sciences **20**, 130 (1963).
- [125] R. Vautard and M. Ghil, *Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series*, Physica D: Nonlinear Phenomena **35**, 395 (1989).
- [126] H. Kantz, *A robust method to estimate the maximal lyapunov exponent of a time series*, Physics letters A **185**, 77 (1994).
- [127] E. R. Deyle, R. M. May, S. B. Munch, and G. Sugihara, *Tracking and forecasting ecosystem interactions in real time*, Proceedings of the Royal Society B: Biological Sciences **283**, 20152258 (2016).
- [128] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Applied mathematical sciences (Springer, New York, 1983).
- [129] H. D. I. Abarbanel, R. Brown, and M. B. Kennel, *Variation of Lyapunov exponents on a strange attractor*, Journal of Nonlinear Science **1**, 175 (1991).
- [130] H. D. I. Abarbanel, R. Brown, and M. B. Kennel, *Local lyapunov exponents computed from observed data*, Journal of Nonlinear Science **2**, 343 (1992).

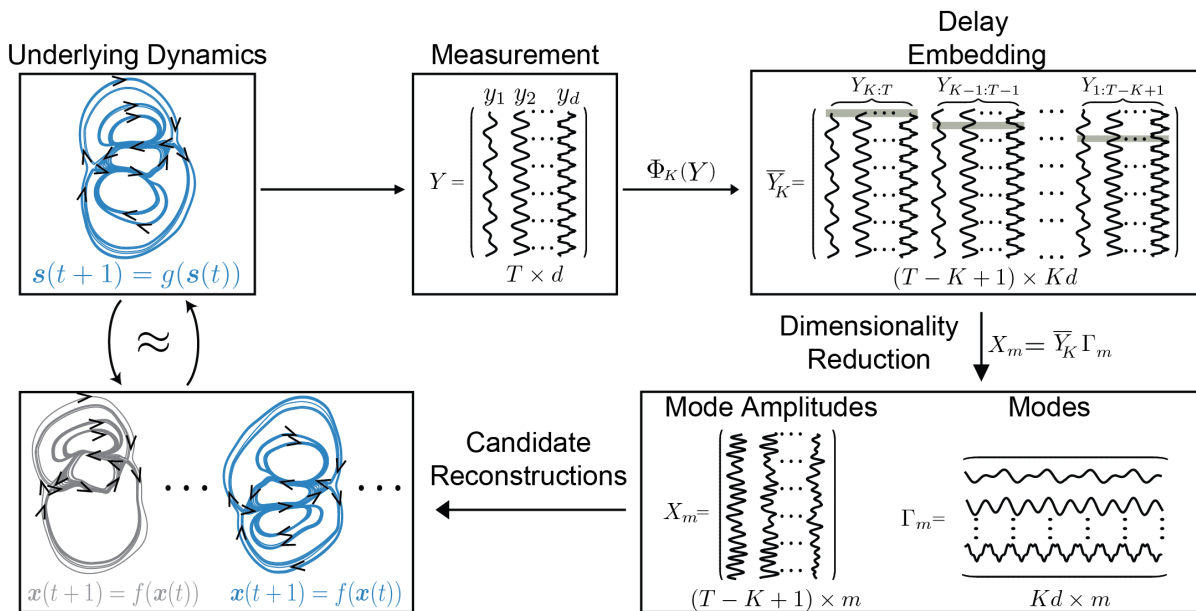


FIG. 1. State Space Reconstruction. (From Upper Left) A d -dimensional time series of an underlying dynamical system is collected in the measurement matrix Y . The embedding operation Φ_K stacks delayed copies of the measurements within a short time window of length K into a matrix \bar{Y} of size $(T-K+1) \times Kd$. Dimensionality reduction of \bar{Y} results in an $m < Kd$ dimensional state space spanned by the basis vectors Γ_m called the modes. The mode coefficients X_m form an approximation of the state space of the underlying dynamics. Each value of K and m results in a different state space reconstruction of the underlying dynamics and we seek embedding parameters which maximize predictability.

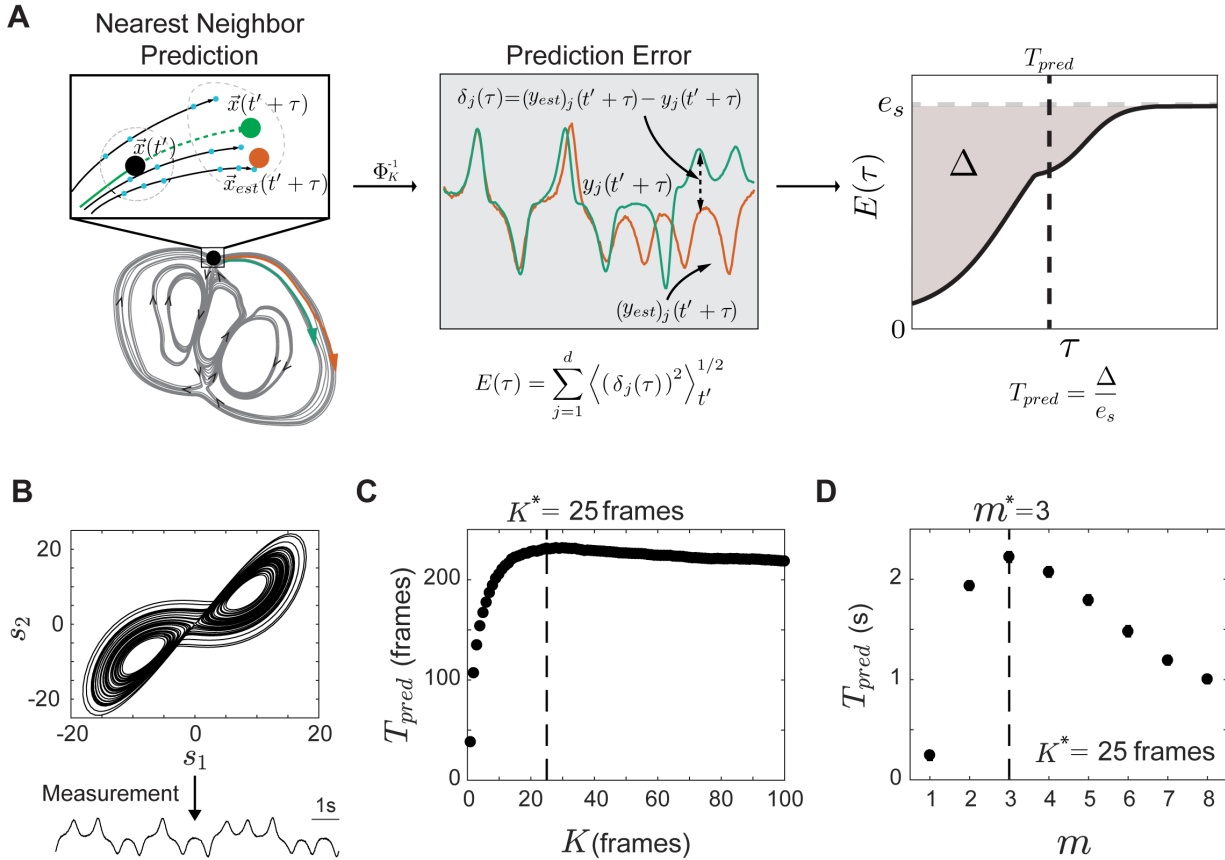


FIG. 2. Reconstruction Parameters for Maximum Predictability. (A) We use a nearest neighbor predictor, which averages the τ -step future of nearest neighbors of a point $\mathbf{x}(t)$ in the candidate reconstruction to estimate future measurements. $E(\tau)$ (Eq. 1) quantifies the error in prediction as a function of τ . In general for stationary complex systems, prediction gets worse with time and $E(\tau)$ grows before saturating to a value e_s , corresponding to the size of the system in state space. We define the characteristic timescale for $E(\tau)$ to saturate to e_s as a new measure of predictability, T_{pred} (Eq. 2), geometrically given by the area between e_s and $E(\tau)$ divided by e_s . (B-D) We apply our reconstruction to the time series corresponding to a noisy measurement of the first coordinate of the Lorenz system with standard chaotic dynamics. (C) The prediction time T_{pred} varies for different K , with a maximum $K^* \approx 0.25$ s. (D) T_{pred} is maximal for the reconstruction defined by K^* and the first three singular vectors, resulting in a 3D embedding of the chaotic attractor, which has fractal dimension $D \approx 2.05$. Errorbars in (C,D) are comparable to the point size.

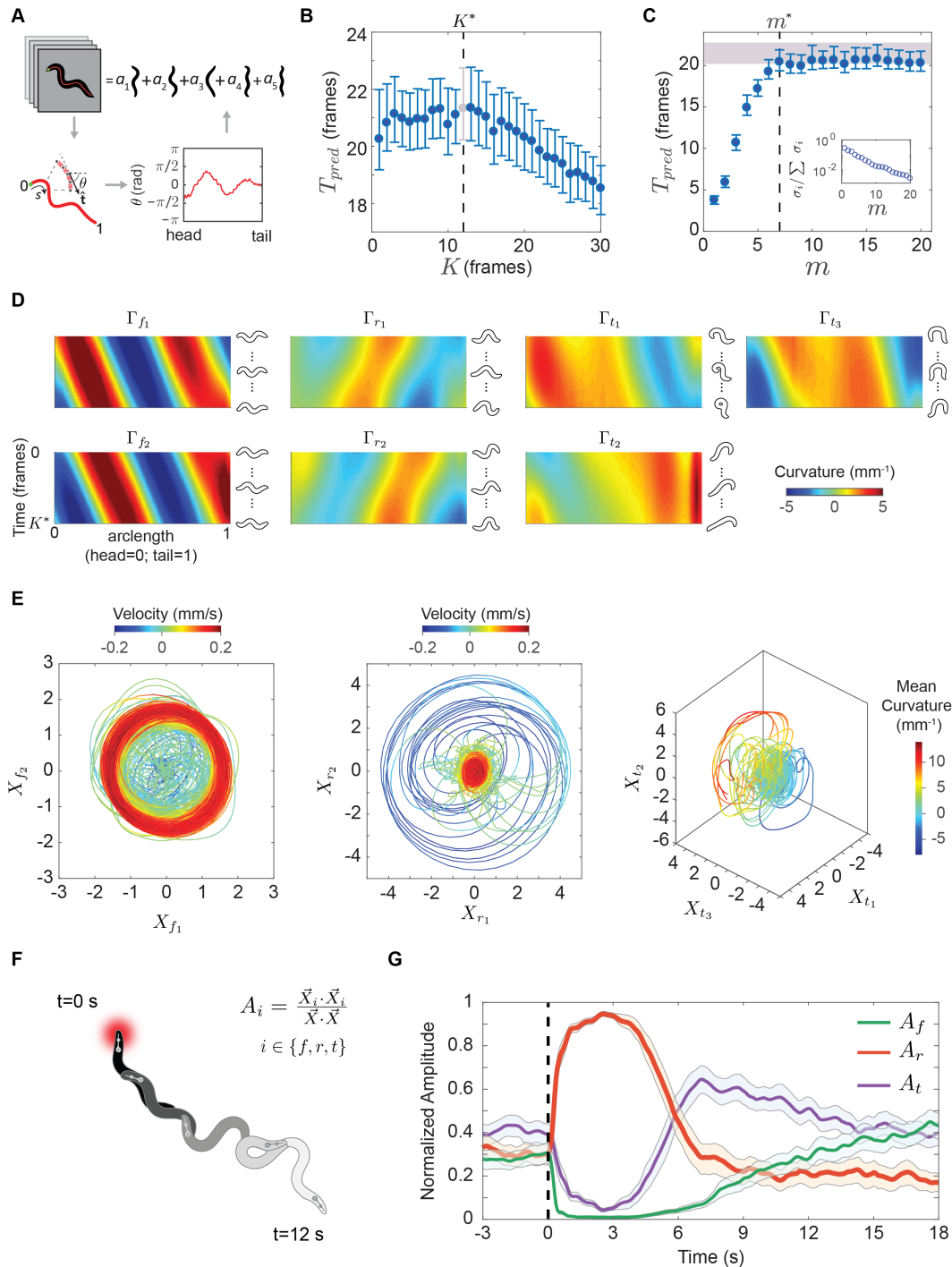


FIG. 3. The Low-Dimensional State Space of *C. elegans* Locomotion. (A) We parameterize worm posture by eigenworm projections (B-E) The embedding process for the foraging behavior of an example worm. (B) T_{pred} as a function of K . We set $K^* = 12$, $\sim 1/2$ cycle of the worm's body wave. (C) T_{pred} as a function of m . Embeddings with dimensions beyond $m^* = 7$ carry little predictive structure. The gray bar denotes $T_{pred}(K = K^*)$. (C, inset) The normalized singular value spectrum does not show an obvious cut-off. (D) We decompose the 7D embedding into linear combinations of independent posture sequences of length K^* , which we denote behavioral modes. We show the modes as curvature kymographs and note that there are two approximately conjugate pairs and a third group with three modes. Γ_{f1} and Γ_{f2} correspond to the body waves of forward locomotion, Γ_{r1} and Γ_{r2} to backward locomotion, and Γ_{t1} , Γ_{t2} and Γ_{t3} to deep body bends. (E) Trajectories visualized as projections onto the above mode combinations. In the (X_{f1}, X_{f2}) , and (X_{r1}, X_{r2}) planes, the trajectories are colored by centroid velocity, negative (blue) for backward locomotion, and positive (red) for forward. Trajectories in the (X_{t1}, X_{t2}, X_{t3}) space are colored by the mean body curvature (blue for dorsal, red for ventral). (F-G) We apply our embedding to the behavior of worms escaping from a heat impulse to the head. (F) Schematic of the response. (G) We project the escape dynamics using the ensemble foraging modes (Fig. S3) and visualize the dynamics through normalized amplitudes of the forward, reversal and turning projections. Reversal dynamics (red) initiate just after the impulse (dotted line) while turning dynamics (purple) are suppressed. The reversal ends with the initiation of an Ω -turn.

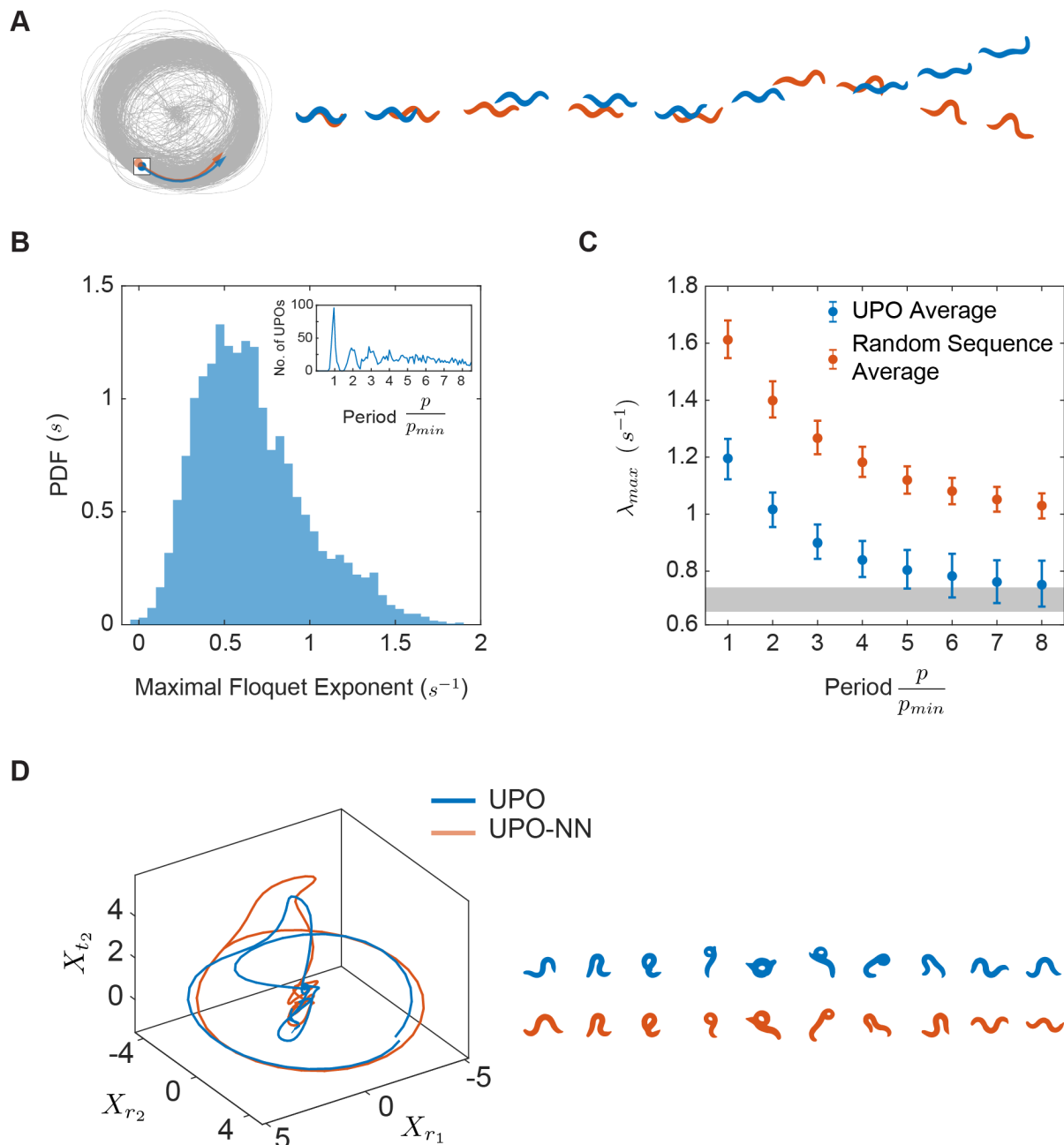


FIG. 4. Chaotic State Space Dynamics Reveal a Deterministic Component of Behavioral Variability (A) Neighboring state space points correspond to similar posture sequences, but the futures of these points diverge at longer times. (B) We identify periodic orbits in the state space using the ensemble embedding and quantify the stability of each periodic orbit through its maximal Floquet exponent. The identified periodic orbits are predominantly unstable, as also observed in known chaotic systems. (B, inset) Histogram of the number of periodic orbits across all worms as a function of their period exhibits peaks at integer multiples of a minimum period p_{min} corresponding to the period of a body wave during forward locomotion. (C) We recover the characteristic divergence of the trajectories λ_{max} (grey bar) by a weighted average of the maximal Floquet exponent of across orbits (blue). We quantify the state-space divergence across worms, Fig. S6, and find an exponential regime, yielding an average maximal Lyapunov exponent $\lambda_{max} = 0.69$ (0.65, 0.74) s^{-1} . In contrast, an average from random sequences of the same length converges more slowly (red). (D) We show an example period-4 orbit composed of a sequence of reversal, deep body bend, and forward movement (blue). This same orbit is revisited at a later time (orange), resulting in a similar sequence of postures for these two long trajectories (right).

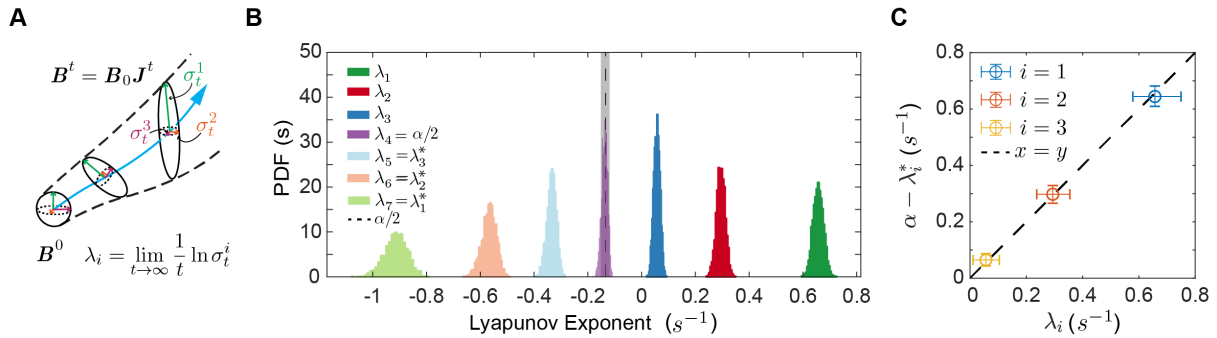


FIG. 5. Symmetric Lyapunov Spectrum Indicates Damped-Driven Hamiltonian Dynamics (A) The contraction and expansion of local volumes, suitably averaged across the full state space, provides important dynamical invariants and is quantified by a collection of exponents: the Lyapunov spectrum. We sketch the transformation of an initial volume B^0 to a new volume B^t by the flow (blue curve) and the local Jacobian J^t . The spectrum of Lyapunov exponents is given by the size of the principal axes of B^t for $t \rightarrow \infty$. **(B)** The Lyapunov spectrum for *C. elegans* foraging dynamics where the distribution for each exponent is constructed from bootstrap samples across different worms. We find two positive (chaotic) exponents driving variability, a near-zero exponent indicative of continuous (non-noisy) dynamics and four negative exponents which drive stereotypy. The spectrum is symmetric about the point $\alpha/2$ which coincides with λ_4 . **(C)** Lyapunov exponents come as conjugate pairs that sum to α , a symmetry suggestive of damped-driven Hamiltonian system.

Supplementary Material

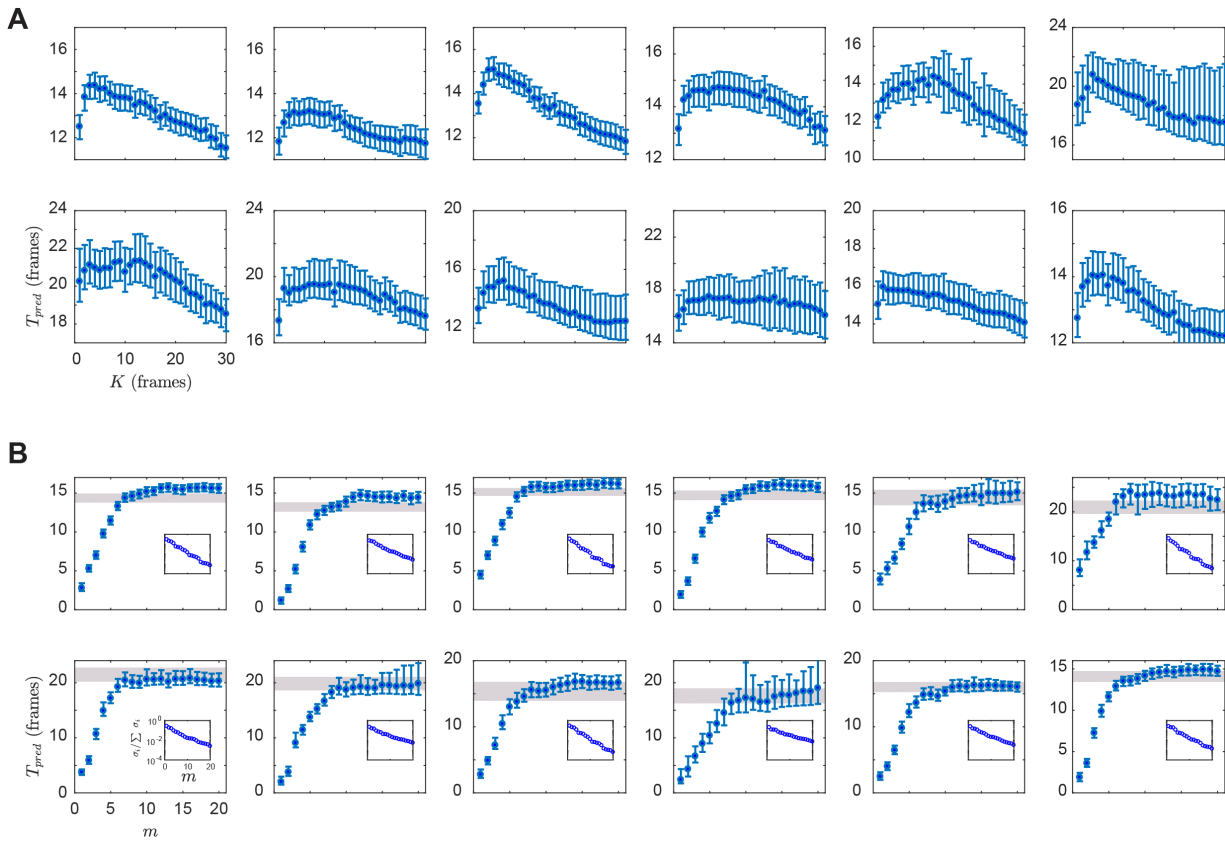


FIG. S1. **State space embedding for all the worms in the foraging dataset.** Individual worm estimates of T_{pred} as a function of K are shown in (A), and as a function of m in (B). We find similar curves across worms, despite the differences in their detailed dynamics. Note that while the distance metric in the SVD space (B) and the space of delays (A) is different, which could result in inconsistencies in the estimation of T_{pred} , we find only minor differences between the maximum T_{pred} in the two cases (gray bar in B). Inset shows the normalized singular value spectrum, which does not have a clear cut-off for any worm.

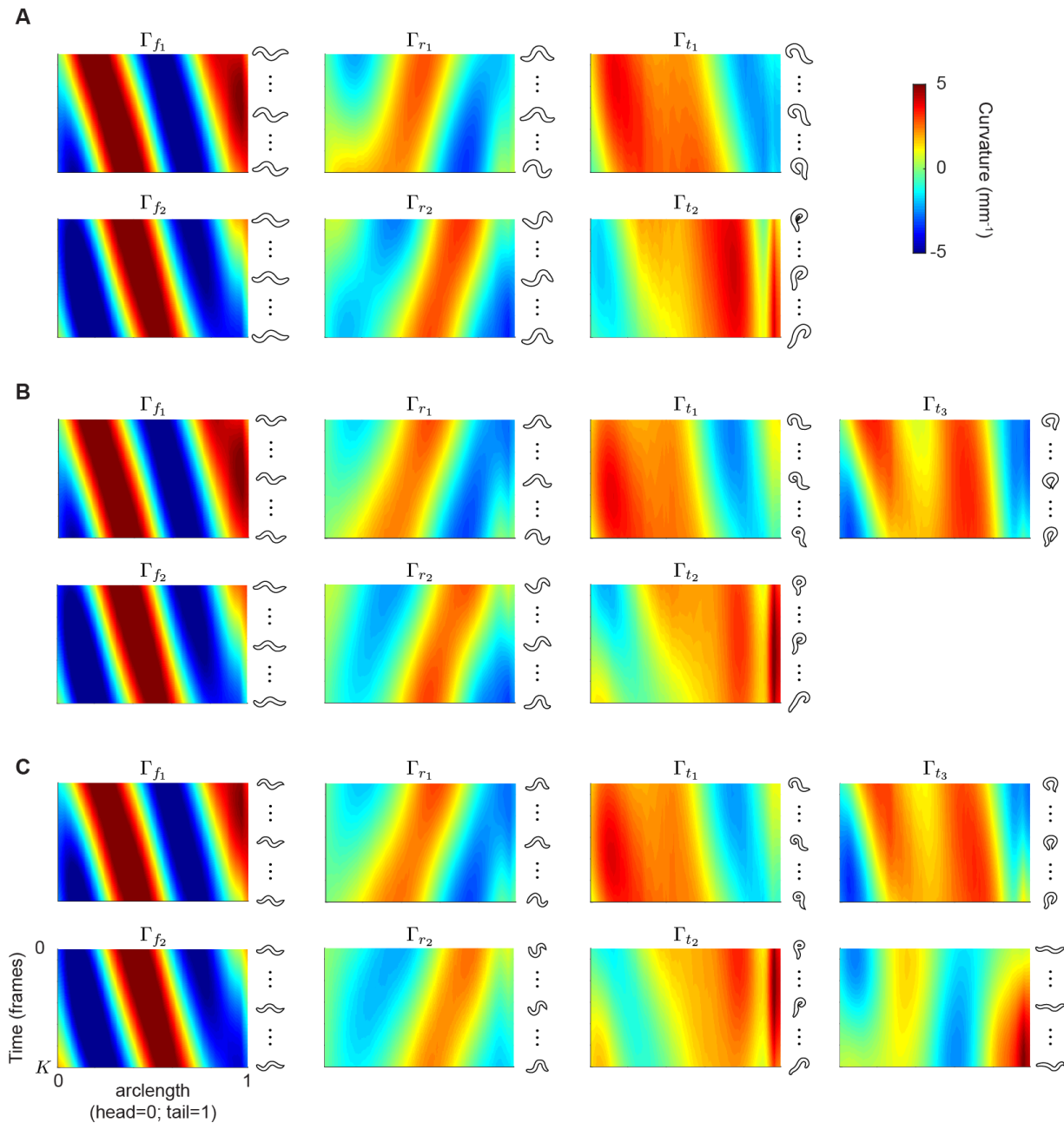


FIG. S2. Dominant state space modes are stable across different embedding dimensions, with distinct groups independently capturing forward, reversal and turning behaviors (A-C) Behavioral modes estimated for the worm in Fig. 3 for dimensions, $m = 6$ (A), $m = 7$ (B), and $m = 8$ (C); embedding window is set to $K = 12$ frames. The modes retain their interpretability across dimensions. In a 6-dimensional embedding, there are two forward, two backward and two turning modes. In 7 dimensions one of the turning modes further splits into an omega-turn like mode (Γ_{t_3}) and a delta-turn like mode (Γ_{t_1}), while Γ_{t_2} changes little. Furthermore, the reversal modes are more separable in 7 dimensions. The 8-dimensional state space retains the forward, reversal and turning dynamics along with an additional and subtle head-bending.

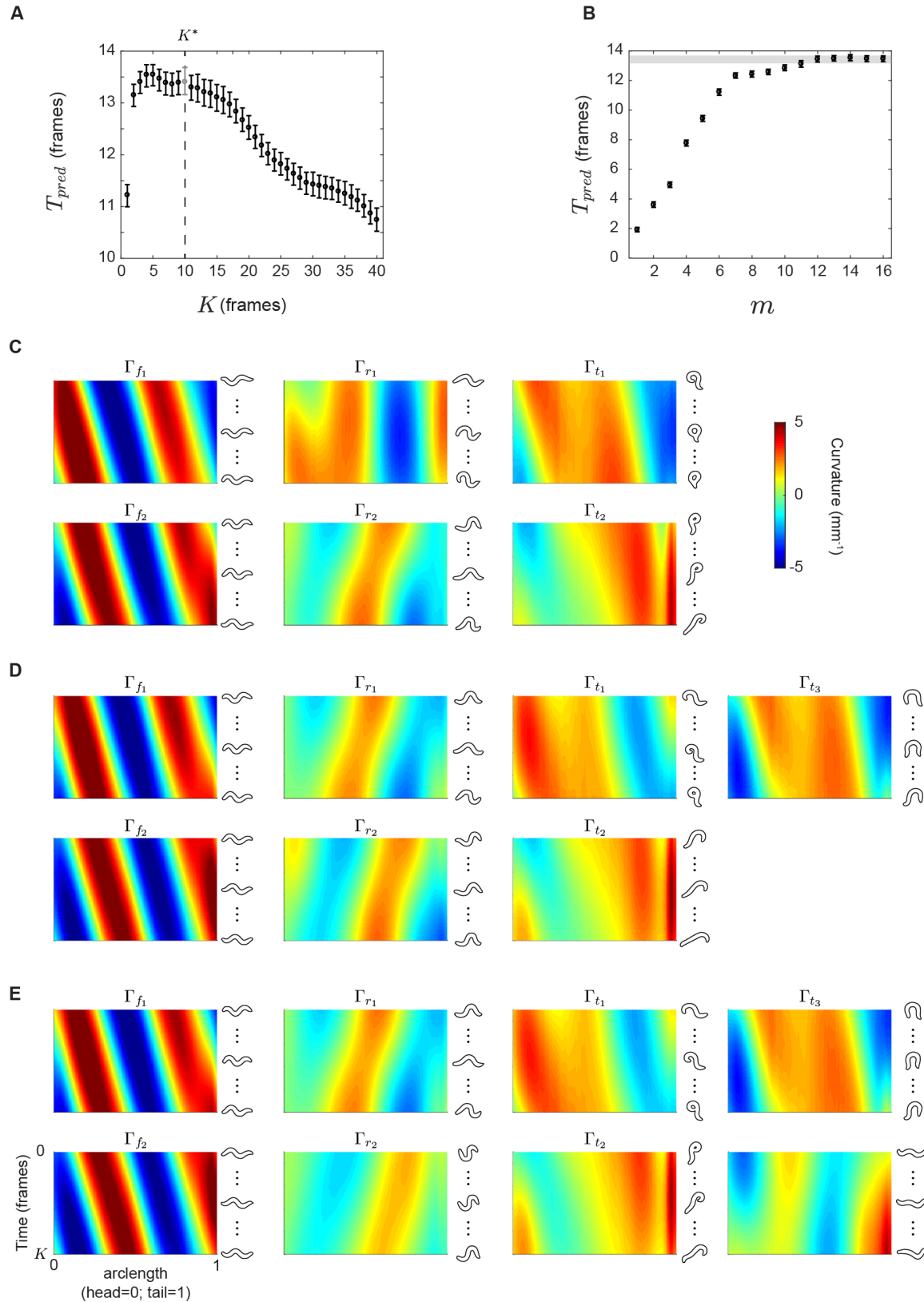


FIG. S3. The ensemble embedding across all ($N=12$) worms is constructed from their concatenated posture time series and characterized by $K^* = 10$ and $m^* = 7$. (A-B) T_{pred} as a function of K and m . We set $K^* = 10$, approximately when T_{pred} begins to decrease, and show $T_{pred}(m)$ at this K^* . We show the resulting modes for $m = 6$ (C), $m = 7$ (D), and $m = 8$ (E), and these are qualitatively similar to those obtained from our representative worm, Fig. S2. The additional modes present for embeddings greater than $m^* = 7$ offer only much smaller improvement in predictability and appear independent of forward, reversal and turning behaviors.

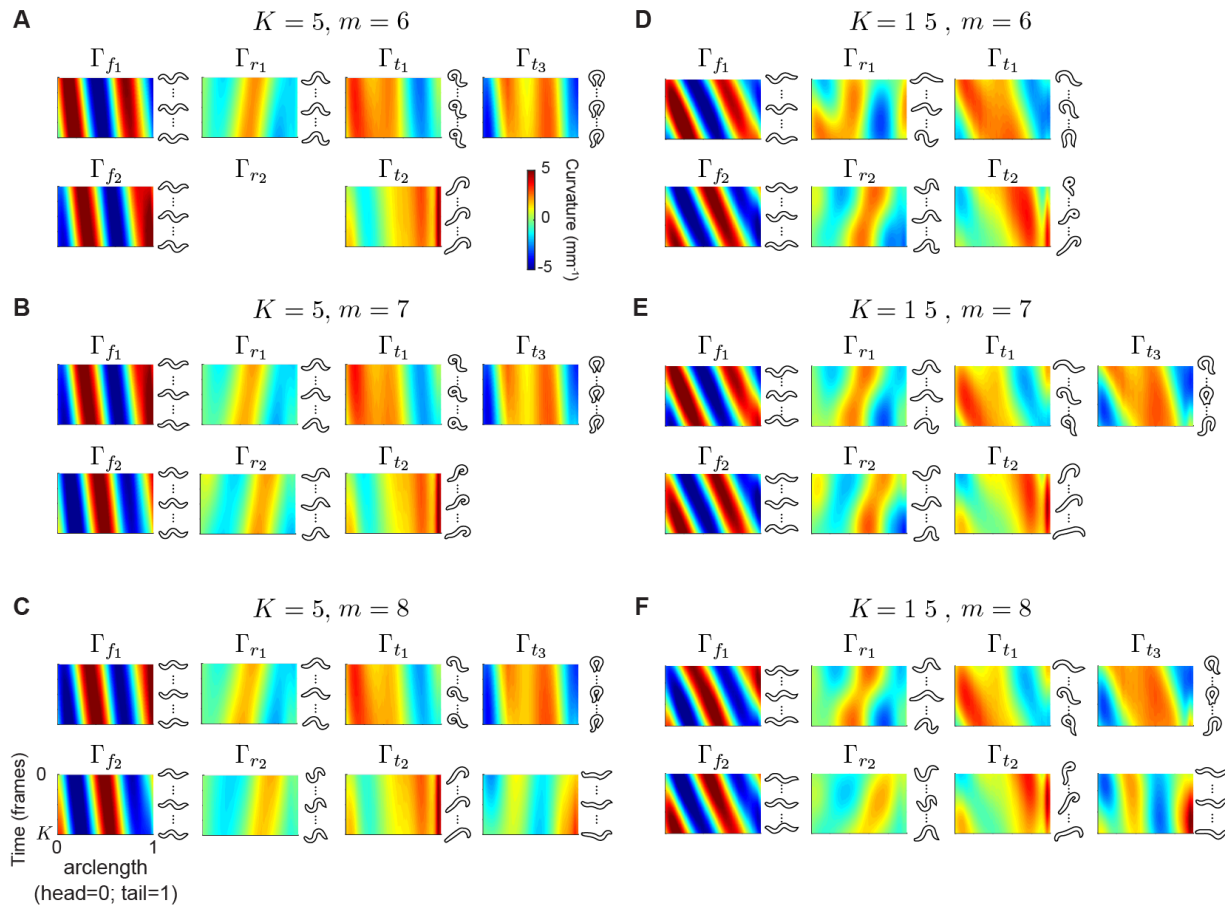


FIG. S4. **Ensemble embedding for different values of K and m** (A-C) Behavioral modes estimated from the ensemble for $K = 5$, and dimensions, $m = 6$ (A), $m = 7$ (B), and $m = 8$ (C). (D-F) Same as above but for $K = 15$. The modes are qualitatively similar across this variation.

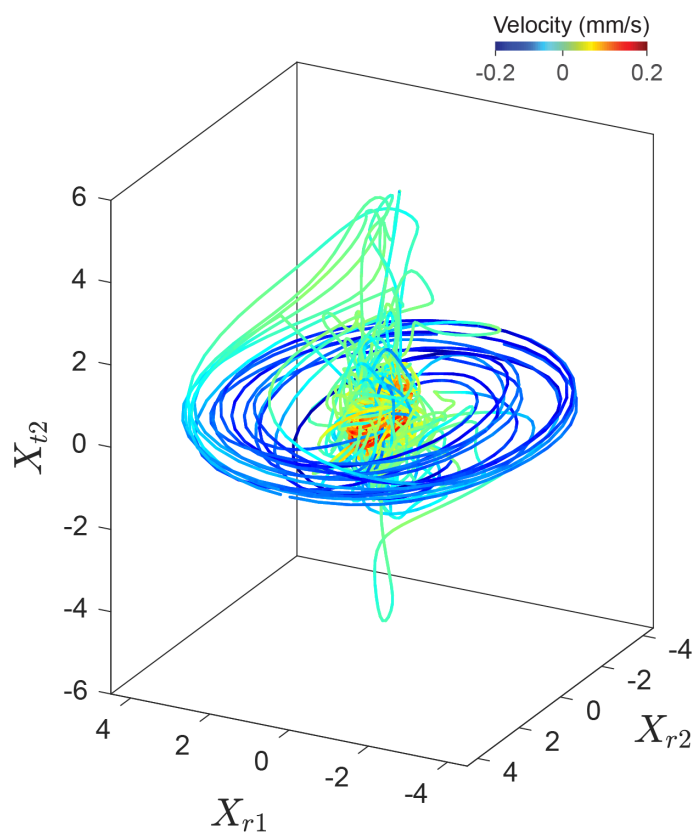


FIG. S5. State space captures a commonly observed sequence where long reversals transition to forward via a deep body bends seen here as a large excitation in X_{t2} as the reversal ends. Here we see that the blue (backward) and red (forward) bundles are smoothly connected via a large transient along the turning mode X_{t2}

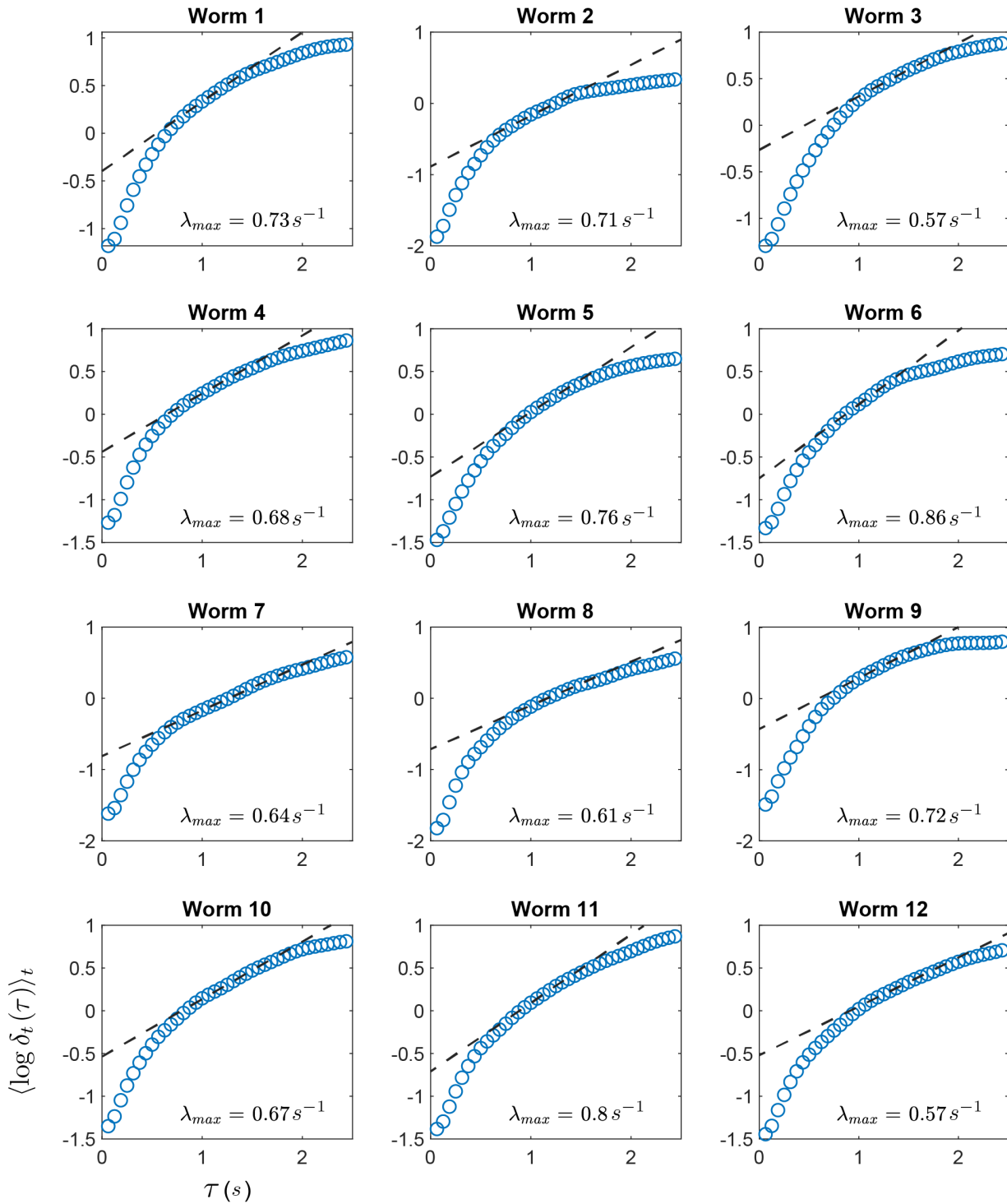


FIG. S6. To quantify the state space divergence we plot the logarithm of the average distance between a trajectory and its nearest neighbors, averaged over several starting reference trajectories. For each worm we find that after a transient, there is linear region showing exponential divergence. The slope of the linear region provides an estimate of the maximal Lyapunov exponent λ_{max} and the positive exponents are an indication of chaos in worm behavior.

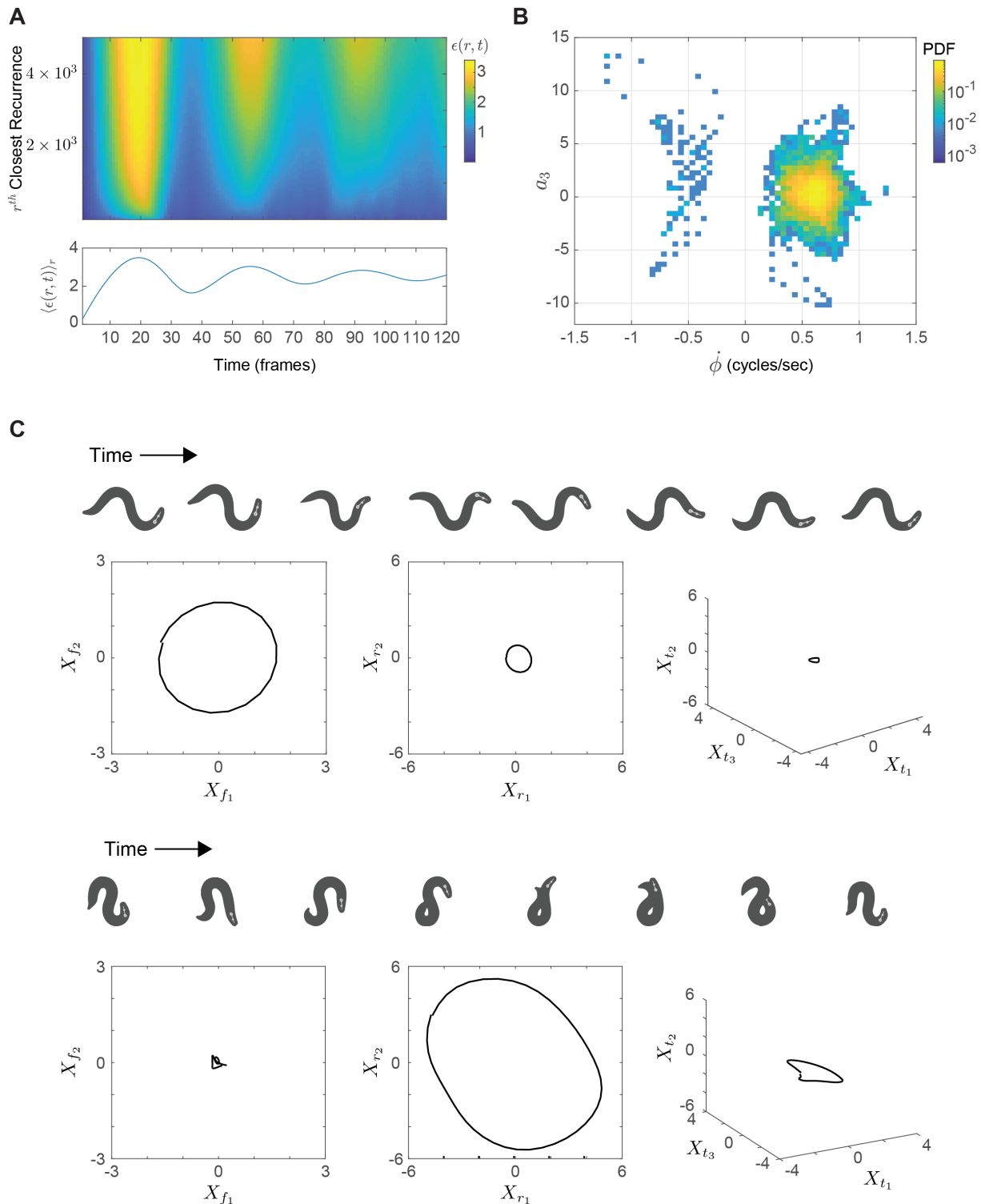


FIG. S7. (A) The recurrence function $\epsilon(r, t)$ from the same worm in Fig 3 for 120 frames and 5000 closest recurrences (top). Local minima of this function at times t_* , as seen in the average $\langle \epsilon(r, t) \rangle_r$ shown below correspond to close recurrences and identify periodic orbits of length t_* . The first local minimum is the smallest period p_{min} , which is 37 frames in this example. For a given value of r , $\epsilon(r, t_*)$ gives the distance threshold at which we must look to find a periodic orbit of length t_* . (B) Probability distribution of phase velocities $\dot{\phi}$ and third eigenworm coefficient a_3 , which is proportional to mean body curvature, across all period-1 orbits of duration p_{min} from all worms in the dataset. We see two clusters corresponding to forward and backward locomotion, as well as orbits with a dorsal or ventral bias (e.g. orbits at bottom right and top left). (C) Example period-1 orbits from the same worm in (A) corresponding to forward (top) and backward (bottom) locomotion.

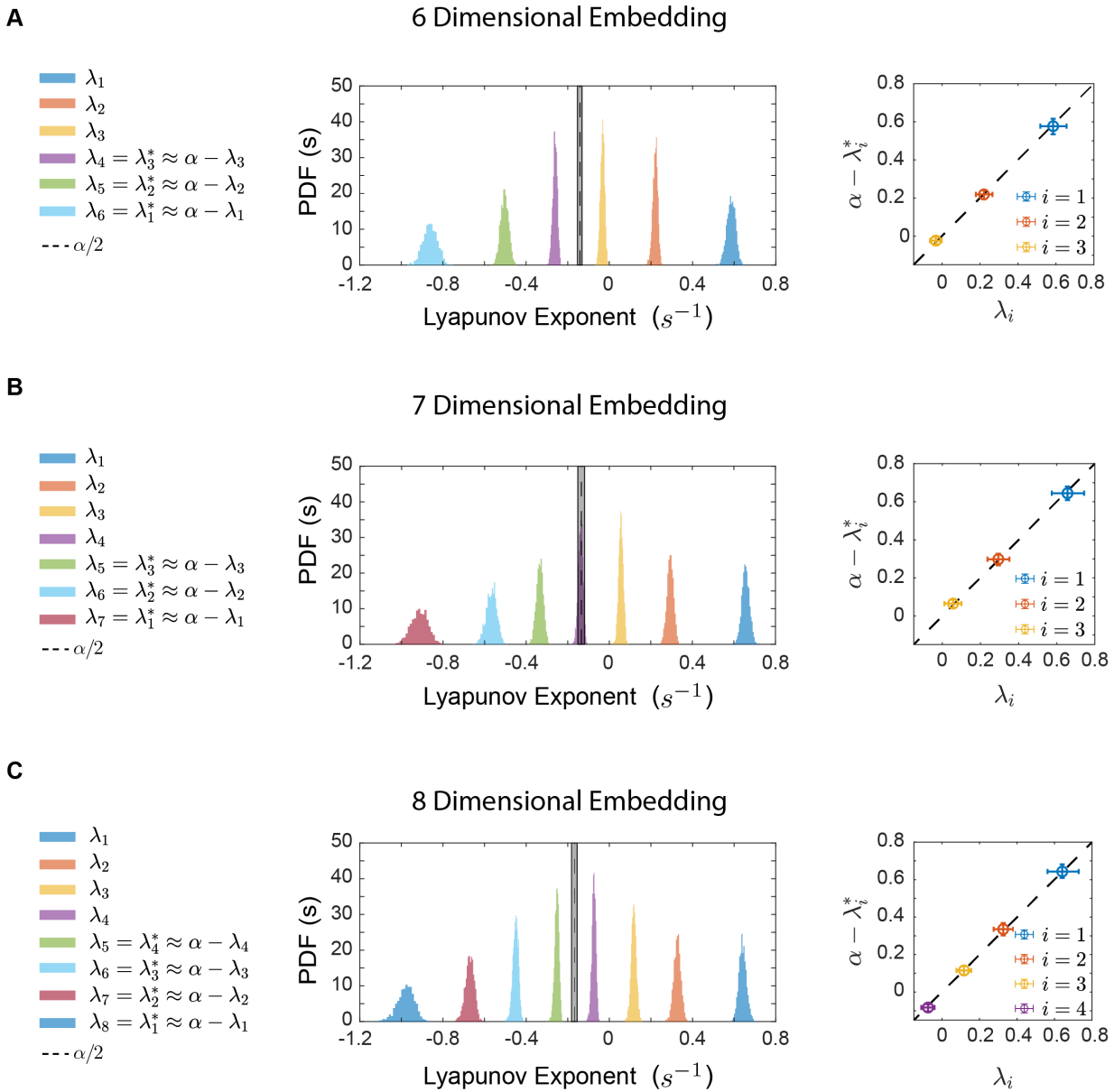


FIG. S8. Lyapunov spectra computed from reconstructions of worm behavior in different embedding dimensions. Conjugate pairing of Lyapunov exponents is robustly observed in dimensions 6 and above.

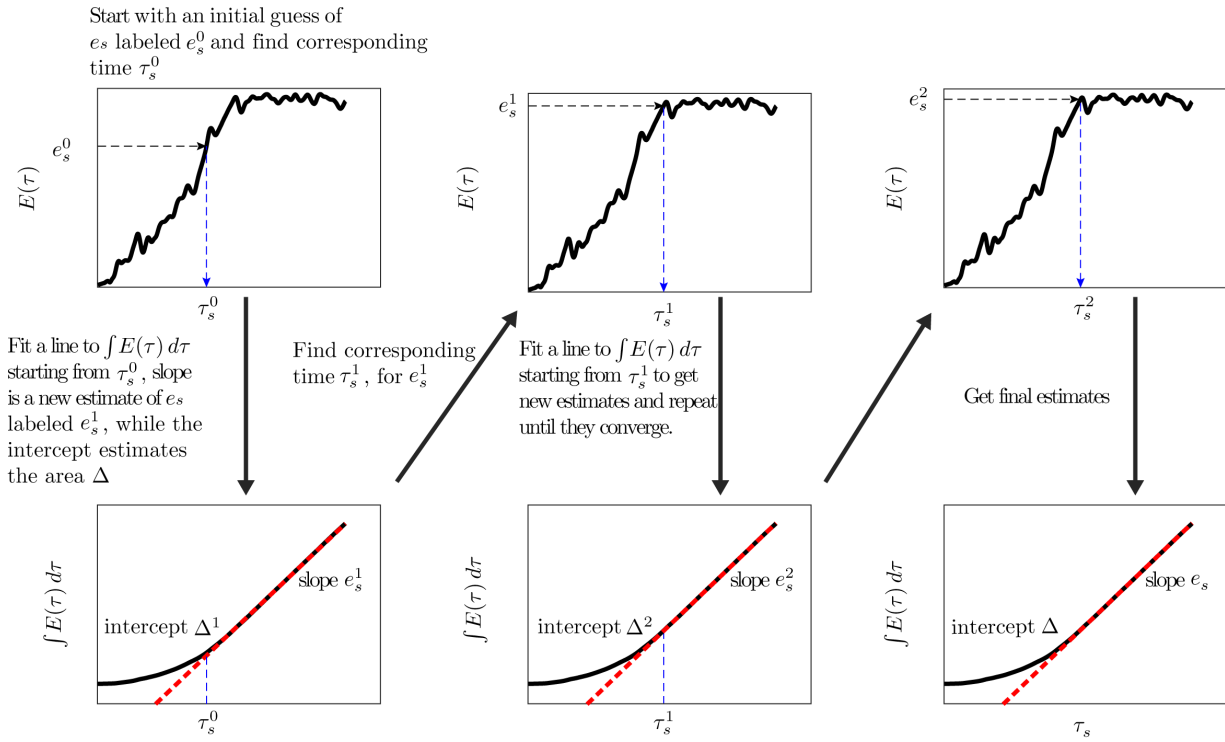


FIG. S9. A schematic showing the fixed point algorithm for robust estimation of the asymptote e_s , and the area Δ .