Supplementary Information

How much of a nervous-system model does behaviour identify?

Chenxi He, Cavendish Laboratory, University of Cambridge
ch2067@cam.ac.uk

Contents

This Supplementary Information provides the extended methods and the complete per-experiment quantities underlying each main-text claim. It comprises: the measurement methods and their controls; the provenance of the external biological datasets; the confirmatory experiments that extend the probes to higher parameter dimension, to named single channels, to the full connectome, to the internal neural state, and to a fifth simulator; and the experiments that identify the manifold with the eigenworm space, establish its complementary tiling by behaviour, and demonstrate the structure across two phyla, in a nematode, an insect and a crustacean, together with one inconclusive control that is reported in full. Every quantity is drawn from a saved result file, named per section, and is regenerated from that file by the accompanying reproduction code.

1 Measuring the manifold: methods

1.1 Multistart parameter drift

For each simulator we run N independent optimisations (separable CMA-ES or SPSA) from independent starts against one *fixed* behavioural target, (Table 1) and record (i) behavioural recovery (fraction of the random-to-target behaviour-distance closed), (ii) the normalised per-dimension dispersion of the N recovered parameter vectors, and (iii) the PCA participation ratio of the solution cloud. A behaviourally-converged but parameter-dispersed cloud is direct evidence of a positive-dimensional equivalence manifold.

Table 1: **Fixed-target multistart optimisation on two simulators.** Each row reports N independent optimisations started from independent points and fitted to a single fixed behavioural target. dim: number of free parameters; N: number of independent runs; success: runs reaching the behavioural target; behav. rec.: median fraction of the random-to-target behavioural distance closed; θ disp.: dispersion of the recovered parameter vectors, normalised per dimension; cloud dim: PCA participation dimension of the solution cloud. A solution-cloud dimension is not by itself evidence of a manifold. Run under the same optimiser and budget, a strictly convex problem with a unique optimum and no manifold at all gives a cloud of 20.2 of 126 dimensions, against the 21.7 measured here, because a participation dimension describes the shape of the scatter and not its size. The quantity that does separate the two is the magnitude of the dispersion: 0.0091 for the convex control against 0.712 for a construction with twenty flat directions, a factor of 78. The FlyGym value of 0.098 falls between them, an order of magnitude above the convex control and a factor of seven below the fully degenerate construction, which is what partial degeneracy should look like and is reported as such. Optimisers are separable CMA-ES (FlyGym) and SPSA (BAAIWorm); the flybody objective overfits per seed without a prior and is therefore excluded here and reported only in the curvature analysis.

sim / method	dim	N	success	behav. rec.	θ disp.	cloud dim

The FlyGym 126-d cloud occupies an effective 21.7 of 126 dimensions. (The flybody objective overfits per-seed without a prior and is excluded from the drift panel; it appears only in the Hessian analysis.) On the full coupled **real modWorm** model (OpenWorm connectome, 279 neurons), the manifold is read from the per-mechanism Hessian (S1.2): only 2 of 7 biophysical-mechanism directions carry behavioural curvature, the rest forming the flat manifold. (A direct random-perturbation probe gives the same qualitative picture but with an assay-dependent magnitude, so we report the manifold by its effective dimension, not by the raw sloppiness ratio, which is not comparable across simulators with different observables.) Files `modworm_REAL_manifold.json` (dispersion probe), `modworm_REAL_b1.json` (Hessian).

1.2 Hessian eigenspectrum

At each converged optimum we form the behavioural-loss Hessian by finite differences (low dimension) or a Lanczos approximation (high dimension) and report the effective dimension (number of eigendirections carrying 90%/99% of the spectral mass) and the count of near-flat directions ($|\lambda|$ at the finite-difference floor). On BAAIWorm, over 5 mechanism-class scalings at ε=5 $\times$ 10^{-3}, the eigenvalues are $\lambda_1 = +348$, $\lambda_2 = -340$ and $\lambda_3 = -214$ (a saddle), with the ion and passive eigendirections flat to machine precision over a parameter L2 length of 2.0; the sign structure is stable across $\varepsilon \in [10^{-4}, 5\times10^{-2}]$. On modWorm, the per-mechanism Gauss–Newton Hessian of the full coupled neural+body model on the OpenWorm connectome (1, 2) (279 neurons, 7 biophysical mechanisms) has effective dimension 2 of 7 at both thresholds with participation ratio 1.29; the stiff mechanisms are the synaptic and gap-junction conductances and the synaptic rise time, while the membrane capacitance and the sigmoid gain carry zero elasticity ($< 10^{-7}$) and so are unconnected in this configuration rather than sloppy, again placing connectivity above single-cell intrinsics. File `modworm_REAL_b1.json`.

1.3 Simulation-based calibration as a manifold probe

An amortised neural posterior (masked autoregressive flow (3), `sbi` (4)) is trained once on the prior, and per-dimension rank-based SBC (5, 6) is run (M/L posterior/rank draws); "pass" is a per-axis Kolmogorov–Smirnov test at α=0.05. Pass count: real modWorm 7/7 (N=300 coupled-model sims), read as 5/5 on the mechanisms that have a measurable effect: two of its seven axes are numerically unconnected (§1.5), and on an axis the data cannot reach the posterior equals the prior, which passes a rank-based calibration test by construction. Calibration counts should be read against the live-axis denominator for the same reason. FlyGym posterior-mean gain rel-error 0.018. The partial pass on FlyGym delimits the calibratable subspace of the manifold; the full pass on the lower-dimensional real-modWorm and BAAIWorm probes reflects a correctly recovered posterior.

1.4 Why the curvature is taken in Gauss–Newton form

A raw finite-difference Hessian is not interpretable away from a minimum. Evaluated over five global mechanism classes on BAAIWorm it carries mixed-sign eigenvalues (+348, −340, −214, 0, 0), so the point is a saddle and the sign of a curvature there says nothing about identifiability. Every dimension reported in this work is therefore taken from the positive semi-definite Gauss–Newton form $J^\top W J$, which is a curvature of the residual and is bounded below by zero by construction. File: `zerocompute_crosscheck.json`.

1.5 Finite-difference validity across simulators

A Gauss–Newton curvature assembled by central differences is meaningful only where the response to a perturbation rises above the numerical floor and remains close to proportional to the step. That condition is not automatic, and we tested it in every simulator whose curvature we report as a dimension. Where it fails, the failure is a property of the simulator rather than of the organism, and separating the two is part of what this measurement is for.

BAAIWorm. Scaling one connection weight, the whitened response is 1.43×10^{-5}, 2.84×10^{-5}, 7.10×10^{-5}, 1.42×10^{-4} and 2.84×10^{-4} at steps of 0.05, 0.1, 0.25, 0.5 and 1.0: the ratio to the step is constant at 2.84×10^{-4} to four significant figures across a twentyfold range. Repeat simulations at identical parameters are bit-identical, so the noise floor is exactly zero and no part of the measured curvature is finite-difference noise.

flyvis. The deposited measurement is taken at a single relative step of 0.2 with an absolute floor of 0.05, the response being the mean of the last five frames of the network's activation binned into 64 index-contiguous channels; no step scan is deposited for this model, and its dimension is reported under the below-cap rule, rising only slightly for the synaptic strengths.

FlyGym. A floor of 0.037 in the whitened joint-angle response is present already at a step of 0.001 and does not fall further, so at a step of 0.05 a large part of the measured difference is stochasticity of the controller rather than parameter sensitivity. Proportionality is recovered only above a step of about 0.25, and a scan across four steps confirms it: the response-to-step ratio is 1.98, 1.39, 1.34 and 1.31 at steps of 0.05, 0.15, 0.25 and 0.50, varying by a factor of 1.51 over the whole range but by only 1.02 across the valid region. Every FlyGym dimension reported in this work is therefore taken at 0.25. The accompanying null, in which columns are built from seed-to-seed variation at identical parameters, gives an effective dimension of 7 (90%) / 22 (99%), which is the floor against which the 11 / 31 measured with 4200 observables should be read.

Prinz–Marder pyloric circuit. This simulator fails the proportionality condition, and it fails it for every observable set we tried, which is why no curvature dimension is reported for it (§2.1).

Changing a single conductance by 0.1% already decorrelates the membrane-potential traces: the mean per-cell Pearson correlation between the baseline and perturbed traces is 0.028 at a step of 0.001 and stays between -0.03 and 0.19 at every larger step, because any parameter change shifts the phase of the rhythm and a pointwise comparison of two phase-shifted oscillations saturates at once. The whitened trace response is correspondingly flat, 1.35, 1.38, 1.39, 1.40 and 1.25 at steps of 0.001 to 0.25, so the Jacobian scales as $1/\delta$ and the normalised spectrum is invariant by construction. A step-size convergence test is therefore uninformative here rather than passed: the effective dimension is unchanged at 25.4, 25.7 and 25.7 for steps of 0.1, 0.05 and 0.025 precisely because it cannot change. The phase-invariant summary statistics used in this literature are the right response in principle, but measured the same way they saturate almost as strongly (median response-to-step ratio varying by 23.2 across a 25-fold step range, against 1 for a proportional response), so they do not rescue the measurement here.

modWorm. The posture response is discontinuous in the parameters. Scaling the synaptic rise time constant changes the body angles by 4×10^{-9} at +0.1% and 3×10^{-8} at +1% (numerically zero), then by 0.141 at +5% and by the same 0.141 at +10%. At the finite step used throughout (δ=0.25), two of the seven mechanism classes, the membrane capacitance (6.6×10^{-8}) and the sigmoid gain (2.5×10^{-8}), change the posture by 10^{-8} or less, that is, not at all: they are unconnected in this configuration rather than behaviourally sloppy, so the model has five mechanisms with a measurable effect and not seven. The leak conductance is a third case: numerically zero at infinitesimal steps but responding at the finite step (0.0195), which is why the infinitesimal five-class curvature of §1.2 and the finite per-mechanism elasticity are reported separately. We report modWorm quantities as finite-step sensitivity statements with both caveats attached; the connectome-scale claims rest on BAAIWorm.

The general point is that an effective dimension is a statement about a measurement as much as about a model, and it should be accompanied by the step-size scan, the noise floor and the count of parameters that do anything at all. Files `E44_baai_fd_controls.json`, `E33_stg_fd_convergence.json`, `E1_flygym_null.json`.

1.6 Assay specifications

Every curvature in this work is fully specified by its perturbation step and scheme, rollout, observable construction, whitening, and seeding. The table collects them; each script named in the results manifest hard-codes exactly these values.

assay	step δ (scheme)	rollout	observables	whitening	seeds
connectome, voltage	0.5 (multiplicative, central)	2000 ms	80 motor neurons × 120 samples = 9600	per-neuron temporal s.d., clip 10^{-3}	deterministic
displaced points	0.5	1500 ms	as above, 300-connection subsample	per-time across-neuron s.d.	point seeds 11/12/21/22
muscle readout	0.5	1500 ms	96 muscle activations over time	as displaced points	deterministic
named conductances	0.5	1500 ms	10,320 trajectory samples	per-neuron temporal s.d.	deterministic
20 class gains	0.25 (log-step)	120 steps	6 path statistics[a]	baseline magnitude	deterministic
FlyGym gains	0.05–0.5 scan; 0.25 reported	0.25–0.5 s at 10^{-4} s	42 joint angles × 100 frames = 4200	baseline magnitude	5 seeds averaged[b]
flyvis	0.2 relative, floor 0.05	20-frame grating	last-5-frame mean, 64 index bins	baseline magnitude $+10^{-6}$	deterministic
modWorm repertoire	0.25 $(1 \pm \delta)$	150 steps	6 kinematics[a]	baseline magnitude	subset cap 20, seed 0
cell panel	0.05 (log-space)	settled trace	12 per parameter[c]	Jacobian row scale	8/24 seeds × 3 folds

Table 2: **Assay specifications.** [a]Net displacement, path length, mean speed, speed s.d., r.m.s. heading change, and forward fraction (worm) or final curvature (modWorm). [b]The repertoire run of §6.5 uses one fixed seed; the controller is deterministic given the seed. [c]Amplitude quantiles (2–98%), spectral magnitudes and autocorrelation at lags 1–4, computed at full resolution and then reduced; optimiser CMA-ES, population 16, at most 400 iterations, seed $31s+7$ for start s; random-start folds $f \in \{1.5, 3, 10\}$. The twelve modWorm behaviours are the four touch presets (gentle/harsh × anterior/posterior) at stimulus scales 0.5, 1 and 2. flyvis probes all 65 biases, all 65 time constants and 200 of 604 synaptic strengths at uniform index spacing.

2 External references (no simulator used)

2.1 E1 — Prinz–Marder pyloric population

The published population of conductance-based stomatogastric models (7) is the canonical demonstration that many parameter sets produce one rhythm, and it is the only system here whose behaviourally equivalent set is enumerated rather than inferred. We use it in that form, and we do not report a curvature dimension for it, for the reason set out below.

What the population shows. Across the 2,365 networks that produce the canonical triphasic rhythm, the parameter cloud spans a participation dimension of 23.2 of 31 conductances, with 23 principal components reaching 90% of the variance and 30 reaching 99%. The behaviourally equivalent set is therefore close to full-dimensional: fixing the rhythm removes roughly a quarter of the conductance space and leaves the rest free. The cloud is not shapeless, however. About 1.1% of conductance pairs correlate at $|r| > 0.3$, recovering the disparate-but-co-regulated structure documented for this circuit (8–11), and the least variable conductances are the PY A-current and four graded inhibitory synapses (coefficient of variation 0.20–0.26) against the most variable, which are the AB/PD leak, AB/PD H-current and LP CaT (0.60–0.73). The worm models give the same placement of coupling (constrained) and fast inward currents (free) from the curvature, while the leak conductances sit on opposite sides in the two systems; the population ordering is obtained without any finite difference. File `E1_pyloric_manifold.json`.

Why no curvature dimension is reported. A Gauss–Newton curvature requires the response to be proportional to the step, and in this simulator it is not, for either observable set. Perturbing one conductance and measuring the whitened response at steps of 0.001, 0.01, 0.05, 0.1 and 0.25, the trace response is 1.35, 1.38, 1.39, 1.40 and 1.25 (medians over points): a 250-fold change in the step moves the response by under a tenth, so the Jacobian scales as $1/\delta$ and the *normalised* spectrum is invariant by construction. This is why a step-size convergence test is uninformative here rather than passed, and it is the correct reading of §1.5. It also explains an otherwise puzzling control: the convergence test of `E33` returns a stable dimension (25.4, 25.7, 25.7 at steps 0.1, 0.05, 0.025) and a leading-subspace cosine of 0.863 and 0.868 between successive steps, which by that test's own preregistered rule reads as real curvature. Both observations are consistent: a quantity that cannot vary with the step also cannot fail a test that only asks whether it varies with the step.

The phase-invariant summary statistics do not escape it either. Scanning all 31 conductances at 12 operating points and three steps, the ratio of the whitened summary-statistic response to the step varies by a median factor of 21.8 across a 25-fold step range, where a proportional response would give 1 and a fully saturated one 25; only 3.4% of parameter–point pairs are proportional. The result is unchanged under a second whitening by the across-population spread of each statistic (median 23.2) rather than by its own magnitude (median 23.9). The pyloric summary statistics are therefore about as saturated in the step as the raw traces, which we attribute to the burst-detection thresholds inside them changing discretely as the rhythm reorganises.

Neither the coarse nor the rich pyloric effective dimension passes the second of the three gates stated in the main text, so neither is reported. The underlying estimate is stable: the 90% effective dimension averages 2.13 over K=8 rhythm-gated points and 2.4 ± 0.9 over seven points of an expanded sweep, and a scan over 160 points holds the running mean flat within its standard error from K=4 to K=24 (2.00, 1.75, 2.08, 2.06, 2.33; point-to-point standard deviation 1.01). The proportionality failure, not the sample size, excludes the number. The sweep gates on burst period alone rather than on bursting, period and AB/PD<LP<PY phase ordering together, and admits 0.56% of 33,000 draws. Files: `E3_pyloric_rank.json`, `e1b_results/e1b_gate.json`, `E33_stg_fd_convergence.json`, `AB2_stg_scan.json`, `AB2b_stg_scaling.json`, `AB2c_stg_allparam.json`, `AB2d_stg_whitening.json`.

2.2 E2 — real neuron variability

Allen Cell Types electrophysiology, mouse and human cortical neurons, read for the eleven intrinsic properties the database defines for every cell. A cell enters only if all eleven are present, which admits 1914 of 2333 and requires no imputation. The coefficient of variation is taken on the raw values and is quoted only for the nine properties measured on a ratio scale; it is not defined for a membrane voltage, whose zero is a convention, so the resting potential and the fast-trough voltage are excluded from the range although they enter the dimension. Across the nine, the coefficient of variation runs from 0.34 for the upstroke-downstroke ratio to 1.85 for spike-frequency adaptation, with the firing rate at 1.11 and the input resistance at 0.48: shape properties are the least variable and rate properties the most. The dimension is computed on z-scored properties, since the eleven carry different units and an unstandardised covariance would follow whichever has the widest numerical range; the participation dimension of the property space is 4.78 of 11, with 7 components reaching 90% of the variance. Real neurons therefore span a wide variability gradient and still occupy a low-dimensional property manifold. Script `scripts/_e2_allen_variability.py`, file `E2_allen_variability.json`.

2.3 E4 — ion-channel-gene conservation

The genes are the 40 *C. elegans* loci encoding the mechanisms the worm models parameterise: the voltage-gated calcium, potassium and cation channels named in the conductance decomposition, the innexins that carry the gap junctions, and the synaptic machinery of the chemical connections. The comparison is therefore between the same mechanisms in the model and in evolution, not between an arbitrary gene set and a model. For each gene we query Ensembl Metazoa for orthologues in four congeners (*C. briggsae*, *C. remanei*, *C. brenneri*, *C. japonica*), take the highest amino-acid percent identity within each species where a gene has more than one orthologue, and average over the species present. Identity runs from 77.94 to 99.91 (median 91.20). The most conserved is the gap-junction innexin `unc-9` at 99.91, followed by the L-type calcium channel `egl-19` at 97.19; the least is the GABA-A receptor `unc-49` at 77.94. The connectivity axis the models make stiff is the

one evolution has held most nearly fixed. The correspondence is between orderings and not between individual genes: `egl-19`, whose conductance is among the sloppiest in the model, is the second most conserved gene at 97.19, since sequence conservation answers to selection on every function a channel serves and not to behavioural identifiability alone. File `E4_channel_conservation.json`.

2.4 E5 — isogenic behavioural manifold

OpenWorm Movement Database single-worm recordings of the N2 Bristol reference strain, one summary vector per animal taken as the `features_means` row of its Tierpsy feature file. Features missing in more than a fifth of animals are dropped, the remainder are median-imputed, features of zero variance are removed and the rest are z-scored before principal components are taken by singular value decomposition, leaving 683 features across 61 animals. Because the features outnumber the animals the decomposition has rank at most 60, so we report the intrinsic-dimension estimates rather than a component count near that cap: the cross-individual participation dimension is 9.75, with 24 components reaching 90% of the variance and 50 reaching 99% (first component 26.6%, second 10.7%). Animals of identical genotype and rearing therefore vary along roughly ten behavioural dimensions, which is the empirical counterpart of a parameter manifold. File `E5_n2_manifold.json`.

3 Unified same-region probe test (result)

Computing both analyses on the full coupled modWorm model in one shared prior region (per-mechanism log-multipliers, prior scale 0.15). **The two geometric analyses agree**: per-mechanism Hessian stiffness versus negative manifold drift give Spearman ρ=0.96 (p=5×10^{-4}, exact permutation p=0.0028; 7 mechanisms). Two of those seven are numerically unconnected (§1.5) and enter both rankings as ties at the bottom, which a rank correlation rewards, so the agreement covers five mechanisms with measurable effects plus two concordant zeros rather than seven independent ranks. **The amortised posterior is a calibration control, not a third ranking**: an NPE/MAF (`sbi`) trained on N=300 real modWorm simulations is well-calibrated on all 7 mechanisms (per-axis SBC 7/7, Kolmogorov–Smirnov $p > 0.05$), confirming the inference is sound; because a calibrated posterior calibrates every axis, SBC serves as a control rather than a per-axis identifiability ranking. We therefore report agreement between two geometric analyses plus a calibration control. Files: `modworm_REAL_sbc.json`, `modworm_REAL_b1.json`.

5 Confirmatory experiments

These measurements extend the picture above to higher parameter dimension, to named single channels, to the full connectome, and to the internal neural state. Behavioural curvature is reported as the Gauss–Newton Hessian H=$J^\top W J$ formed from the observable Jacobian J_{ik}=$\partial\,\mathrm{obs}_i/\partial\theta_k$ (whitened by per-observable scale W), the Fisher-information form that quantifies which parameter directions behaviour constrains.

5.1 Higher-dimensional Hessian: FlyGym 48-d

The behavioural Hessian over the 48 actuator gains of the FlyGym walking controller, computed from ten summary scalars, has effective dimension 1 (90%) / 1 (99%) with participation ratio 1.006. That figure is set by the observable count: with ten summary scalars the Hessian can have rank at most ten, and it has exactly ten non-zero eigenvalues, of which one carries essentially all of

the mass. Measured against 4200 joint-angle observables the effective dimension is 11 (90%) / 31 (99%), against a noise floor of 7 / 22 obtained from seed-to-seed variation at fixed parameters.

That figure is quoted at a step of 0.25 rather than the 0.05 used elsewhere, because this controller is proportional only above about 0.25 and the difference matters. Repeating the measurement at four steps gives a whitened response of 0.0989, 0.2083, 0.3361 and 0.6563 at steps of 0.05, 0.15, 0.25 and 0.50, so the response-to-step ratio falls from 1.98 to 1.39, 1.34 and 1.31: it varies by a factor of 1.51 over the whole range but by only 1.02 across the two valid steps, which is what proportionality looks like once the stochastic floor stops contributing. The rich effective dimension follows, 16 / 38 at the invalid step and 12 / 31, 11 / 31 and 11 / 29 as the step enters the valid region, so the value at 0.05 was inflated by roughly a third and 11 / 31 is the figure this work reports. The coarse reading is 1 / 1 at every step, so the rank argument is unaffected by the choice. At the largest step 10 of 1925 legs drove the physics divergent and two of the 48 gains had no usable leg, a reminder that the step which buys proportionality is close to the step at which this simulator stops integrating. File `AB4_flygym_delta.json`. The controlled observable-count comparison holds the parameters and step fixed and changes only how much is recorded, so it supplies the reported rank result. A separate ten-scalar calculation at the same step (δ=0.05) and with the same observables returns 3 / 6, participation ratio 2.15, and a spectrum spanning 19.5 orders of magnitude, of which 3.85 lie above the finite-difference floor; the implementation difference between the two calculations is unresolved. Under either ten-scalar estimate, recording 4200 joint angles raises the dimension by roughly an order ($1 \to 16$ or $3 \to 16$). Files `E1_flygym_rank.json`, `flygym_hessian48.json`.

5.2 Uniform channel-class scaling: stiff/sloppy decomposition

This measurement scales each named conductance by one multiplier applied uniformly to all 136 cells, so its 20 coordinates are mechanism-class gains and not the model's per-cell conductance space, which has roughly 136×16 dimensions. Three of the 20 are not channels at all but the global synaptic, gap-junction and motor-output gains. Every dimension below is therefore a statement about 20 class-gain directions, and §5.2.1 measures how much of the model's curvature those directions actually reach. Replacing the five global mechanism-class scalings with 20 named mechanisms (16 individual ion-channel conductances, the passive leak, the synaptic and gap-junction weights, and the motor output), the class-gain behavioural elasticity ($\mathrm{d}\ln\mathrm{obs}/\mathrm{d}\ln\theta$) is measured under two observable sets. Through the full motor-neuron trajectory (10,320 observables; the assay the main text reports, file `E45b_baai_named_channels.json`) the effective dimension is 3 (90%) / 6 (99%) over the sixteen named conductance-class gains, with the inward rectifier IRK (123.5), the sodium leak NCA (117.7) and the potassium conductances EGL-36 (86.0) and EGL-2 (44.6) stiffest, and SLO-1/EGL-19 (0.09), UNC-2 (0.17) and SHL-1 (0.44) sloppiest. Mechanism names are resolved through the model's own mechanism table rather than by string manipulation of the conductance name: `gbshl1` lives in mechanism `shl1`, but `gbslo1` lives in `slo1_egl19` and `slo1_unc2`, so a name-derived lookup would fail to reach them. A perturbation that fails to reach the model raises an error rather than returning a zero column. Through six summary kinematics the same decomposition gives 2 (90%) / 3 (99%) of 20 (eigenvalues $43.2, 5.84, 0.88, 0.47$, remainder ≈ 0): the coarser assay reports fewer constrained directions, as the rank analysis predicts, while the stiff/sloppy identities agree. The stiff mechanisms are the passive leak, the gap-junction and synaptic weights, the motor output, and a few potassium conductances (EGL-36, EGL-2, IRK); the sloppiest are the fast voltage-gated calcium channels EGL-19 (elasticity 0.49) and UNC-2 (0.31). The calcium-activated potassium conductances are not uniform: through the trajectory assay SLO-1/UNC-2 (6.68) and KCNL (15.1) rank mid-table while SLO-2 (0.57–0.78) and SLO-1/EGL-19 (0.09) are the sloppiest of the sixteen. Behaviour is set by wiring, passive properties and a subset of potassium channels, and

is nearly invariant to the fast calcium and calcium-activated potassium conductances, the conductances known to be compensated in small circuits (12). The infinitesimal five-mechanism curvature of §1.4 and this finite class-gain elasticity are distinct measures; the elasticity is the behaviourally meaningful one and is reported as primary. File `baai_perchannel_hessian.json`.

5.2.1 How much curvature the uniform directions reach

The class-gain coordinates are a projection of a much larger space, and the size of that projection is measured rather than assumed. Resolving one mechanism cell by cell on modWorm, the 279 per-neuron leak conductances in place of a single leak multiplier, gives effective dimension 1 (90%) / 2 (99%) with participation ratio 1.06, against 2 / 2 and 1.29 for the seven-multiplier probe: the low dimension survives a fortyfold refinement of the granularity. The uniform direction within that 279-dimensional space, however, carries only 0.295% of the curvature, and the per-neuron elasticities span 4.2×10^{-10} to 2.3×10^{-5}. Two things follow and we state both. The dimension is robust to granularity. The identity of the constrained directions is not: which cells carry the curvature cannot be recovered from a class-gain probe, and reading it requires the per-cell Hessian rather than the class-gain one. Wherever this work names a stiff or sloppy mechanism, the claim is at the level of the mechanism class. File `E37_granularity_modworm.json`.

5.3 Full-connectome manifold probe

Over the full 3076 connection weights (1992 synaptic + 1084 gap-junction) we perturb every weight individually and assemble the Gauss–Newton curvature against 9600 motor-output observables, so the rank is not limited by the measurement. The effective dimension is 8 at 90% and 47 at 99%, participation ratio 2.55: 3068 of the 3076 directions are flat. Only 4 weights produce no measurable change, so the flat directions are degeneracy rather than unconnected parameters. Per-connection elasticities span seven orders of magnitude (first percentile 6.5×10^{-6}, ninety-ninth 59.7). Splitting by connection type, chemical synapses carry 87% of the curvature and gap junctions 13%, from 1992 and 1084 connections respectively, so per connection a chemical synapse contributes 3.63 times more. The median elasticities, however, differ by only 1.20× (0.0836 synaptic against 0.0694 gap-junction), so the 87/13 split is carried by a heavy tail of a few strong synapses rather than by a uniform per-connection advantage. The degeneracy of this model therefore lies principally in the connectome, not only in the single-cell biophysics. The measurement is not a property of the published parameter point: displacing the entire weight vector to four new operating points ($\theta_i\, e^{\mathcal{U}(-\ln f, \ln f)}$ per connection, two points each at f=1.25 and f=1.5) and repeating the probe over an identical 300-connection subsample gives effective dimension 8–12 (90%) / 28–30 (99%) at every point, with the spectrum spanning 14 orders of magnitude throughout, against 9/29 at the published point on the same subsample. The stiff subspaces moreover coincide across points: the mean pairwise principal cosine is 0.89 (range 0.78–1.00), where randomly oriented subspaces of the same dimension give 0.162 ± 0.019 (File extttnull_subspace_overlap.json). The displaced-point assay uses a 1500-ms rollout and per-time whitening against the published point's 2000 ms and per-neuron whitening, so the comparison is of subspaces, not of eigenvalues. The low dimension, and the identity of the constrained directions, are properties of the region of parameter space the model occupies. Nor does the number depend on which layer of the model is read. Repeating the full 3076-connection probe through the model's own motor readout, the 96 body-wall muscle activations rather than the motor-neuron membrane potentials ($900 \times 96 = 86{,}400$ observables, again in excess of the parameters), gives effective dimension 3 at 90% and 11 at 99% with participation ratio 2.53 and a spectrum spanning 14 orders of magnitude. The readout closest to behaviour is therefore the

more concentrated of the two, and the value reported in the main text is the more conservative. Files `E42_baaiworm_connectome.json`, `E58_baai_multipoint.json`, `E60_baai_muscle_full.json`.

5.4 Behaviour-based prediction of the internal neural state

Rolling out BAAIWorm (NEURON (13)) with behaviour-recovered parameters and reading the internal membrane voltage of all 136 neurons, the generated neural dynamics are low-dimensional (PCA participation dimension $\approx$ 4.7–4.8) with mean absolute pairwise correlation 0.37–0.39, the same regime as the real whole-brain calcium recordings of Kato et al. (14) (mean absolute correlation 0.21). Behaviour-only parameter generation thus reproduces the gross statistics of the unobserved neural state, though the recovered and default arms are similar and the comparison is qualitative (membrane voltage versus calcium fluorescence; behaving versus immobilised). File `NEW_EXPERIMENTS_RESULTS.json` (field `b2_behaviour_to_neural`).

5.5 Mutant mechanism localisation

Knocking each named mechanism to 0.4$\times$ and asking whether the resulting behavioural-shift signature identifies the perturbed mechanism: the sloppy calcium channels are behaviourally near-silent (Fig. 1b) (signature norm EGL-19 0.07, UNC-2 0.12, SLO-1 and SLO-2 0.09–0.16, against 0.45 for KCNL and 0.89 for the sodium leak NCA) and fall below the 0.2 signature-norm working threshold, so this probe does not assign them, while the detectable mechanisms (synaptic, passive, motor) are mutually confusable (cosine >0.9 with several others). Behaviour alone does not uniquely localise the perturbed mechanism, which is the mechanism-level expression of the behavioural-equivalence manifold. File `baai_b3_localise.json`.

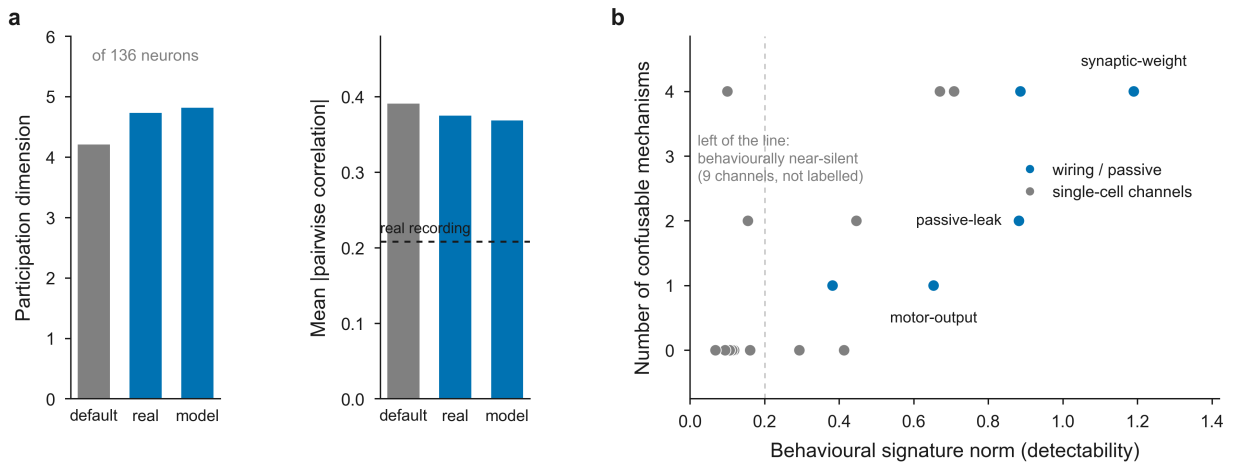

Figure 1: **Internal neural state under behaviour-recovered parameters, and the behavioural detectability of single-mechanism knockdowns. a** Mean absolute pairwise correlation of the internal membrane-voltage dynamics of all 136 BAAIWorm neurons, under default parameters and under parameters recovered from behaviour alone against a real ("real") or model-generated ("model") target. Dashed line: the same quantity computed from the real whole-brain calcium recordings of Kato et al. (14). The generated dynamics are low-dimensional (PCA participation dimension $\approx$ 4.7–4.8) with mean absolute correlation 0.37–0.39, against 0.21 in the recordings. **b** Behavioural detectability of an in-silico knockdown of each named mechanism to 0.4×, plotted against the number of other mechanisms whose behavioural signature it is confusable with. Blue: wiring and passive mechanisms; grey: single-cell channels; only the three mechanisms named in the text are labelled. Dashed line: the detectability threshold below which a knockdown is behaviourally near-silent. The fast calcium channels and the SLO-1 and SLO-2 conductances fall below that threshold (signature norm 0.07–0.16) and so cannot be localised from behaviour, whereas KCNL (0.45) and the sodium leak NCA (0.89) sit above it, while the mechanisms that are detectable are mutually confusable (cosine $>$ 0.9 with several others); behaviour alone therefore does not uniquely localise a perturbed mechanism. Comparison in **a** is qualitative: membrane voltage in a behaving model against calcium fluorescence in an immobilised animal.

5.6 External mutant-phenotype check (null)

Comparing real gene-knockout behavioural-effect magnitude (Yemini OWMD, 43 strains) against the computed stiff/sloppy class and against gene conservation: the stiff-versus-sloppy contrast is not significant (Mann–Whitney p=0.70) and the effect–conservation correlation is weak (Spearman ρ=0.33, p=0.42, n=8 overlapping genes). The comparison spans a granularity gap that we have measured rather than assumed. What the curvature ranks is a mechanism-class gain, one multiplier applied to a conductance across all cells; what a knockout removes is one gene in every cell that expresses it. Resolving a single mechanism cell by cell shows the uniform direction carrying only 0.295% of the curvature and per-cell elasticities spanning five orders of magnitude (§5.2.1), so a class-gain ranking should not be expected to predict a per-gene phenotype magnitude. Homeostatic compensation widens the same gap from the biological side, letting individual conductances vary severalfold while the functional combination is preserved (8–10, 12). We therefore read the stiff/sloppy geometry from the behavioural probes rather than from per-gene phenotype magnitudes. File `E3_mutant_phenotype.json`.

5.7 A fifth simulator: *Drosophila* connectome model

To extend the stiff/sloppy decomposition to a connectome-constrained *neural* model of the fly (the flybody/FlyGym musculoskeletal simulators have no conductance-level parameters), we add the connectome-constrained optic-lobe network of Lappalainen et al. (15), built on the adult-*Drosophila* connectome (16, 17) (`flyvis`; 45,669 neurons, 734 free parameters in three named groups: cell-type biases, cell-type time constants, synapse-type strengths). On a synthetic drifting-grating stimulus, the finite-difference Hessian over the 65 per-cell-type biases measured from 64 output channels has effective dimension 43 (90%) / 58 (99%) of 65 (participation ratio 36.3). That figure is a property of the assay, and this model is where the paper demonstrates it directly: reducing the recorded output channels from 64 to 32, 16, 8 and 6 moves the effective dimension to 24, 13, 7 and 5, close to one constrained dimension per channel, so the bias-only number tracks how much was recorded rather than the circuit. Measured across all three of the model's free parameter groups, where sampled parameters (330) outnumber neither the observables nor the readout budget per parameter, the effective dimension is 14 (90%) / 32 (99%), the figure comparable with the other systems and the one the main text reports. The identity of the stiff axes does not depend on the assay: the stiffest cell types are the elementary motion detectors T4a–d and their medulla inputs (Tm4, Tm16), the sloppiest the photoreceptors (R5, R6) and lamina wide-field cells (Lawf1/2): the motion-computing pathway is behaviourally load-bearing, the input gain is not. (We compare simulators by effective dimension rather than by a raw perturbation-sensitivity ratio: across the four motor models the latter ranges over three orders of magnitude with the observable and parameterisation.) Files `flyvis_b1.json`, `flyvis_manifold.json`.

5.8 Cross-system stiff/sloppy convergence

The crustacean pyloric circuit enters this comparison through its enumerated population rather than through a curvature, because it admits no curvature dimension (§2.1). Across the 2365 networks producing the same rhythm the least variable conductances are the graded inhibitory synapses and the PY A-current, and the most variable are the AB/PD leak and H currents and the LP CaT (the fly-visual and pyloric figure of the main text), which places coupling on the constrained side and the fast inward currents on the free side. That is the split the worm models give from the curvature for those two classes, while the leak conductances sit on opposite sides in the two systems, where the fast voltage-gated calcium channels EGL-19 and UNC-2 are sloppy (§5.2). Across four models spanning two phyla, a nematode and two arthropods, behaviour or output is set by network wiring and by slow and passive properties, and left free along the fast inward currents. The crustacean agreement compares two different measurements, a variability ranking against a curvature ranking, and is reported as a convergence of ordering and not of quantity. File `STG_b1analog.json`.

6 The identifiable subspace: identity, tiling, and cross-species universality

6.1 Identification of the manifold with the eigenworm postural space

We reduce real *C. elegans* skeletons (OpenWorm Movement Database, N2; 8,432 frames) and model body postures to tangent-angle eigenmodes by PCA. Real posture has effective dimension 3.79 (49-point skeletons) / 4.14 (17-point), with four eigenworms capturing 0.96 / 0.94 of postural variance, reproducing the classical four-eigenworm decomposition (18). Two independently built worm models produce behaviour that lives in this same real eigenworm space: the modWorm posture

subspace aligns with the real eigenworm basis at mean principal-angle cos 0.773, and BAAIWorm at cos 0.68–0.69 (67–72% of variance). Within each model the stiff mechanisms (synaptic, gap-junction, synaptic rise and fall time) displace the eigenworm trajectory more than the sloppy ones, but only modestly: the median displacement is 1.145 for the stiff mechanisms against 0.976 for the sloppy, a ratio of 1.17, with the individual mechanisms spanning 0.85 (gap) to 1.84 (fall). The ordering agrees in direction with the curvature but not in magnitude, so we read this as consistency rather than as a demonstration that sloppy mechanisms leave posture unchanged. The low-dimensional behavioural-equivalence manifold is related to the neural control of the worm's ∼4-dimensional posture. These alignments are significant against a null of random four-dimensional subspaces: drawing random k=4 subspaces in a d-dimensional posture ambient and computing the same mean principal-angle cosine, the modWorm value 0.773 has p=0.0014 and BAAIWorm 0.683 has p=0.043 even in a deliberately conservative d=10 ambient, falling to $p<10^{-4}$ at d=20 and d=48 (the full tangent-angle space); the random-subspace mean is 0.55/0.38/0.24 at d=10/20/48 (20,000 draws each). Files `EW_eigenworm.json`, `EW_baai_local.json`, `NCS_p1p5.json`.

6.2 Independence of the low dimension from observable resolution

Re-measuring the real modWorm behavioural Hessian with 7,200–14,400 posture observables (full body-segment angle and curvature time series, rather than six summary kinematics) leaves the effective dimension at 2 (90%) / 3 (99%) of 7, with spectrum $[1, 0.16, 0.013, 0.002, 0, \dots]$. The low dimensionality is a property of the behaviour, not of how coarsely it is sampled. File `RICH_observable.json`.

6.3 Complementary tiling: distinct behaviours constrain distinct axes

In BAAIWorm we contrast two behaviours, food-directed chemotaxis (food close; base net displacement 0.061) and plain locomotion in a weak gradient (food far; 0.231), and perturb either the chemosensory parameters or the motor parameters. The chemosensory parameters are stiff during chemotaxis (behavioural sensitivity 0.872) but sloppy during locomotion (0.122): a 7.15× context shift. The motor parameters show the opposite contrast (0.109 chemotaxis vs 0.357 locomotion, ratio 0.30). Each behaviour makes stiff the axis its function needs, and the two behaviours make *different* axes stiff, so the identifiable subspace across behaviours contains directions from both rather than only those of either one. File `BAAI_chemo.json`.

6.4 Growth and saturation of combined-behaviour curvature

On the real modWorm model, the effective dimension of the per-behaviour curvatures summed over a behaviour set grows with the number of behaviours and then saturates: across twelve behaviours the summed curvature reaches 4 (99%) / 2 (90%) of 7, which against the five live mechanisms leaves one residual direction that no tested behaviour constrains. Four genuinely distinct locomotor presets (gentle/harsh × anterior/posterior touch) each constrain 2–4 of 7 mechanism directions, their leading stiff subspaces only partially overlapping (mean pairwise cos 0.54, range 0.35–0.77), and their combined curvature is 5 (99%) / 2 (90%). Quantifying the complementarity directly: the four behaviour-specific stiff directions span an effective 2.98 of 4 dimensions (a single shared direction would give 1; the participation dimension of their singular values), and each behaviour contributes a mean 0.535 component orthogonal to the span of the other three (per-behaviour 0.56/0.65/0.46/0.47). To establish that this growth is not the trivial consequence of non-identical subspaces, we compare the observed geometry against two explicit nulls, with 20,000 draws each. Under a *redundant* null, in which every behaviour constrains one shared direction, the span would be 1, the novel fraction 0 and the mean pairwise cosine 1. Under a *random-orientation* null,

with directions drawn uniformly on the sphere in the same ambient space, the span is 3.50 (95% interval 3.00–3.88), the novel fraction 0.735 and the mean pairwise cosine 0.312. The observed values (span 2.978, novel fraction 0.535, cosine 0.525) lie between the two and are more aligned than random (p=0.020 for the span, p=0.082 for the novel fraction, p=0.016 for the cosine). These three tests form one family and we correct them together: under Holm's step-down procedure the cosine becomes 0.047 and the span 0.041, both still below 0.05, while the novel fraction was already not significant. The correction does not change the conclusion. The tests are over the orientation of directions given these four behaviours, and do not license an inference about behaviours in general. The behaviours therefore share substantial structure while each still contributes a distinct component. With only four behaviours, a bootstrap over behaviours is weak (resampling can draw duplicates and collapse the span), so we report the null comparison as the primary statistic and the bootstrap interval (span 1.45–2.98) only for completeness. File `null_model_tiling.json`. Two of the seven mechanism classes have no measurable effect at this step (§1.5), so the ambient dimension that these numbers should be read against is five and not seven, and the saturation value of four leaves one residual direction rather than three. The shuffled control below sets the same ceiling from the other side.

To check that this saturation is structure rather than an artefact of the estimator, each behaviour's Jacobian is scrambled while its marginal distribution is preserved, and the summed curvature recomputed. The shuffled control gives effective dimension 4.0 (90%) and participation ratio 3.54–3.86, against 2.5–3.0 and 1.75–1.86 observed: the measured spectrum is more concentrated than a scrambled one with the same marginals. The gap is one dimension, and the shuffled value of 4.0 is also the ceiling allowed by the five mechanisms with measurable effects, so this control bounds the estimator rather than establishing a large margin. File `E35_dof_vs_behaviours_modworm.json`. The degree-of-freedom curves against behaviour count, for modWorm and for FlyGym, are the source of Fig. 3c,d. Files `SAT_saturation.json`, `MB_behaviour_specificity.json`, `E35_dof_vs_behaviours_mo` `E36_dof_vs_behaviours_flygym.json`.

6.5 Cross-species: the same growth structure across two phyla, in a nematode, an insect and a crustacean

Fly (FlyGym; 48 actuator gains, four descending turn commands, joint-angle trajectories). Measured at the proportional step of 0.25 (§1.5), at one fixed controller seed (dispersion is across command subsets, not seeds), the effective dimension of the summed curvature barely moves: 12.25 ± 3.03 at one behaviour, then 12.67, 12.50 and 13.00 at two, three and four, with the 99% figure running 30.75, 31.67, 32.00, 32.00. The increase of 0.75 over the whole repertoire is a quarter of the spread across single behaviours, and the curve is not even monotone, so this is a saturated shape and not a measured gain. The same construction at the invalid step of 0.05 returns $16.5 \to 19.0$, larger in both endpoints and in the increment, which is the same inflation the step-size scan finds for the single-behaviour figure. The four commands drive one pattern generator, so they are aspects of one action rather than functionally distinct behaviours, and this is the fly counterpart of the nematode saturation in §6.4. Files `E36b_flygym_validstep.json`, `E36_dof_vs_behaviours_flygym.json`.

The same summed-curvature construction was applied to two further simulators, and they end in different places. *Mammal* (Virtual Rodent (19, 20), `dm_control` (21), MuJoCo (22); 38 actuators): the drives are deterministic open-loop sinusoids from a fixed initial state, so the differences are parameter sensitivity, and a six-behaviour repertoire drives a still-rising combined-behaviour curve ($3.67 \to 5.93 \to 7.67 \to 9.27 \to 10.5 \to 12.0$ at 99%). The absolute count is not comparable with the other species, which are parameterised differently, and the rodent uses open-loop drives rather than

trained natural behaviours, so we report the structure and not the number. *Fly larva* (larvaworld; 10 behavioural-module parameters) does not survive its own noise check and we withdraw it. Each condition had been run as a single realisation, and one of the perturbed parameters sets the rate of the intermitter, the module that generates stochastic crawl and pause bouts. Repeating the baseline 12 times at identical parameters gives a whitened run-to-run spread of 0.41 on net displacement, 0.24 on bend and heading, and 0.09–0.13 on the speed and path statistics. Against that floor the median signal-to-noise across the ten parameters is 1.21, and three of them (`turner.w_ec`, `turner.w_cc`, `interference.attenuation`) fall below their own floor: about a third of the measured difference is parameter sensitivity and the rest is realisation noise, a worse ratio than the FlyGym case of S1.5. A larva measurement that would support a dimension needs averaging over animals or a longer window, which we did not run. The claim that different actions engage different axes rests instead on the nematode contrast of S6.3, where repeat simulations are bit-identical. Files `RODENT_manifold.json`, `LARVA_noise_floor.json`, `LARVA_manifold.json` (withdrawn).

6.6 Concordance of two independently constructed worm models

Measured identically (per-mechanism Gauss–Newton elasticity, the same six observables, δ=0.25), modWorm (coupled-ODE, 7 mechanisms) and BAAIWorm (NEURON, 20 named channels) agree at the class level: the network-wiring axes (synaptic, gap-junction) carry a stiff fraction of 0.833 (modWorm) / 0.868 (BAAIWorm), the single-cell biophysical axes only 0.083 / 0.158. "Wiring stiff, single-cell sloppy" is thus a property of the worm, not of one implementation. The agreement is at the level of the class and not below it: within the stiff classes the two models order the mechanisms in opposite directions (modWorm ranks synaptic highest, BAAIWorm the passive leak; Spearman $\rho = -1$ over the n=3 shared classes). Class-level agreement with fine-level disagreement is what a class-level property predicts, and it is the same conclusion the granularity analysis of §5.2.1 reaches from the other side. File `G2_crosssim_align.json`.

6.7 Are sloppy directions coupled combinations or unused parameters?

Whether a sloppy direction is a coupled combination of parameters or simply one parameter the behaviour does not use is a question about eigenvector orientation, so it can only be asked where the eigenvectors are resolved. On BAAIWorm they are: over the sixteen conductance-class gains every eigenvalue lies at least five orders of magnitude above the double-precision floor of the leading eigenvalue (λ from 2.26×10^{4} down to 1.50×10^{-3}, floor 2.26×10^{-8}), and the simulator's own noise floor is exactly zero. Measured there, the participation of each eigenvector over the mechanism coordinates runs from 2.88 for the stiffest, loading mainly on the sodium leak NCA, and 3.25 for the third, down to 1.02 for the sloppiest, which is almost entirely SLO-1/EGL-19. The stiff directions are therefore mild combinations of two to three mechanism classes, while the sloppy directions are close to single classes. File `AB7_baai_participation.json`.

The measurement is taken on BAAIWorm rather than on modWorm because a participation describes an eigenvector's orientation, and an orientation is defined only where the eigenvalue is. The modWorm mechanism spectrum ends at -6.5×10^{-16} and -1.5×10^{-15}, numerically zero and of indeterminate sign, so the participations its trailing eigenvectors produce, stiff 3.42 and sloppy 4.68 of 7, are not fixed by the data and we do not use them. On the evidence that is well posed the claim is that stiff directions are combinations of a few mechanism classes, and not that sloppy directions are behaviour-preserving coupled combinations.

The same modWorm run carries a finite-step control worth reporting on its own terms. Moving a fixed distance along each eigenvector, at α=0.6 the stiff eigenvector changes behaviour by 0.580,

the sloppy eigenvector by 0.787 and a single parameter by 0.764: the sloppy direction moves behaviour *more* than the stiff one (ratio 0.74). A local curvature ordering therefore need not survive displacements of this size, which is the caution the multipoint probe of §5.3 addresses directly on BAAIWorm; modWorm additionally fails the differentiability gate of §1.5. File `CP_coupling.json`.

6.8 Robustness of the effective-dimension measurement

Three controls establish that the low effective dimension is neither a single-point nor a method artefact. *(i) Multiple manifold points:* at five independent points on the real modWorm manifold the per-mechanism Hessian effective dimension at 99% is $[4, 3, 1, 3, 4]$ of 7: low everywhere, not at one optimum. A denser probe at 24 random points drawn in the same prior region (σ=0.15) gives 90% effective dimension 1–3 and 99% dimension 2–5 at every point, so the low dimension is a property of the region rather than of the published parameter set; the mechanism ordering correlates across points at Spearman 0.63 ± 0.19 and the identity of the single stiffest mechanism varies (shared by 38% of point pairs), so the leading direction rotates within a stiff subspace that itself stays low-dimensional. File `E7_E11_modworm.json`. *(ii) Ruler calibration:* on constructed spectra of known rank the estimator is exact (flat-k spectra recover k=2, 5, 10; uniform-d recover d=8, 20; an end-to-end sloppy geometric spectrum recovers 2). *(iii) Statistics of probe agreement:* the Hessian-stiffness versus manifold-drift ranking agree at Spearman ρ=0.964, exact permutation p=0.0028 (5040 permutations), bootstrap 95% CI $[0.565, 1.00]$. Files `mw_D1R6.json`, `D2_ruler_control.json`, `D3_modworm_stats.json`.

6.9 Robustness of the low dimension to the spectral-mass threshold

The effective dimension is the number of eigenvalues reaching a chosen fraction of the spectral mass; the main text reports the 90% and 99% values. Recomputing across thresholds 80/90/95/99%, and alongside the threshold-free participation ratio ($\mathrm{PR} = (\sum \lambda)^2 / \sum \lambda^2$), the low-dimensional conclusion is insensitive to the choice (Table 3): every system stays low-dimensional at all thresholds, and the participation ratio, which uses no threshold, is ≤ 3.4 throughout. The rodent combined-behaviour curvature is the highest, consistent with the repertoire dependence measured above, not because of a threshold artefact.

system	dim	effective dimension at threshold 80% / 90% / 95% / 99%	PR
modWorm 7-mechanism	7	2 / 3 / 3 / 4	2.23
modWorm rich-observable	7	1 / 2 / 2 / 3	1.35
BAAIWorm per-channel	20	1 / 2 / 2 / 3	1.34
larva combined (withdrawn, S6.5)	10	— / — / — / —	2.07
rodent combined (6 behaviours)	38	4 / 6 / 7 / ≥ 10	3.35

Table 3: **Effective dimension across spectral-mass thresholds and by the threshold-free participation ratio.** Effective dimension at a threshold q is the smallest number of eigenvalues whose cumulative magnitude reaches a fraction q of the total spectral mass; the participation ratio $(\sum_i \lambda_i)^2 / \sum_i \lambda_i^2$ uses no threshold at all. Every system remains low-dimensional at all four thresholds and by the threshold-free measure, so the conclusion does not depend on where the cut is placed; the rodent combined-behaviour curvature is the largest, as the behavioural-repertoire result predicts. The rodent row uses the stored top-ten spectrum of the combined curvature, so its 99% entry is a lower bound; the full-spectrum combined curvature has effective dimension 12 (§6.5). Curvature is the positive semi-definite Gauss–Newton form in every row. The first two rows are two observable sets for the same modWorm measurement: the 7 mechanisms reported in the main text measured from six summary kinematics and from 7,200 posture observables. A 30-dimensional per-neuron modWorm measurement is not listed: it was evaluated as a raw finite-difference Hessian at a zero-loss anchor, carries two negative eigenvalues and a condition number of 1.2×10^{18}, and so is not a Gauss–Newton form and not comparable with the other rows.

6.10 A different action reaches directions one action cannot

Section S6.4 shows that further aspects of one action converge on the subspace that action already determines. Here we ask the complementary question directly, by measuring the same BAAIWorm circuit under two different actions and combining the curvature. Two parameterisations are used: the 20 named conductance, synaptic, gap-junction and passive scales of S5.2, and the leak conductance of each of 14 neuron classes. In both, food-directed chemotaxis alone renders four directions behaviourally distinguishable, plain locomotion in a weak gradient alone renders four, and the combined curvature $H_{\text{chemo}} + H_{\text{loco}}$ renders six. Two four-dimensional subspaces whose combined span is six-dimensional share two directions and each reaches two the other cannot; this is a rank statement and requires no threshold and no null. It stands in contrast to the within-action result, where twelve behaviours of the same kind saturate at four directions and never exceed it (S6.4).

The size of that gain must be reported alongside its existence, because the two answer differently. Measured by the effective dimension that is this work's primary quantity, the combined curvature does not move: chemotaxis alone gives 2 (90%) / 4 (99%), locomotion alone 1 / 2, and the combined curvature 2 / 4 in the 20-parameter probe and 1 / 4 in the 14-class probe, where the 90% figure is *lower* than chemotaxis alone. The reason is visible in the spectrum: the combined eigenvalues are 346.7, 39.0, 6.00, 4.12, 1.179, 1.038, so the two directions that raise the rank from four to six carry 0.30% and 0.26% of the spectral mass, against 1.51% and 1.03% for the third and fourth. A different action therefore reaches directions the first cannot, and those directions enter the spectrum alongside the ones already counted rather than at its floor. The effective dimension at 90% does not move because the leading eigenvalue carries 87% on its own, which is a statement about how concentrated the spectrum is and not about whether the new directions are there. That is the same conclusion the within-action result reaches from the other side, and it is why the abstract says the subspaces overlap more than they complement rather than that a repertoire multiplies

what is identified.

The identity of the constrained parameters changes with the action in the same way. Ranking neuron classes by their weight in the leading curvature direction, chemotaxis gives AWC, DB, VA and AIZ, whereas locomotion gives DD, VA, VD and DA: of four, one is shared, the ventral-cord A-type motor neurons, and the chemosensory and first-layer interneuron classes are exchanged for the inhibitory D-type motor neurons. The leading stiff subspaces of the two actions overlap at cos 0.46 (20-parameter) and cos 0.53 (14-class), against the chance level $\sqrt{k/n}$ for k-dimensional subspaces in n dimensions of 0.39 and 0.46 respectively; unlike the within-action comparison of S6.4, which exceeds its null at p=0.016, these sit at the chance level. No draw-based null is computed for these two runs, so the rank statement, which needs none, is the primary evidence and the subspace cosines are reported as description rather than as a test. Both measurements use six world-frame path observables, fewer than the parameter count, so this section reports a rank and not an effective dimension: the first of the three gates, that observables outnumber parameters, is a requirement for reporting a dimension, whereas a rank is admissible whenever it falls below the observable cap. The observed rank of four lies below the single-action cap of six, and the combined rank of six lies below the cap of twelve that two actions afford. The combined value of six coincides numerically with the single-action cap, so the margin rests on the two-action cap rather than on a comfortable gap, and a wider observable set would settle it (§6.5 does this for the rodent). Files `B6_union_crossbehaviour.json`, `B6b_cellclass_union.json`.

7 Ground-truth verification at the cell scale

7.1 Panel, protocol and estimator

Eleven published cell models, five neuronal (Hodgkin–Huxley (23), Hodgkin–Huxley with an A-type current (24), Wang–Buzsáki (25), Morris–Lecar (26), FitzHugh–Nagumo (27)) and six intracellular oscillators (Kholodenko MAPK (28), repressilator (29), Goodwin (30), Selkov (31), Brusselator (32), Goldbeter mitotic (33)), are implemented directly from their published equations and verified against the published dynamical structure: the two models carrying the main claims agree with independent second transcriptions to within 0.1%, and each of the others reproduces a structural signature of its source (the Brusselator Hopf threshold at $B = 1 + A^2$, the Goodwin requirement of a Hill coefficient above eight, the one-third-period phase symmetry of the repressilator, the Sel'kov and FitzHugh–Nagumo oscillation windows, the Goldbeter period of 25 minutes, fast spiking at 60 Hz in the Wang–Buzs'aki neuron). They are measured with the identical Gauss–Newton estimator used throughout, in log-parameter space, with the observable count pinned to a fixed multiple of the parameter count so no model is measured from fewer observables than parameters. Recovery follows the organism-scale generation protocol (34): the target is the published parameter set and the search starts from a random point $\theta_i^{\text{init}} = \theta_i^* e^{\eta_i}$, $\eta_i \sim \mathcal{U}(-\ln f, \ln f)$, drawn independently of the truth, at $f = 1.5$, 3 (the organism-scale regime) and 10. The loss is built from phase-invariant features of the settled trace (amplitude quantiles, spectral magnitude, autocorrelation), computed at full resolution and only then reduced, which keeps observables in excess of parameters while removing the phase sensitivity that makes pointwise trajectory losses untraversable for oscillators: pointwise losses converged on 11% of the same starts at f=3 and on none at f=10, against 47% and 36% for the feature loss. The discarded transient fraction is chosen per model by stationarity of the settled trace (quarter-range agreement), which matters only for the A-current neuron, whose default window straddles a silent stretch. A fit counts as converged when the behavioural distance falls below 0.02; only converged fits contribute parameter errors. Files `cell_panel3_result.json`, `cell_panel3b_result.json`.

7.2 Both ends of the geometry, against truth

The 2–4-parameter neuronal models are recovered exactly: Hodgkin–Huxley converges on 23/24 random starts at median parameter error 6.2×10^{-5} in the primary panel and on 72/72 at 1.9×10^{-5} in the HH-family rerun, Morris–Lecar on 17/24 at 3×10^{-7}, the Wang–Buzsáki and A-current variants on 39/72 and 15/72 at $<3\times10^{-4}$ (their lower convergence rates reflect landscape difficulty, not degeneracy: every converged fit lands on the truth). The larger biochemical cascades are degenerate in the direct sense: MAPK matches the behaviour to 10^{-3} while its parameters end at median error 0.40, farther from the truth than the random start (0.31), with individual rate constants compensating each other by factors up to 8 in the displayed example; Goodwin and the Goldbeter oscillator end at 0.10 and 0.17. The MAPK trace in Fig. 5 has parameter error 43% and behavioural distance 3%; an equal-length off-manifold displacement has period ratio 0.38. Across the eleven models the converged parameter error rises with parameter count (Spearman ρ=0.70, p=0.017). The estimator itself is validated in both directions: it reports the canonical sloppy spectrum on the model family where sloppiness was described, and exact identifiability where recovery proves it (Fig. 5a,b). Files `cell_panel3_result.json`, `cell_panel3b_result.json`, `mapk_showcase.json`.

7.3 Effective dimension is not recoverable dimension

On Hodgkin–Huxley, a grid of three parameter counts (3, 5, 8: conductances; + leak reversal and capacitance; + gating-rate scalings) × four observable sets (1 spike count, 7 F–I and spike-shape features, 200- and 1000-point voltage traces) gives a 90% effective dimension of 1 in all twelve cells, while the achieved recovery error spans 4% (three conductances against the F–I assay; g_{Na}, g_{K} within 4–5%) to 38% (eight parameters against the same assay, where the leak and capacitance are absorbed by the gating-rate scalings) to 56% (the cell-scale ground-truth figure of the main textc). The effective dimension measures how concentrated the curvature is, not how many parameters recovery can pin down; the trajectory-observable cells additionally fail for a different reason (the behaviour is never matched), which is a property of the loss landscape, not of identifiability. This grid also reconciles the exact single-neuron recoveries here with the conductance degeneracy reported for such models (7, 8): the two literatures occupy different corners of the same table, few parameters against a multi-condition assay versus many parameters against a single trace. Files `cell_ladder_result.json`, `cell_diag_result.json`.

7.4 Network size does not dilute the constrained fraction

Coupling N identical Hodgkin–Huxley cells through fixed sparse excitatory synapses, with three conductances free per cell and the observable budget pinned per parameter, the constrained fraction barely moves as N grows from 2 to 64 (Fig. 2) (6 to 192 parameters): 0.33 to 0.30 at a 10× observable budget, with the effective dimension rising linearly at close to one constrained direction per cell throughout, while the eigenvalue spectrum widens from 2.1 to 4.3 decades. Size alone therefore deepens the degeneracy but does not dilute the constrained fraction, so the 0.26% constrained fraction of the connectome-scale model (§5.3) is not a generic consequence of its parameter count. File `cell_scale_result.json`.

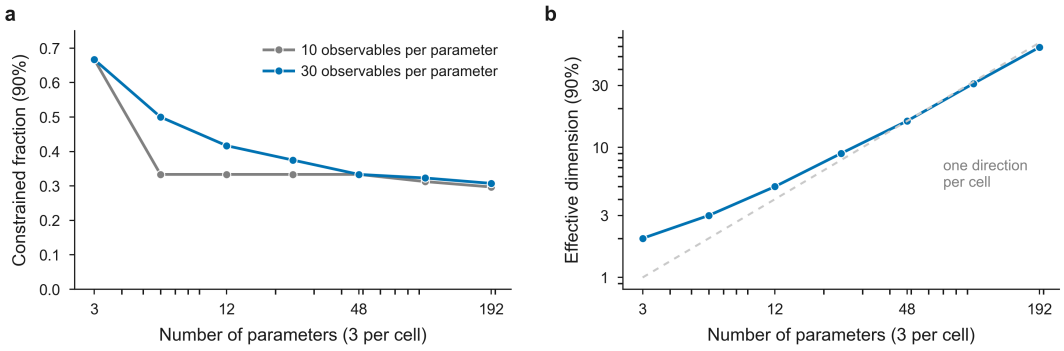

Figure 2: **Network size deepens the degeneracy but does not dilute the constrained fraction.** **a** Constrained fraction at 90% of the curvature as identical Hodgkin–Huxley cells are coupled into a fixed sparse network, three free conductances per cell, at two observable budgets. From 6 to 192 parameters the fraction moves from 0.33 to 0.30 at the 10× budget and from 0.50 to 0.31 at the 30× budget, two orders of magnitude away from the 0.26% of the connectome-scale model. **b** The effective dimension itself rises close to linearly, at about one constrained direction per cell (dashed reference), so the flat fraction is not a ceiling artefact; the widening eigenvalue spectrum (2.1 to 4.3 decades) deepens the degeneracy without diluting the fraction.

8 Reproducibility

All external datasets are public: Prinz–Marder STG model population (7); Allen Cell Types; WormBase ParaSite / Ensembl Metazoa; OpenWorm Movement Database. All simulators are the published releases cited in the main text, run at the settings given above. Each result section names its saved result file; every reported quantity and figure is regenerated from those files by the accompanying reproduction code.

References

[1] Balázs Szigeti, Padraig Gleeson, Michael Vella, et al. OpenWorm: An open-science approach to modeling Caenorhabditis elegans. *Frontiers in Computational Neuroscience*, 8:137, 2014.

[2] Padraig Gleeson, David Lung, Radu Grober, et al. c302: a multiscale framework for modelling the nervous system of *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society B*, 373(1758):20170379, 2018.

[3] George Papamakarios, Theo Pavlakou, and Iain Murray. Masked autoregressive flow for density estimation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

[4] Alvaro Tejero-Cantero, Jan Boelts, Michael Deistler, Jan-Matthis Lueckmann, Conor Durkan, Pedro J Gonçalves, David S Greenberg, and Jakob H Macke. sbi: A toolkit for simulation-based inference. *Journal of Open Source Software*, 5(52):2505, 2020.

[5] Sean Talts, Michael Betancourt, Daniel Simpson, Aki Vehtari, and Andrew Gelman. Validating Bayesian inference algorithms with simulation-based calibration. *arXiv preprint arXiv:1804.06788*, 2018.

[6] Joeri Hermans, Arnaud Delaunoy, François Rozet, Antoine Wehenkel, Volodimir Begy, and Gilles Louppe. A trust crisis in simulation-based inference? Your posterior approximations can be unfaithful. In *Transactions on Machine Learning Research*, 2022.

[7] Astrid A Prinz, Dirk Bucher, and Eve Marder. Similar network activity from disparate circuit parameters. *Nature Neuroscience*, 7(12):1345–1352, 2004.

[8] Eve Marder and Adam L Taylor. Multiple models to capture the variability in biological neurons and networks. *Nature Neuroscience*, 14(2):133–138, 2011. doi: 10.1038/nn.2735.

[9] Mark S. Goldman, Jorge Golowasch, Eve Marder, and L. F. Abbott. Global structure, robustness, and modulation of neuronal models. *Journal of Neuroscience*, 21:5229–5238, 2001.

[10] Adam L Taylor, Jean-Marc Goaillard, Davi Bhatt, and Eve Marder. How multiple conductances determine electrophysiological properties in a multicompartment model. *Journal of Neuroscience*, 29(17):5573–5586, 2009.

[11] Leandro M Alonso and Eve Marder. Visualization of currents in neural models with similar behavior and different conductance densities. *eLife*, 8:e42722, 2019.

[12] Eve Marder and Jean-Marc Goaillard. Variability, compensation and homeostasis in neuron and network function. *Nature Reviews Neuroscience*, 7:563–574, 2006.

[13] Michael L. Hines and Nicholas T. Carnevale. The NEURON simulation environment. *Neural Computation*, 9(6):1179–1209, 1997.

[14] Saul Kato, Harris S Kaplan, Tina Schrödel, et al. Global brain dynamics embed the motor command sequence of *Caenorhabditis elegans*. *Cell*, 163(3):656–669, 2015.

[15] Janne K. Lappalainen, Fabian D. Tschopp, Sridhama Prakhya, Mason McGill, Aljoscha Nern, Kazunori Shinomiya, Shin-ya Takemura, Eyal Gruntman, Jakob H. Macke, and Srinivas C. Turaga. Connectome-constrained networks predict neural activity across the fly visual system. *Nature*, 634:1132–1140, 2024.

[16] Louis K. Scheffer, C. Shan Xu, Michal Januszewski, et al. A connectome and analysis of the adult Drosophila central brain. *eLife*, 9:e57443, 2020.

[17] Sven Dorkenwald, Arie Matsliah, Amy R. Sterling, et al. Neuronal wiring diagram of an adult brain. *Nature*, 634:124–138, 2024.

[18] Greg J Stephens, Bethany Johnson-Kerner, William Bialek, and William S Ryu. Dimensionality and dynamics in the behavior of *C. elegans*. *PLoS Computational Biology*, 4(4):e1000028, 2008.

[19] Josh Merel, Diego Aldarondo, Jesse Marshall, Yuval Tassa, Greg Wayne, and Bence Ölveczky. Deep neuroethology of a virtual rodent. *arXiv preprint arXiv:1911.09451*, 2019. International Conference on Learning Representations (ICLR) 2020.

[20] Diego Aldarondo, Josh Merel, Jesse D. Marshall, Leonard Hasenclever, Ugne Klibaite, Amanda Gellis, Yuval Tassa, Greg Wayne, Matthew Botvinick, and Bence P. Ölveczky. A virtual rodent predicts the structure of neural activity across behaviours. *Nature*, 632(8025):594–602, 2024.

[21] Saran Tunyasuvunakool, Alistair Muldal, Yotam Doron, Siqi Liu, Steven Bohez, Josh Merel, Tom Erez, Timothy Lillicrap, Nicolas Heess, and Yuval Tassa. dm_control: software and tasks for continuous control. *Software Impacts*, 6:100022, 2020.

[22] Emanuel Todorov, Tom Erez, and Yuval Tassa. MuJoCo: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5026–5033, 2012.

[23] Alan L. Hodgkin and Andrew F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4): 500–544, 1952.

[24] John A. Connor and Charles F. Stevens. Prediction of repetitive firing behaviour from voltage clamp data on an isolated neurone soma. *The Journal of Physiology*, 213:31–53, 1971.

[25] Xiao-Jing Wang and György Buzsáki. Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *Journal of Neuroscience*, 16:6402–6413, 1996.

[26] Catherine Morris and Harold Lecar. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal*, 35:193–213, 1981.

[27] Richard FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1:445–466, 1961.

[28] Boris N. Kholodenko. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *European Journal of Biochemistry*, 267(6): 1583–1588, 2000. doi: 10.1046/j.1432-1327.2000.01197.x.

[29] Michael B. Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403:335–338, 2000.

[30] Brian C. Goodwin. Oscillatory behavior in enzymatic control processes. *Advances in Enzyme Regulation*, 3:425–437, 1965.

[31] Evgeni E. Sel'kov. Self-oscillations in glycolysis. 1. a simple kinetic model. *European Journal of Biochemistry*, 4:79–86, 1968.

[32] Ilya Prigogine and René Lefever. Symmetry breaking instabilities in dissipative systems. II. *The Journal of Chemical Physics*, 48:1695–1700, 1968.

[33] Albert Goldbeter. A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. *Proceedings of the National Academy of Sciences*, 88:9107–9111, 1991.

[34] Chenxi He. Generating behaviour-equivalent parameters for whole-organism simulators from tracking video. Preprint, 2026. URL `https://doi.org/10.21203/rs.3.rs-10486658/v1`. Preprint.