Global modules robustly emerge from local interactions and smooth gradients

Mikail Khona[a],[1,2,*] Sarthak Chandra[b],[2,*] and Ila Fiete[c2]

[1] *Physics, MIT*

[2] *Brain and Cognitive Sciences, and McGovern Institute, MIT*

Modular structure and function are ubiquitous in biology, from the organization of animal bodies and brains to the scale of ecosystems. However, the mechanisms of modularity emergence remain unclear. Here we introduce the principle of *peak selection*, a process by which purely local interactions and smooth gradients can result in global modular organization. It can lead to the self-organization of discontinuous module boundaries from a smooth global gradient, unifying the positional hypothesis and the Turing pattern formation hypothesis for morphogenesis. Applied to the brain's grid cell networks, peak selection results in the spontaneous emergence of functionally distinct modules with discretely spaced spatial periods. Applied to ecological systems, a generalization of the process results in discrete systems-level niches. The dynamics exhibits emergent self-scaling to variations in system size and "topological robustness" [1] that renders module emergence and module properties insensitive to most parameters. Further, peak selection confers robustness within modules. It ameliorates the fine-tuning requirement of continuous attractor dynamics even in single grid cell modules. It makes a detail-independent prediction that grid module period ratios should approximate adjacent integer ratios, furnishing the most accurate match to data to date. Additional testable predictions promise to bridge physiology, connectomics, and transcriptomics. In sum, our results indicate that local interactions combined with low-information global gradients can drive robust global module emergence.

INTRODUCTION

Modular structures are ubiquitous in natural systems, from body structures to circuits in the brain, and from ecological niches to human communities. This is probably so because they are robust to localized perturbations [2, 3], can be faster to adapt if the world requires sparse or modular changes [4], or can permit flexible, high-capacity computation through compositionality [5–10]. In these senses, modularity is the crux of biological organization.

The prevalence of modularity raises critical questions about its evolutionary, developmental, or ecological origins: Modular solutions to a given problem form a vanishingly small subset of all possible solutions, thus from an evolutionary perspective it is unclear how these solutions are found and selected. From the perspective of development, the question is how modular structures form, and whether module features such as size, number, and boundary locations need to be genetically instructed or spontaneously emerge through unfolding physical processes such as symmetry breaking. From an ecological perspective, the question is how modular structures emerge among interacting species even without a shared task or function that the system is seeking to optimize.

One hypothesis for the developmental emergence of structure, which is our primary focus here, is the positional information hypothesis espoused by Lewis Wolpert (Fig. 1a): Gene expression generates spatial morphogen concentration gradients, and different downstream genes become activated in spatially localized regions by thresholding the morphogen concentration [11, 12]. In line with this hypothesis, body segmentation in Drosophila [13] is controlled by spatial bands of expression of a family of genes (the gap genes) that are activated by different concentrations of maternally deposited Bicoid RNA (Fig. 1b-c). Modular gap gene expression precedes and directs modular body segmentation. A distinct hypothesis by Alan Turing is the idea that structured patterns can

[a] Email: mkhona@nvidia.com

[b] Email: sarthakc@mit.edu

[c] Email: fiete@mit.edu

[*] These authors contributed equally to this work.

spontaneously emerge from local competitive interactions, minimizing or removing the need for genetic instruction [12, 14] (Fig. 1d). Supporting this hypothesis is evidence that digit formation in hand morphogenesis is under control of spontaneous pattern formation [15]. Another clear example of Turing-like pattern formation comes from the grid cell system in the medial entorhinal cortex (MEC) of mammalian brains, Fig. 1e-f. MEC neurons fire in triangular grid-like activity patterns as a function of explored space when animals navigate [16]. Underlying these spatially periodic responses are intrinsic periodic activity patterns in the cortex [17, 18]. Extensive evidence [19–23] links these activity patterns with continuous attractor network (CAN) models based on Turing-like pattern formation [24–26]. However, it remains unclear whether and how these CAN models for single grid cell modules translate to the formation of the multiple discrete modules of grid cells with distinct periods found in the brain [27].

The positional and Turing processes for structure emergence have distinct properties and predictions [12, 28–30]: The positional mechanism is susceptible to noise in copy number [12, 31–33] and requires separate downstream genetic cascades to specify how and where each structure forms. Its prediction is that modular structure or function are driven by modularity in gene expression, which runs counter to at least some experimental studies that find that modular function can exist without evidence of modularity in the underlying cellular or molecular properties [31, 34–47]. Positional processes possess scale invariance: the formed structures are self-scaling to occupy a constant fraction of the system as the system size is varied. The pattern forming mechanism typically only produces structure of a single scale, given by the width of the local lateral interactions. Thus, these models do not explain emergence of structures of multiple scales. However, pattern formation is robust to noise.

We hypothesize that pattern forming and positional mechanisms can be unified into a combined process that exhibits the strengths of both, allowing modularity to emerge via self-organization from local interactions without the need for modularity in gene expression, and such that the resulting process is scale-invariant (Fig. 1g). We show that such a process can explain the emergence of multi-scale structure in the form of multiple grid cell modules in mammalian cortex (Fig. 1h-i), and is robust to most parametric variation and noise. The model produces strikingly accurate predictions about the sequence of successive spatial period ratios in grid cells, improving substantially on existing models. The process exhibits a "topological robustness" property that substantially eases the usual fine-tuning requirements of continuous attractor models of grid cells [24]. It also generates numerous predictions for future physiology, transcriptomics, and connectomics experiments in the system.

Analyzing the underlying dynamical mechanisms of the process allows us to extract a general principle for global module emergence with smooth global gradients and local variations, which we call the *peak selection* principle. We then apply the peak selection principle to very different problems, showing the emergence of modular multi-species niches in an interacting ecological system and in a non-interacting system with two distinct external drives.

GENERALIZATION OF SINGLE-MODULE CONTINUOUS ATTRACTOR MODELS

Grid cells in the mammalian medial entorhinal cortex (MEC) of mammals exhibit spatially periodic response patterns as animals explore open spaces [16]. Before considering mechanisms for the formation of multiple discrete and functionally independent grid modules along the long (dorsoventral or DV) axis of MEC, Fig. 1h-i, , we extend the theory of single grid cell modules.

The properties of grid cells within a module are consistent with continuous attractor neural network (CAN) models [19–26]. CAN models involve a linear, Turing-like instability driven by strong competitive local interactions between neurons, leading to spontaneous pattern formation and the verified prediction that the states of the circuit of thousands of cells lie on a two-dimensional set with the topology of a torus, Fig. 1e-f [22, 24].

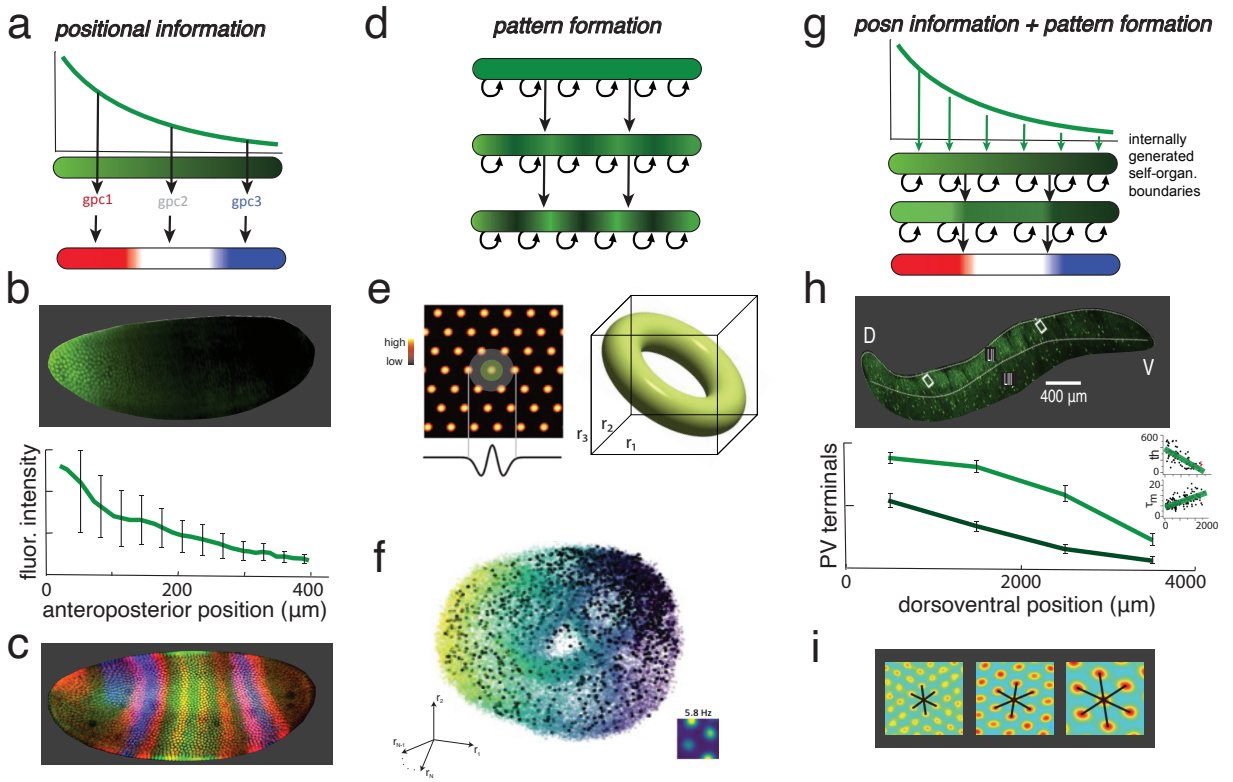


FIG. 1. **Positional versus pattern-forming mechanisms for structure formation and our hypothesis.** (a) The positional hypothesis: global gradients are thresholded by different downstream gene expression cascades to generate structure [11]. (b) Fluorescence image of maternally deposited protein *bcd* RNA (based on maternal bicoid RNA deposition) early in development of the Drosophila embryo [48] sets up a polarity gradient. (c) A downstream gene-protein expression cascade, including gap and pair-rule genes, sets up body segment-defining bands by thresholding the bicoid gradient (immunofluorescence image adapted from [49]; segmentation figure adapted from [13].) (d) Spontaneous self-organized structure emergence (pattern formation) through competitive lateral interactions [14]. (e-f) The continuous attractor neural network (CAN) model for single grid cell modules [24] is based on Turing instability based on local interactions, and its predictions are consistent with the experimental data [19–21, 23]. These include the prediction of a continuous set of stable states with toroidal geometry across waking and sleep, and its recent confirmation [22]. (g) Our hypothesis: Positional and pattern forming mechanisms can interact to lead to structure emergence that exhibits the strengths of both mechanisms. Module boundaries are determined by emergent pattern formation not modular gene expression, but an overall gradient permits the system to exhibit self-scaling with system size. (h-i) The long dorsoventral (DV) axis of medial entorhinal cortex (MEC; image of layers II and III) [35] exhibits smooth-seeming gradients in multiple cellular properties, while along the same axis, grid cells are organized into discrete modules with discontinuous jumps in their spatial periods (adapted from [27]).

Existing CAN models are based on two interaction profiles: a center excitation-surround inhibition (Mexican hat) shape [24, 50] (including an inhibition-only version [24]), or a uniform local inhibition shape (which we term a "Lincoln hat" [51]). Given the differences in these profiles, we sought to explore the implications for experiment: whether there was something unique about these profiles and to map the space of interactions that could generate grid cell-like responses. We derived a set of simple conditions on the neural input-output transfer function and the local interaction kernel W that we hypothesize would be sufficient for grid-like patterning (SI Sec. A): the neural transfer function is not an odd function; and the interaction kernel is such that it is radially symmetric; its integral is negative (inhibition dominated: $\int W(x, x')dx' < 0$); and that

it is sufficiently strong. These conditions define an infinitely large set of distinct interaction profiles. We sampled randomly from this set and numerically implemented the network dynamics with rectified linear neurons, finding that all sampled profiles produce grid-like patterning (Fig. 2a). These results significantly expand the generality of CAN models for single grid cell modules. The experimental implication is that the interaction profiles of past models [24, 25, 50, 51] may not predict the profiles unearthed in future connectomic studies, but that the uncovered profile might conform to the more-general conditions defined here.

LOCAL INTERACTIONS FOR MULTI-SCALE GLOBAL MODULARITY EMERGENCE

Grid cells exhibit discrete jumps in period and independent functionality along the long (DV) axis of entorhinal cortex [27]. Physiological experiments reveal that several biophysical quantities are graded along this axis, including in the neural time-constant, the strength of synaptic interactions, the width of connectivity, and other properties [35, 41, 45, 52], Fig. 1h. We replaced the translation-invariant interaction kernel $W(|\mathbf{x}-\mathbf{x}'|)$ in CAN grid cell models by a kernel $W^g(|\mathbf{x}-\mathbf{x}'|;\sigma(n_{DV}))$ with a slowly graded width $\sigma(n_{DV})$ along the DV axis of the model neural sheet (SI Fig.11a; n_{DV} refers to DV location and $\mathbf{x}$ refers to the general 2-dimensional position on the neural strip). This variation of the local interaction width in the CAN model still produced hexagonally arranged activity bumps, with a growing period (SI Fig. 11b). However, the variation in pattern period was smooth, without emergent modularization (SI Fig. 11c).

Because global modularity and local patterning involve two spatial scales, we reasoned that two scales of lateral interaction might be necessary to generate both. At the same time, just as local interactions can lead to globally periodic structure, we hypothesized that the addition of a second *local* interaction might be sufficient to induce global modular structure. However, the sum of two local kernels is simply another local kernel and from the previous section we should expect no difference in results, unless the second interaction is distinct from the first in some way beyond a mere shape difference. We therefore considered two types of local interaction with the following key difference: the first is graded across the DV axis while the second remains fixed. The combined interaction is:

$$W(\mathbf{\Delta x};\sigma(n_{DV})) = W^g(\mathbf{\Delta x};\sigma(n_{DV})) + W^f(\mathbf{\Delta x}). \tag{1}$$

Both interactions are local and much smaller in width than the DV length (L) of the cortical sheet. We assume that the fixed interaction width (d) is larger than the largest width (σ_{max}) of the graded interaction. Remarkably, the addition of such a fixed-scale interaction leads the network to spontaneously decompose into a few discrete modules, with coherent periodic activity patterns locally and discontinuous jumps in period globally, Fig. 2b-d.

As before, there is broad latitude in the shapes of the interaction kernels W^g and W^f, so long as one is graded and the other is slightly wider but fixed in width along the DV axis, Fig. 2f-h (setting the fixed interaction to be narrower than the graded interaction resulted in a grid pattern whose period varied in a sawtooth-like instead of step-like modular fashion, SI Fig. 24; this result was also predicted by our analytic theory described below). This combination of a graded-width and a fixed-scale interaction also produces robust and spontaneous decomposition of dynamics into discrete modules in 2D network models, Fig. 2e,i. (See SI Sec. D 8 for additional model results for 2D networks.) The formation of K modules involves approximately K times the number of neurons as in single-module continuous attractor models (SI Sec. D 9 a, SI Fig. 22). We have used kernel width as a general proxy for some gradient in effective interaction along the DV axis. All these results generalize if we instead consider gradients in two other biophysical properties: the cellular time-constant and the synaptic strength instead of connectivity width (SI Sec. E) [35, 41, 45, 52–55].

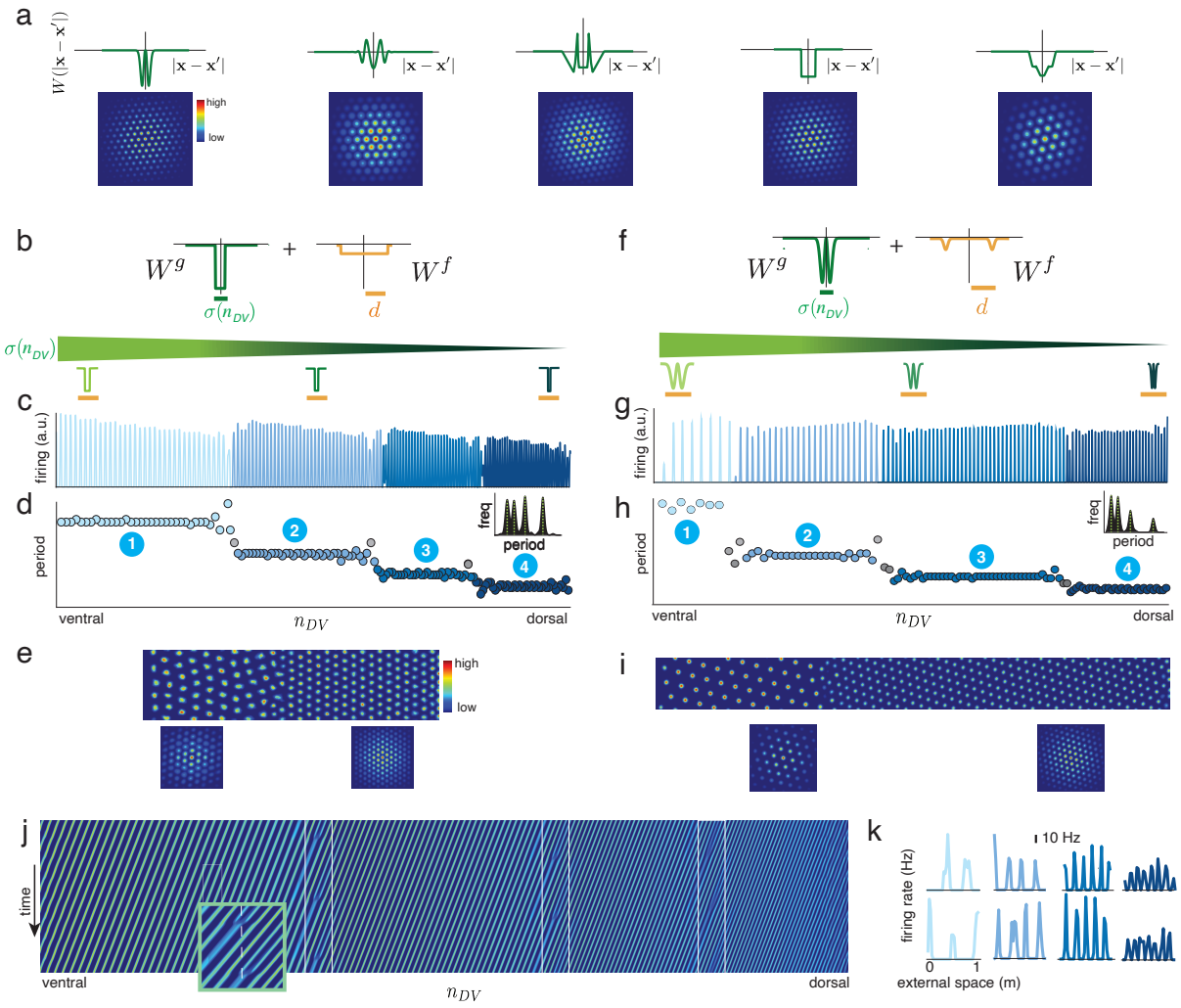


FIG. 2. **Two *local* interactions, with graded and fixed widths, respectively, lead to global module emergence.** (a) Generalization of CAN grid cell models: 5 examples from an infinite set of distinct local interaction kernel shapes that can lead to grid-like patterning. (b) Combining two local interactions, one whose width ($\sigma(n_{DV})$) scales smoothly along the DV axis ($W^g_{n_{DV}}$, green) and a broader but still-local one whose width (d) remains fixed along the neural strip. Interaction widths indicated below the gradient are drawn to scale relative to the activity shown in (c). (c-e) The two interactions from (b) lead to spontaneous emergence of modules with distinct periods in 1-dimensional (c) neural strip, with extracted periods shown in (d). The same kernels applied to a 2-dimensional neural sheet (e), with the 2d autocorrelation function of the local (single-module) patterns in the neural sheet (bottom). (f-i) Same as (c-e), but for a different pair of interaction kernels $W^g_{n_{DV}}, W^f$ with distinct gradient shape $\sigma(n_{DV})$ and endpoints $(\sigma_{\min}, \sigma_{\max})$ from (c-e). (j) The response of the 1-dimensional neural strip shown over time when the network is driven by a smoothly graded velocity input, white lines highlighting the temporal evolution of dynamics at the module boundary (inset: magnification of the first boundary). (k) The independent velocity-driven pattern dynamics in each module result in regular periodic spatial tuning curves (shown are 2 cells per module). *See Methods for parameter and simulation details.*

Strikingly, the modules that emerge, Fig. 2c-e, g-i, are much larger than the widths of either local interaction, Fig. 2b, f (interaction widths shown to scale).

Formed modules are functionally independent

We probed whether the emergent modules are functionally independent units. In single-module CAN models, velocity inputs drive the pattern to flow at a direction and speed proportional to the velocity. The network is thus an integrator of the velocity signal[24, 56–59]. For the formed modules to independently perform velocity integration, their patterns must flow independently and their phases pass each other discontinuously at module boundaries, even though connectivity is equally continuous within and across module boundaries. It seems intuitively unclear whether this could hold. However, when we drove all modules with a common velocity input, we found that the patterns flowed independently. The positions of the boundaries between modules remained fixed and sharp, and the phases on either side of a boundary updated independently so that there were spatiotemporally discontinuous dislocations in phase across the boundary, Fig. 2j. This dynamics results in veridical and independent velocity integration within and across modules, so that all cells (even those close to the module boundaries) have periodic spatial tuning curves, Fig. 2k. We conjecture that this independence arises from the integer-based selection of number of modules and module boundaries. At each moment, as the velocity input drives updates in the pattern phase across the system, the network performs module formation which forces a break across modules at the same location. This discrete jump at module boundaries preserves the structure and dynamics within each module, allowing each module to function independently.

The independence of phase updating does not imply an independence in the angular tuning of modules: if all modules receive a common rotated velocity input, their spatial tuning curves will rotate. The only way to induce independent rotation of tuning in the grid modules is to provide separately rotated velocity inputs to each.

ANALYTICAL THEORY OF MODULARIZATON: PEAK SELECTION, TOPOLOGICAL ROBUSTNESS, AND SELF-SCALING

The generality and robustness with which discrete modules emerge from the combination of a fixed-scale and a graded-scale local interaction suggests a general principle at work. Starting from an initial condition of uniform activity, the network exhibits nearly immediate (within 1-2 biophysical time-constants τ) signs of modularization, Fig. 3a. Modularization begins before most neurons have crossed their nonlinear thresholds, and unfolds concurrently with local periodic patterning (Fig. 3a,e). The system also exhibits localized eigenvectors (SI sec H; similar Anderson localization in condensed matter physics [60]). These phenomena suggest that patterning and modularity might both be explained by a unified linear instability-based theory. We derive such a theory, summarizing it below with details in SI (Sec.B). Besides establishing how, why, and when modularity emerges, the theory accurately predicts the discrete pattern periods of all modules, the number and sizes of modules, and the locations of module boundaries (explored below).

We considered how small perturbations evolve from an initial state $s_0(n_{DV})$. In the local neighborhood of each DV location on the neural sheet (the neighborhood is assumed to be larger than the interaction width $\sigma(n_{DV}), d$ but much smaller than the full sheet width L), the local interaction $W(\mathbf{\Delta x}; n_{DV})$ changes only a little, and we can approximate it to be spatially uniform to solve the dynamics using Fourier modes (SI sectionD 8). As in the analytical theory of the earlier section (Generalization of single-module continuous attractor models), the network forms a patterned state within each neighborhood, with the *spatially varying* inverse periods as a function of DV location n_{DV}:

$$\frac{1}{\lambda(n_{DV})} = \arg\max_k \{\tilde{W}^f(k) + \tilde{W}^g(k; n_{DV})\}, \tag{2}$$

where $\tilde{W}$ is the Fourier transform of W.

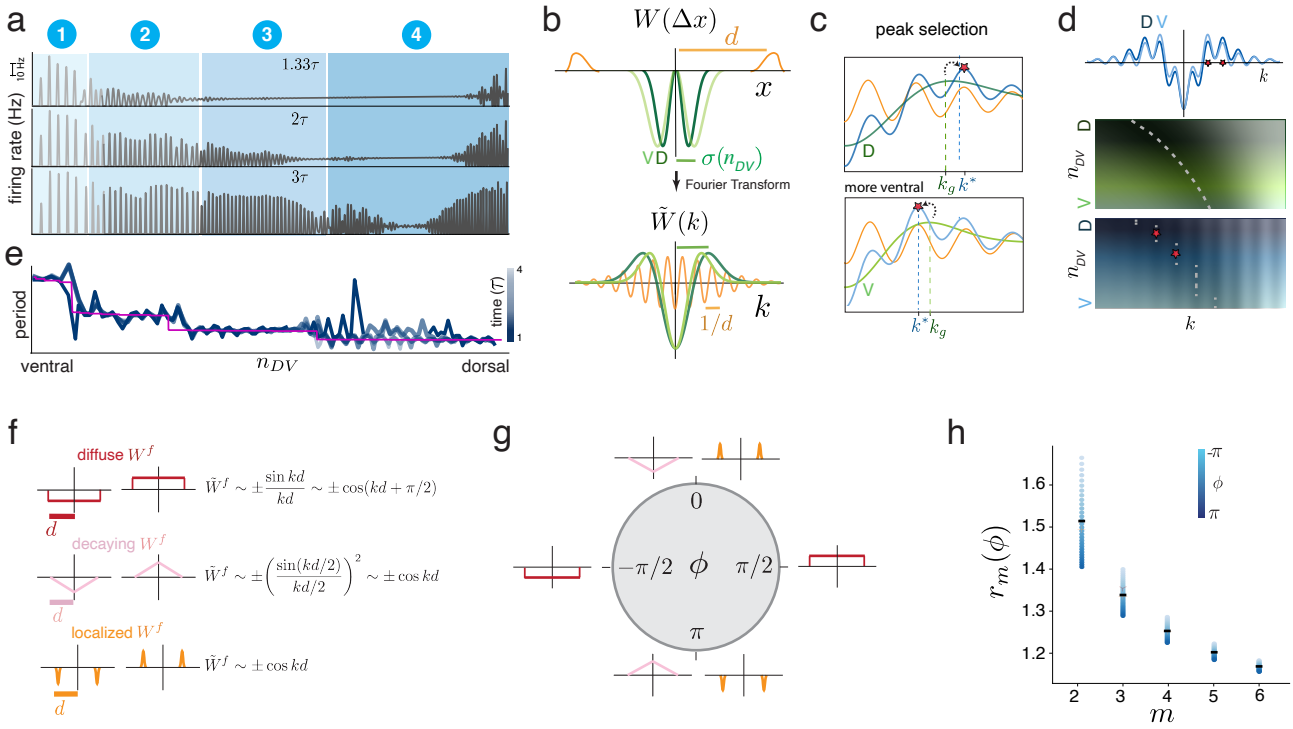


FIG. 3. **Theory of module emergence: multi-scale linear instability and topological peak selection** (a) Snapshots of population activity within a few neural time-constants (τ) of initializing the dynamics at a uniform state. Modules appear in situ at the same time as local patterning, before most neurons have hit their nonlinear thresholds, defined as the point when the presynaptic inputs to a neuron reach the non-linear region of the input-output relationship. (b) Top: Schematics of fixed-width (orange) and graded (green) interaction kernels. Kernels from two different DV locations (designated D for dorsal and V for ventral). Bottom: their Fourier transforms. (c) Peak selection process: The global maxima in Fourier space (blue) are based on combining the graded interaction (green) with the fixed interaction (orange). As the green peak slides across, the global maximum (marked by red star) jumps abruptly from the position of one orange peak to the next. (d) Top: Summed Fourier transform of the two local interactions (darker (lighter) blue: more dorsal (ventral)). Middle: The location of the maximum of the graded interaction varies smoothly as a function of DV location. Bottom: the maximum of the summed interaction jumps discontinuously (bottom). (e) Dark to light blue curves: activity pattern periods from (a) for early to late times after initialization. Module boundaries and periods remain unchanged from the earliest time-points. Pink: theoretical prediction of periods and module boundaries from Eq. 2. (f) Left: Example simple fixed-scale interaction profiles that produce modularization: profiles can be roughly categorized as diffuse, decaying, or localized. Right: the dominant terms in their Fourier transforms. (g) The Fourier phases of the interactions in (a). (h) Theoretically predicted sequence of period ratios for any value of ϕ (blue circles), for module numbers 2-6. Black markers denote averages for each value of m. Any dependence on ϕ and thus the shape of the fixed-scale interaction is weaker for higher module numbers (smaller period/dorsal modules). *See Methods for parameter and simulation details for (a) and (e).*

We can understand how the peridocity varies spatially as follows. Suppose the fixed interaction kernel W^f is *simple* in the sense that there is a single dominant length-scale d (details in SI Sec. D); then $\tilde{W}^f(k) \sim \cos(kd - \phi)$), which has closely spaced peaks (every $\sim 1/d$ in k-space; Fig. 3b, bottom, orange) and a phase ϕ. These local maxima remain the same for each DV location in the neural strip because W^f is not graded.

By contrast, $\tilde{W}^g_{n_{DV}}$ exhibits a broad Fourier peak (of scale $\sim 1/\sigma(n_{DV}) \gg 1/d$). The width and location of this peak for different DV locations contract smoothly (as $\sigma(n_{DV})$ increases; Fig. 3b, bottom, green). The W^g interaction drives spatial patterning and its graded variation is ultimately responsible for changes in period through the smooth variation of the peak of $\tilde{W}^g(k)$. However,

the narrow peaks of $\tilde{W}^f(k)$ determine the specific values of the maxima of the sum of $\tilde{W}^f$ and $\tilde{W}^g$, while the smoothly moving peak of $\tilde{W}^g(k; n_{DV})$ performs "peak selection" on these possibilities to define the global maximum (Fig. 3c and SI Movie 1). As the broad peak of $\tilde{W}^g(k; n_{DV})$ smoothly sweeps through the set of narrow local maxima of $\tilde{W}^f$ (as n_{DV} is varied), the global maximum remains at one of the narrow maxima, then abruptly and discontinuously jumps to the next peak of $\tilde{W}^f$, generating a constant period within modules and abrupt changes in period between modules, Fig. 3c-d. The spatial periods are determined by the maxima of $\tilde{W}^f$, which occur at

$$\lambda_m^{-1} = k_m^* \approx \left\{ \frac{2\pi m + \phi}{d} \middle| m \in \mathbb{Z}^+ \right\}. \tag{3}$$

In other words, the periods are determined by the width d of the fixed-scale interaction W^f, with different module periods given by this scale divided by integers m. Module periods are independent of the length L of the neural strip.

Permitted values of the integers m are given by which local maxima of $\tilde{W}^f$ fall within a range $[\eta/\sigma_{\max}, \eta/\sigma_{\min}]$, determined by the range of scales of the graded-width interaction (η is a fixed proportionality constant; SI Sec. D 7 for details). From this, we can determine the number of allowed modules, which is the set of integers m that fit in the following interval:

$$\frac{\eta d/\sigma_{\max} - \phi}{2\pi} \leq m \leq \frac{\eta d/\sigma_{\min} - \phi}{2\pi} \tag{4}$$

The phase $\phi \in [-\pi, \pi]$ is a constant that reflects the only influence of the shape of W^f on the formed modules (Fig. 3f-g): If W^f is uniformly diffuse across its width d, then $\phi \approx \pm\pi/2$ (sign determined by whether it is excitatory or inhibitory). If W^f is locally concentrated around d, then ϕ is close to 0 or π (for excitatory or inhibitory interactions, respectively); a decaying W^f also leads to ϕ close to 0 or π. Intermediate values of ϕ can be obtained by interpolating between these interaction shapes (See Fig. 12 for several examples).

The analytical expression for module periods (Eq. 2 evaluated on the Fourier transform of W) exactly predicts the values from numerical simulation (Figs. 3e, 4b, SI Fig. 12). The even simpler analytical expression for period in Eq. 3 with ϕ computed from W^f and without free parameters, also exactly predicts module periods from numerical simulation across diverse lateral interaction shapes (Figs. 3a, 4b, SI Fig. 12).

Period ratio prediction and parameter invariance

The (inverse) module period expression of Eq. 3 supplies a quantitative prediction about adjacent module period ratios. Period ratios have been characterized experimentally [27] and are the subject of several theoretical models [61, 62], however these findings consider all adjacent module period ratios to have a single value. By contrast, our model's period ratio predictions vary with module: the period ratio of the mth module to the $m+1$th module is:

$$\frac{\lambda_m}{\lambda_{m+1}} \equiv r(m, \phi) = \frac{(m + 1 + \phi/2\pi)}{(m + \phi/2\pi)}. \tag{5}$$

The module period ratio prediction is strikingly and completely independent of any scale: neither $d, \sigma_{n_{DV}}$, nor L, even though the module periods themselves scale with d. In other words, module period ratios in our model are completely independent of the widths of the interaction kernels, or the size of the cortical strip. The ratios are also independent of the functional form of gradient, $\sigma_{n_{DV}}$, by which the graded interaction varies. The only parameter dependence in the period ratios is through the scalar phase ϕ. For $\phi = 0$, The ratios of adjacent modules are simply successive integer ratios, with the integer indexing the module number (Fig.3h).

The extreme invariance of predicted period ratios to almost all parameters is due to robustness arising from a "topologically protected" [1] process (SI Sec.D 7): possible solutions to the dynamics of patterning form a topologically discrete set, with periods given by the successive integers which are generated by the peak selection process. Through all continuous variations in the parameters of the system, it must settle into one of these solutions. Thus, the modularization process and resulting predictions about module period ratios is a topologically protected robust process immune to nearly all parameteric variation.

We implemented and theoretically analyzed models where the graded quantity is either the neural time constant or the strength of recurrent connectivity (rather than the width of neural connectivity), SI Sec. E. All the above results remain qualitatively unchanged, and the period ratio predictions remain qualitatively and quantitatively unchanged.

Spontaneous self-scaling of modules

Above, we noted that module periods are independent of the cortical strip length L and period ratios are independent of both L and the widths of the interaction kernels.

Interestingly and surprisingly, the number of formed modules is also independent of system size, if the minimum and maximum widths ($\sigma(0) = \sigma_{\max}$, $\sigma(L) = \sigma_{\min}$) of the graded interaction remain the same as the network size L is varied. We can see this by setting $\sigma_{\min}, \sigma_{\max}$ to be constant (independent of L) in Eq. 4, and deriving the number of modules N_{mod} to be:

$$N_{mod} = \left\lfloor \frac{\eta d}{2\pi\sigma_{\min}} - \frac{\phi}{2\pi} \right\rfloor - \left\lceil \frac{\eta d}{2\pi\sigma_{\max}} - \frac{\phi}{2\pi} \right\rceil$$

where $\lfloor\ \rfloor$, $\lceil\ \rceil$ indicate the floor and ceiling operations, respectively. The number of modules is determined by the interplay (difference) between the width ratios of the two local kernels, without depending on their specific widths or even the values of their ratios. This expression is independent of L.

As a result, a prediction of our analytical results is that the module formation process self-scales to the system size, such that each module must grow in size (but with unchanged periodicity within the module) as the system size (cortical strip length) is increased. Indeed, in numerical simulations where we hold $\sigma_{\min}, \sigma_{\max}$ fixed as we scale the cortical strip by several factors, each module scales in size with the overall network size (Fig.4a), and the number of modules and module periods remain the same (Fig.4a-b). Thus, if the neural sheet is large, the module sizes can be orders of magnitude larger than any of the lateral interaction scales, $\sigma_{\min}, \sigma_{\max}, d$, resolving the mystery of what sets the scale of individual modules, why they are unrelated to the local interaction scales, and why they are global in size, Fig. 2b-d,f-h. This scale-free nature of the emergence of modules is an entirely novel feature for pattern-forming models.

The number of formed modules is also insensitive to the shape of the gradation $\sigma(n_{DV})$ of the graded parameter, depending as above only on its maximum and minimum values (Fig.4c-d). The only effect of changes in the shape of $\sigma(n_{DV})$ is on the relative sizes and boundary locations of the different modules, rather than their number. (We can make a more approximate but still fairly accurate prediction for where module boundaries will form, Fig. 3e, pink curves and Fig. 4d; see SI D 7 for details.)

Finally, even if we smoothly vary one of the extremal values of $\sigma(n_{DV})$, the number of modules remains fixed until the change becomes large enough to accommodate one additional or one less module. At that point, we will obtain different numbers of modules (Fig.4c-d). However, we again we see remarkable robustness in the modules that do form: as the number of formed modules is reduced by increasing $\sigma_{\min}$, the formed module periods are identical to the periods of some of the modules in the bigger set (Fig.4c-d).

In sum, the dynamics of module formation exhibit an emergent invariance or self-scaling property with brain size, automatically adjusting to the size of the substrate. Variations in brain size across individuals and species can, via this property, be accommodated to generate a fixed number of modules with a fixed set of periods without tuning any biological parameters. Positional information models (like the French flag model, Fig. 1a [11]), which apply fixed thresholds to a global gradient, would also be expected to yield self-scaling, however here the modules and boundaries are *self-generated* through recurrent interactions, and do not require externally imposed thresholds or genetic control. An additional consequence is that increasing the number of modules involves simply changing the value of one endpoint ($\sigma_{\max}$ or $\sigma_{\min}$) rather than creating new gene expression cascades for each added module (Fig. 4e).

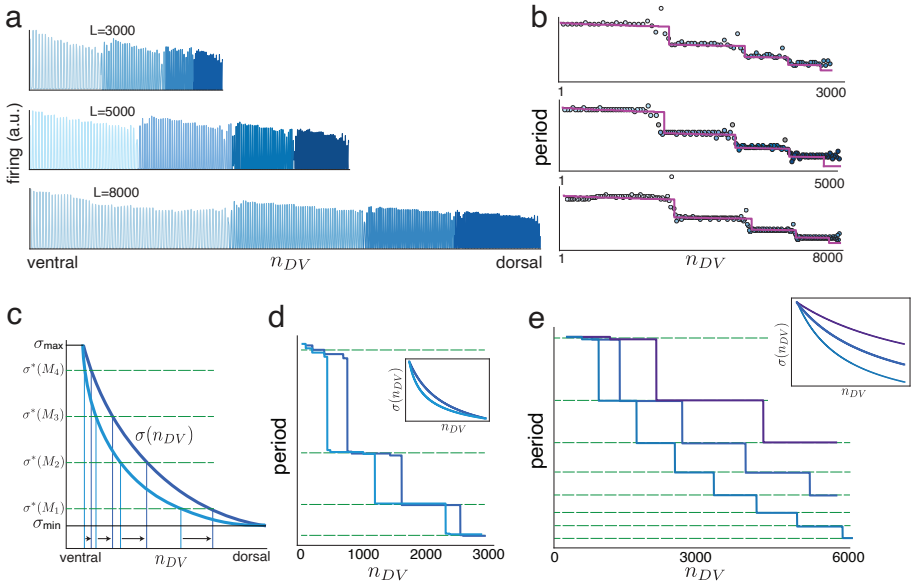


FIG. 4. **Emergent self-scaling of modules with system size.** (a) Increasing the size of the neural sheet while holding constant the minimum and maximum graded interaction widths and the fixed interaction width, the within-module periods remain the same but module sizes expand so that the system has the same number of modules regardless of system size. (b) Extracted periods from results in (a). The neural axis is scaled (normalized) by network size to compare relative module sizes; the period axis is the same across plots (preserved periods in each module). Pink: analytical predictions from Eq. 2. (c) Different functions (shapes) for the monotonically graded interaction width $\sigma(n_{DV})$ are predicted theoretically to result in the same number of modules if the minimum and maximum values of the width $(\sigma_{\min}, \sigma_{\max})$ remain unchanged. Shape changes only affect the detailed positions of module boundaries. (d) k^* calculated from numerical Fourier transform of interaction matrix with two different gradient shapes, holding $\sigma_{\min}$ and $\sigma_{\max}$ fixed. Module number and periods remain unchanged, while boundaries shift. (inset) The shapes of the gradient in the width of the primary pattern-forming interaction for the two choices of gradient shapes. Green dashed lines are scales corresponding to each local maxima of the secondary interaction. (e) k^* calculated from numerical Fourier transform of interaction matrix with three significantly different values of $\sigma_{\min}$ while holding the spatial extent of the system fixed. The number of formed modules changes from 3 to 5 to 8, while the periods of the first few modules (that are common across all three simulations) remain unchanged. (inset) The shapes of the gradient $\sigma(x)$ in the primary pattern-forming interaction for the three choices of gradient shapes. *See Methods and SI SecD 7 for parameter and simulation details.*

Neural data matches detailed predicted sequence of period ratios

Our prediction of period ratios is module-specific and depends on the phase ϕ, Fig. 5a. First, we compare our prediction of period ratios with existing data is to average the predicted values

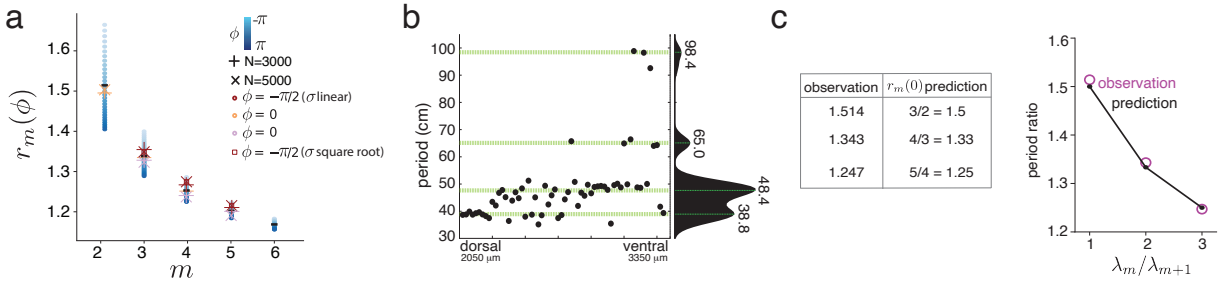


FIG. 5. **Comparison of precise period ratio predictions with data** (a) Period ratio predictions from 3h together with numerical simulation (other symbols) of neural circuit models with the set of fixed-scale interaction profiles shown in 3f, same color code. Numerical simulations with all combinations of network size and weight profiles are shown. σ(linear) and σ (square-root) denote two different profiles of $\sigma(x)$ (refer to Eq.11 for the functional form of the profiles). (b) Observed periods of grid cells from multiple modules [27] (c) Successive period ratios computed from the observation (left column), and predicted period ratios for $\phi = 0$ (middle column). Ratios match predicted values with $R^2 = 0.999$ (right column). *See Methods for parameter and simulation details.*

across 4 modules and over all phases ϕ (SI Sec. D 6). This yields a predicted value of 1.37, in good agreement with experimental results of an average ratio of ~ 1.42 across animals, as reported in [27]. For the animal shown in Fig. 5b, the average module period ratio is 1.368.

Next, we compare our more fine-grained successive period ratio predictions with published per-module period values, selecting the best-fit value for ϕ. Our prediction with $\phi = 0$ matches the sequence of observed period ratios from [27] strikingly well, Fig.5c, as well as other datasets in which multiple grid periods ratios are available from single individuals (SI Sec. E 1).

PEAK SELECTION ENHANCES ROBUSTNESS WITHIN AND ACROSS MODULES

Robustness within individual continuous attractor networks

Above, we found that peak selection-based multi-module emergence leads is robust and invariant to variations in parameters, function shapes, and the form of the global gradients. Here we further report that two-scale interactions and the peak selection principle make the dynamics within single grid modules resistant to several forms of weight heterogeneity and activity perturbation. The requirement that continuous attractor models possess a high level of weight homogeneity (i.e., perfect translation invariance) to generate a continuum of fixed points [23, 24, 63, 64] is a well-known Achilles heel that has led to debate about whether these models accurately describe the biological circuits. This susceptibility to noise is one of the fundamental open problems for most continuous attractor models[23].

We simulated multi-bump CAN models [24] for one grid module (i.e. no gradient in the pattern forming interaction), and added inhomogeneities of two types to the weights: noise in the shared radial structure of all interaction weights and i.i.d. noise in each weight (SI Fig.20 and SI sec D 9 for details and visualizations). These inhomogeneities were sufficient to nearly destroy pattern formation (Fig. 6a, left). However, inclusion of a second wider-scale local interaction (also subject to the same amplitude of added noise) results in robust and homogeneous pattern formation (Fig. 6a, right). We quantified this effect, finding stronger regularity in the formed pattern despite weight heterogeneity with the broader interaction than without, Fig. 6b. The susceptibility to heterogeneity and the gains from a broader interaction scale held whether the model was simulated with periodic or aperiodic boundary conditions (one version of CAN grid cell models consists of a single activity bump [50, 65, 66]; this version, which is likely inconsistent with the observation

that single modules exhibit a multi-bump pattern on the cortical sheet [67], would not benefit from the addition of a broader interaction profile). The formed attractor states are continuous enough to path integrate with fidelity despite weight inhomogeneity (SI Fig. 21). The same enhancement of within-module period regularity, despite the addition of significant weight noise, holds for simulations with graded weights that result in the formation of multiple modules, Fig. 6c.

Conceptually, the broader secondary interaction likely enhances pattern regularity because the narrower peaks it induces in Fourier space imposes discrete period selection, which greatly reduces the pattern variance-driving effects of noise, Fig. 6c: the wider local interaction "focuses" the dynamics to a narrower region than specified by the narrower local interaction alone (Fig. 6d).

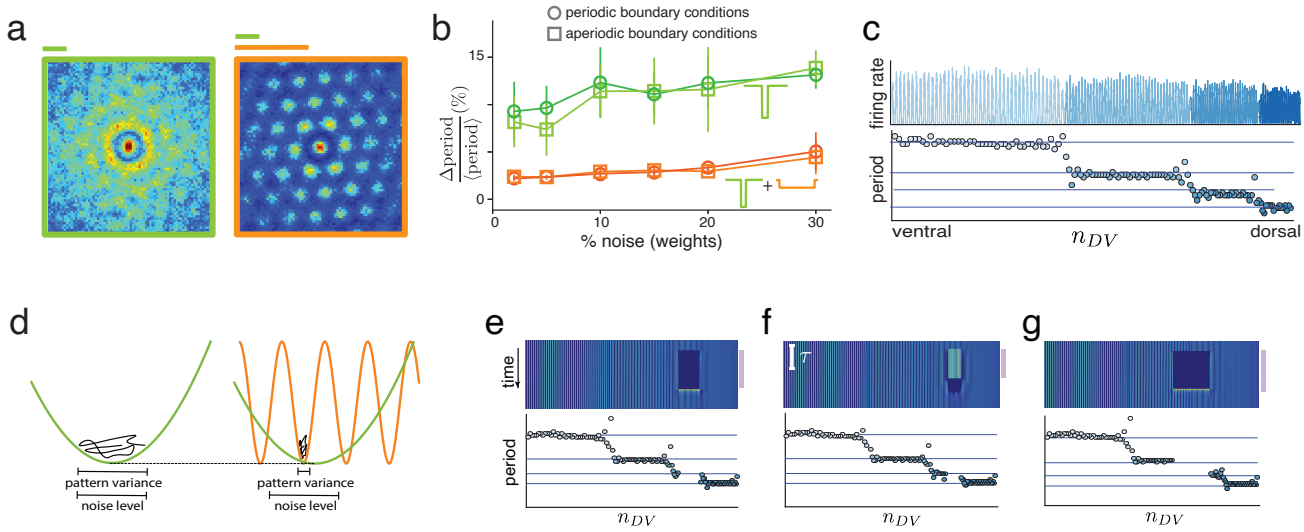


FIG. 6. **Enhanced robustness to weight heterogeneity, noise, and activity perturbation by peak selection.** (a) Left: Weight heterogeneity (here, radial asymmetry and i.i.d. noise) quickly destroys discernable pattern structure in multi-bump continuous attractor models [24]). (Simulation of a 100x100 neuron network with kernel W_g as in Fig. 2d-e; scale shown by green bar.) Right: Addition of a secondary wider local interaction (scale shown by orange bar), with noise in both sets of weights, rescues patterning. (b) Variability in 1-dimensional patterning versus the magnitude of added noise in the weights, for single-scale weights (green), and for networks with two local interaction scales (orange) with both periodic (circles) and aperiodic (squares) boundary conditions. Pattern variation is the ratio of the standard deviation to the mean of the pattern period. (c) The same as (a), showing regularity in period despite the addition of noise in a 1-dimensional setting. (d) The mechanism for enhanced within-module robustness: the broader local interaction scale enforces a narrower set of solutions in the energy landscape than possible with the pattern-forming interaction alone. (e-g) Inter-module dynamical independence: (e) An entire module is transiently silenced for 50 ms; (f) a large fraction of a module is externally driven by large-amplitude fixed, random, independent perturbations; (g) a continguous region that spans two modules is transiently silenced. In all cases, the perturbation remains local so neighboring regions and modules are unaffected, and the perturbed module recovers within one neural time-constant after removal of the perturbation. *See Methods for parameter and simulation details.*

Across-module robustness to large-scale activity perturbations

We next probed whether perturbing activity within entire modules or large across-module regions of the network will affect patterning in the rest of the modules, as might be expected given that the lateral interaction weights span across modules.

We entirely silenced activity in one module (mimicking optogenetic inactivation), to find that the other modules, their periods, and even adjacent module boundary locations remained stable, Fig.6e. When we force a subset of cells in one module to persistently fire at randomly selected rates between 0 and 20Hz, patterned activity is also disrupted in immediately adjacent regions of the module, Fig.6f, but the patterns are immediately restored upon removal of the forcing drive. During this perturbation in one module, no other modules are disrupted. If a region partly spanning two modules is silenced, the dynamics and periods in spared parts of the two modules remain unchanged, and the boundary re-emerges at its pre-perturbation position after removal of the suppressive input (Fig.6g). In all cases, pre-perturbation states are restored within one cellular time-constant ($\approx \tau$). These findings contrast with existing models of module formation in which modules interact in a stacked architecture [68]: these models exhibit cascading dependencies between modules, so that perturbation of one module will have propagating effects in all downstream modules.

GENERALIZED ENERGY LANDSCAPE VIEW OF MODULE EMERGENCE

We hypothesize that the principle of peak selection could, applied in domains other than the Fourier space of a translationally invariant spatial system, supply a general mechanism for modularity emergence without periodic pattern formation. The theory can be generalized in two steps. First, by translating the linear dynamics of Fourier modes into nonlinear dynamics on an general energy landscape, and next translating the Fourier peaks and troughs into multiple rugged local optima in the energy landscape (SI Sec. F).

Consider an arbitrary state variable x whose dynamics flow downhill on a generalized energy landscape (Lyapunov function) $L(x;\theta)$, such that $dx/dt = -\nabla_x L(x;\theta)$, where θ is some parameter. The solution to the dynamics is some fixed point state $\bar{x}$. Suppose the energy function is given by a sum of two terms (Fig. 7a):

$$L(x;\theta) = (1-\alpha) f_1(x) + \alpha f_2(|x - g(\theta)|) \tag{6}$$

where the function $f_1(x)$ is rugged with multiple similar-depth minima in the state space x and a function $f_2(x;\theta)$ that has a single broad minimum at x^* such that $x^* = g(\theta)$, where g is some monotonic function. The location of the broad minimum in the state space x smoothly moves as the parameter θ is smoothly varied (Fig. 7b). The dynamics of x can be viewed as regularized optimization on a rugged loss landscape, with a regularizer $f_2(|x-g(\theta)|)$ that acts as a prior biasing solutions $\bar{x}$ toward $x^* = g(\theta)$. In our numerical simulations, the landscape is first governed by the broad quadratic term, then sculpted by the rugged landscape (with α gradually decreasing with time starting from $\alpha = 1$, SI Sec. G). Smoothly varying the parameter θ results in a set of modular solutions $\bar{x}$ (Fig. 7c).

This version of topological peak selection generalizes the notion beyond linear instability, Fourier modes, spatial embedding, or periodic solutions. This general setting — in which a fixed function generates multiple local minima in some space and another "selecting" function that generates one broad minimum in that space, with the position of the minimum smoothly moving as another parameter is varied — can now be applied to generate modularity across diverse settings. Next, we explore modularity emergence with this principle at the ecological scale, for the emergence of spatial ecological niches and coordinated spawning in corals.

SELF-ORGANIZATION OF ECOLOGICAL NICHES AND COORDINATED CORAL SPAWNING THROUGH PEAK SELECTION

Consider a set of (N) species interacting cooperatively or competitively. Species are indexed by i, and s_i denotes their population levels. We follow the literature in the field to model the

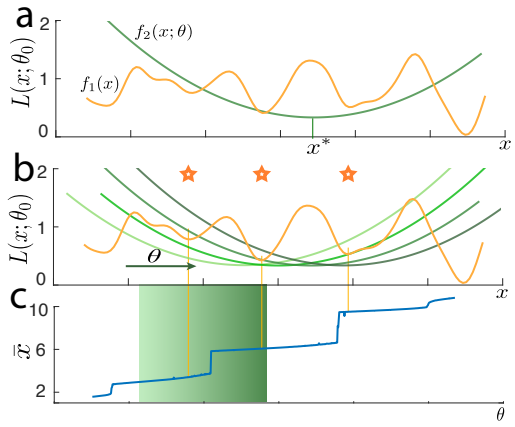


FIG. 7. **Generalized peak-selection mechanism leads to modularity emergence.** (a) Energy landscape (Lyapunov function) for dynamics of the abstract state variable x consisting of a rugged multi-minimum function and a smooth, broad single-minimum function with minimum located at x^*. (b) As a parameter θ is varied, x^* varies as $g(\theta)$, where g is some monotonic function. (c) The resulting fixed points $\bar{x}$, as a function of the smoothly varied θ, form sets with a constant value, followed by an abrupt jump to a new set of values, and so on in a series of discrete steps, defining a set of discrete modules. (See SI Sec. G for simulation details.)

dynamics of the population at a given location x and time t by a Hopfield network, defining the interaction between two co-localized species i and j by an interaction strength W_{ij}, which can take positive (cooperative) or negative (competitive) values (Fig. 8a-b) [70–74]:

$$s(i,x;t+1) = H\left[-\delta_i s(i,x;t) + \sum_{j,x'} W(i,j)K(x,x')s(j,x';t) + \mathbf{b}_i \cdot \mathbf{r}^g(\theta)\right] \tag{7}$$

where δ_i is the death rate of species i; $H[.]$ is a rectifying nonlinearity (populations levels are non-negative); and $K(x,x') \equiv K(x-x')$ is a spatial interaction kernel (two species interact only if their ranges have some overlap, separated by no more than the width of the kernel, which is assumed to be much smaller than L, the size of the environment). Each species has different resource needs, specified by a resource feature vector ($\mathbf{b}_i$ of dimension M (sampled i.i.d. for each species). The environment supplies resources $\mathbf{r}^g$. Ignoring the spatial aspect, the interactions W induce a rugged landscape that would result in a set of attractors in the state space, Fig. 8c, with each attractor representing a potential stable configuration relative population levels of the N species. (See Methods for details.) Next, we construct a parametric resource gradient by linearly interpolating between two random vectors $\mathbf{r}_1^g, \mathbf{r}_2^g$ drawn from $\{-1,1\}^M$, Fig. 8c (colored line). Different values of the input along this parametric gradient "tilt" the landscape: when we slice the landscape along this gradient direction and consider one value of the input or another, the minima remain the same but their relative amplitudes vary, Fig. 8d.

Finally, we assume that the parametric resource gradient is spatially organized (in other words, the parameter θ for the gradient $\mathbf{r}^g(\theta) = (1-\theta)\mathbf{r}_1^g + \theta\ \mathbf{r}_2^g$ is some monotonic function of space, $\theta = g(x)$). Such variations – such as in sunlight, temparature, humidity, precipitation – are common features of ecosystems and are believed to shape population dynamics and nice formation [75–84].

We initialize this ecological model at a spatially homogeneous state, with continuous resource gradients, and examine the structure of the population vector and its normalized correlation matrix $C(x,x')$, (Fig. 8e-f). The population state self-organizes, through the process of peak selection, into a steady state with modular niches (Fig.8e-h): clusters of species form stable groupings of relative population abundances over local regions of the environment, with sharp and highly correlated

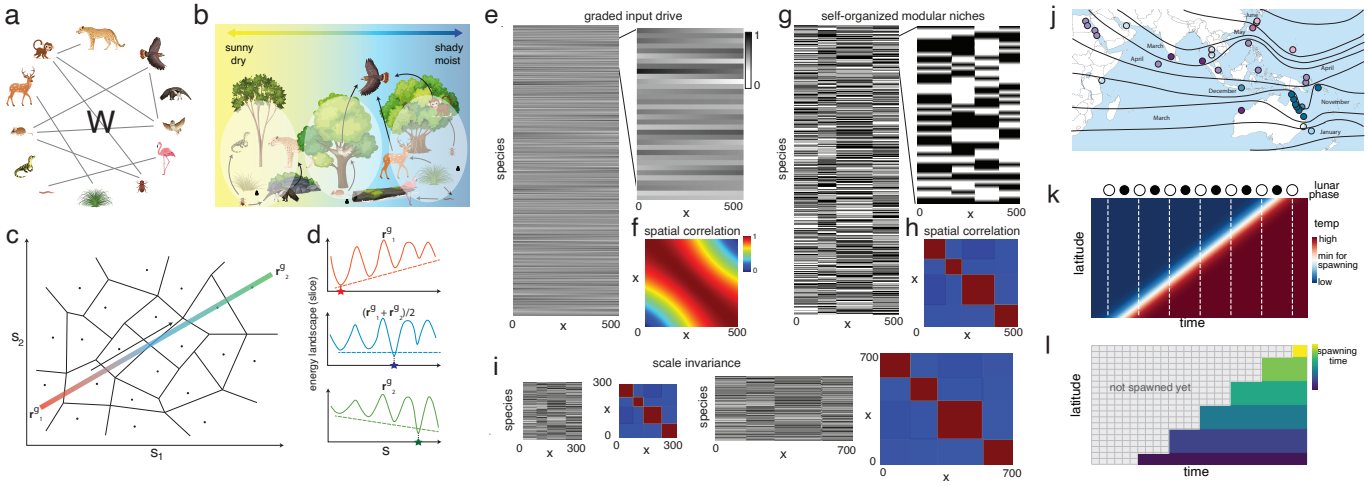


FIG. 8. **Self-organization of ecological niches and synchronous spawning through peak selection**: (a) Schematic web of competitive and cooperative interactions between species. (b) Embedding of the species from (a) into an environment with smoothly varying spatial resource gradients. (c) Schematic of the state space of the interacting species system as in (a), ignoring the spatial distributions of species and resources, has multiple attractor states (dots; basins boundaries depicted by black lines). Multi-colored line: We set up a resource gradient $\mathbf{r}^g(\theta) = (1-\theta)\mathbf{r}_1^g + \theta\ \mathbf{r}_2^g$ by smoothly and linearly interpolating between two random resource vectors. (d) The same one-dimensional slice/view of the energy landscape in (c), at three sampled values along the input resource gradient (top, middle, and bottom, respectively). The resource gradient "tilts" the energy landscape, varying the relative heights of the local optima. (e-h) Plots of species distributions as a function of space (x) (e,g), and their spatial cross-correlations (f,h). (e-f): Initial distribution. (g-h): After convergence of the dynamics there is an emergent self-organization of four ecological niches. (i) Global scale invariance: changing the spatial size of the system (L) while maintaining the local interaction kernel (K) width and the end values of the resource gradient function results in the same niche structure. (j) Data on a number of coral reefs and their mass spawning dates (adapted from [69]). Note the widespread bands of synchronized spawning across the Indo-Pacific (same-color circles). (Open circles: data not available.) (k) Model for synchronized coral spawning: ocean temperature is assumed to undergo seasonal variation, with temperatures rising at lower latitudes first, and eventually in equatorial regions. Along with this spatiotemporal temperature variation, the lunar cycle provides a faster time-scale periodic cue towards coral spawning (l) Spawning occurs in the model at the full-moon lunar phase in all regions where the ocean temperature is warmer than a given threshold. The dynamics results in a step-like variation across latitudes in coral spawn times. *See Methods for parameter and simulation details.*

boundaries leading into distinct patterns of relative population abundances in different parts of the environment (Fig. 8g-h).

As with grid cells, the system exhibits the theoretically predicted property of global scale invariance: if the resource values at the endpoints of the environment are held fixed as the size (L) of the environment is scaled, even though the spatial interaction width (set by $K(x, x')$) is not, the number of formed niches and their composition remains unchanged (Fig. 8i). While we only examine the symmetric $W(i, j)$ case here, we expect this mechanism for self-organization of modular niches to also hold in the more general non-symmetric $W(i, j)$ case. Examining such species interaction networks could lead to spatio-temporal dynamics in species niche formation, which would be an interesting future direction.

Finally, we consider a simplified model of the synchronization and timing of mass spawning by the *Acropora* genus of corals (the dominant shallow-water coral genus in the Indo-Pacific [85]). Across broad geographic swaths, *Acropora* coral colonies spawn in an impressive coordinated burst lasting only a few hours, once a year (Fig. 8j). The day of this spawning depends on multiple factors, including surface water temperature and lunar phase[69, 85–87]: the water must have warmed sufficiently, and the lunar phase should be a few days after the full moon. In the temporal domain,

temperature varies gradually (temporal gradient), while the lunar phase changes periodically and more rapidly. While temperature provides a broad permissive seasonal range for spawning, the faster-cycling lunar cycle locks in a specific spawning day within the broad warming trend, permitting broad synchronization across the population. The peak selection model predicts an additional form of emergence, in the form of spatial organization: surface temperature varies both temporally and spatially, gradually changing with latitude. The model predicts large step-like jumps in the spawning day as a function of space, with different and widely separated coral colonies at similar latitudes spawning on the same day, but others sufficiently offset in latitude spawning a month later, in a step-like spatial progression, Fig. 8k-l). This is broadly consistent with data from coral species in the Indian and Pacific ocean[69], Fig. 8j).

DISCUSSION

Summary We have shown how structure can robustly and spontaneously emerge at multiple scales from purely local interactions. The peak selection theory shows how modular structure can emerge without modular genetic specification (development) or modular interactions (development and ecology). The mechanism exhibits useful features of both positional and pattern formation processes [12]: structure is noise-tolerant as in pattern formation processes yet self-scaling as in positional processes. In the context of pattern formation systems (with multi-bump models for single grid modules being one example), the two-scale peak selection mechanism yields higher robustness to connectivity inhomogeneity beyond the capabilities of conventional attractor models [24].

In the context of grid cells, this work extends and robustifies continuous attractor models [24–26], from single modules to multiple modules, from dependence on specific interaction profiles to an infinite set of kernels for grid emergence, and from dependence on homogeneous weights to a weaker dependence on weight homogeneity. It connects to observed DV gradients in MEC [35, 41, 45, 52–55], and also potentially more generally to observed gradients that underlie discontinuous function in cortex [34, 46].

Predictions The model forms a mechanistic bridge between distinct scales and levels in the grid cell system, showing how biophysical properties, synaptic connectivity, and activity might interact, and providing predictions about whether and when perturbations on one level might constrain and affect outcomes on the other. Relationships between levels in the model provide low-dimensional "knobs" for experimental manipulation and testing. Under the assumption that spatial gradients in cellular properties are themselves driven by gradients in gene expression [88, 89], we highlight a few predictions for connectomics (C), transcriptomics (T), development (D), and physiology (P) studies and connections between them: 1) Functionally discrete modules with sharp boundaries can emerge without sharp boundaries in gene expression or cellular and synaptic properties (T,C,P). It is possible that these emerging modules are later consolidated by differential gene expression patterns, or that gene expression remains graded in the adult (D,T). 2) One type of interaction (synaptic strength or width) or biophysical property (like time-constant that affects interaction strength) that is fixed along the DV axis of MEC, and another that smoothly varies along it. These two types of interactions may reside in different cells and synapses. Alternatively, a single set of synapses might contain both scales, such as a synaptic connectivity profile whose shape is graded along the DV axis, but whose maximal cutoff radius is invariant across the DV axis. (C) 3) The detailed adjacent grid period ratio prediction, given by successive integer ratios or by integer ratios with a correction given by the coarse form of the (fixed-scale) interaction profile (P). 4) Invariance of module number and periods to brain size, if endpoint interactions are fixed or change only slightly; specific predicted variation in module number if endpoint interaction properties are varied (T,C,P). 5) Predicted shift of module boundaries locations but not module number with gradient shape (T,C,P). 6) Predicted relationship between lateral interaction shape

and period ratios, through the scalar variable ϕ. (C,P) Predictions (3-6) can be probed on intra- or interspecies levels. 7) Independence of dynamics between modules and high robustness of dynamics to activity perturbation within and across modules: Effects of perturbation to activity are localized to the module it is applied to, without a cascading effect across modules [68]. Entirely suppressing one module should not alter others, and suppressing half a module boundary should not shift the rest of the boundary (P). 8) Following MEC-wide silencing, activity patterning in all modules should re-emerge independently and in parallel, rather than sequentially [68] (P). 9) If all grid cells receive a common velocity input, then grid modules should rotate in tandem under any remapping because their relative orientations on the cortical sheet are fixed (P); if they are able to rotate independently, it would require independent velocity inputs (P). 10) If two different cell types or processes mediate the graded and fixed components of the lateral interactions, then a manipulation to turn off the fixed-scale interaction should lead to patterning but with continuous variation of period across the DV axis instead of discrete modules.

More broadly, as a mechanistic circuit model at the level of neurons and synaptic connectivity, the model is fully "computable". Thus, it is a tool for generating predictions about the result of any perturbation for which our model has a corresponding parameter or variable, which experimentalists might want to apply.

Related work Our focus is on the theory and mechanisms of the emergence of multi-modular function from prespecified weights and is thus complementary to work that models the *learning* of weights in MEC, through biologically plausible Hebbian-like rules [90] or backpropagation-based learning [91–95]. The learning models generally do not produce multiple modules, and in the rare cases where they seem to, the circuit connectivity that produces them are unknown [91, 95]. The only other work that proposes a network mechanism for multi-grid module emergence [68] has a distinct (stacked) initial and final architecture, and its predictions on dynamics and connectivity as noted above are interestingly and distinguishably different.

The peak-selection principle for module emergence is both an instance and a generalization in some ways of the idea of spatial bifurcation for the emergence of discrete function from smooth gradients [34, 96]. It permits a number of distinct modules to form from smooth variations in the spatial dimension, but the broader theoretical framework generalizes to variations along abstract parametric dimensions, and generalizes from linear instabilities and pattern formation to nonlinear systems like parameteric gradients in nonlinear Hopfield-like systems. We have shown three distinct flavors of peak selection-based modularity emergence: peak selection in a pattern-forming process interacting with smooth gradients in interaction parameters for grid cells; peak selection via a smoothly varying regularization term in dynamics on a rough landscape in the general Lyapunov function approach; and peak selection in a symmetry-breaking process (which is more general than Turing-like pattern formation) based on initial conditions or input gradients that tilt the landscape, with applications to ecological niche emergence and geographically extended bands of synchronized coral spawning.

These concepts provide dynamical and mechanistic principles for *how* modular structure can emerge without modular components, in contrast to normative models that focus on *why* or *when* modular structure is favored [4, 9, 97–105]. Such a mechanistic understanding connects with literature on the emergence of discrete function in brain areas, the spinal cord and the body, despite the existence of smooth biophysical gradients[31, 34–38, 40–47, 96, 106, 107].

An important vein in developmental research is to explore how structure can emerge with precision in the presence of noise [31, 32, 108, 109]. Some solutions within the positional hypothesis involve spatial or temporal integration of noisy gradients [12, 29, 108]. Pattern forming mechanisms confer robustness to noise because the patterned state is much lower-dimensional than the overall state space [110]. Our observation that peak selection contributes significant additional robustness raises the intriguing question of whether such a mechanism might assist in tandem with positional mechanisms during morphogenesis [111].

Acknowledgements

We are grateful to Mehran Kardar for helpful discussions and to John Widloski, Akhilan Boopathy and Ling Dong for comments on the manuscript. This work was supported by the MathWorks Science Fellowship to MK, the Simons Foundation through the Simons Collaboration on the Global Brain, the ONR, and the Howard Hughes Medical Institute through the Faculty Scholars Program to IRF.

Conceptualization, writing: MK, SC, IRF; Coding, analysis: MK, SC.

[1] David Thouless. *Topological quantum numbers in nonrelativistic physics*. World Scientific, 1998.

[2] Marta Sales-Parda. The importance of being modular. *Science*, 6347(357):128–129, 2017.

[3] Anurag Limdi, Alfonso Pérez-Escudero, Aming Li, and Jeff Gore. Asymmetric migration decreases stability but increases resilience in a heterogeneous metapopulation. *Nat Commun*, 9(1):2969, 07 2018.

[4] Jeong-Man Park, Man Chen, Dong Wang, and Michael W Deem. Modularity enhances the rate of evolution in a rugged fitness landscape. *Phys Biol*, 12(2):025001, Mar 2015.

[5] Wilhelm von Humboldt. On the diversity of human language construction and its influence on the mental development of the human species. Cambridge University Press, 2005.

[6] J A Fodor and Z W Pylyshyn. Connectionism and cognitive architecture: a critical analysis. *Cognition*, 28(1-2):3–71, Mar 1988.

[7] Ila R Fiete, Yoram Burak, and Ted Brookings. What grid cells convey about rat location. *J Neurosci*, 28(27):6858–71, Jul 2008.

[8] Sameet Sreenivasan and Ila Fiete. Grid cells generate an analog error-correcting code for singularly precise neural computation. *Nat Neurosci*, 14(10):1330–7, Sep 2011.

[9] Ferran Alet, Tomas Lozano-Perez, and Leslie P. Kaelbling. Modular meta-learning. In Aude Billard, Anca Dragan, Jan Peters, and Jun Morimoto, editors, *Proceedings of The 2nd Conference on Robot Learning*, volume 87 of *Proceedings of Machine Learning Research*, pages 856–868. PMLR, 29–31 Oct 2018.

[10] Man Yi Yim, Lorenzo A Sadun, Ila R Fiete, and Thibaud Taillefumier. Place-cell capacity and volatility with grid-like inputs. *Elife*, 10, May 2021.

[11] Lewis Wolpert. Positional information and the spatial pattern of cellular differentiation. *Journal of theoretical biology*, 25(1):1–47, 1969.

[12] Jeremy BA Green and James Sharpe. Positional information and reaction-diffusion: two big ideas in developmental biology combine. *Development*, 142(7):1203–1211, 2015.

[13] Lucia Durrieu, Daniel Kirrmaier, Tatjana Schneidt, Ilia Kats, Sarada Raghavan, Lars Hufnagel, Timothy E Saunders, and Michael Knop. Bicoid gradient formation mechanism and dynamics revealed by protein lifetime analysis. *Mol Syst Biol*, 14(9):e8355, 09 2018.

[14] AM Turing. The chemical basis of morphogenesis. *Philos Trans R Soc Lond BB*, 237(641), 1952.

[15] Jelena Raspopovic, Luciano Marcon, Laura Russo, and James Sharpe. Digit patterning is controlled by a bmp-sox9-wnt turing network modulated by morphogen gradients. *Science*, 345(6196):566–570, 2014.

[16] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E.I. Moser. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052):801–806, 2005.

[17] J Heys and D Dombeck. Cellular resolution optical imaging of medial entorhinal cortex. In *SFN Poster*, 2014.

[18] Yi Gu, Sam Lewallen, Amina A Kinkhabwala, Cristina Domnisoru, Kijung Yoon, Jeffrey L Gauthier, Ila R Fiete, and David W Tank. A map-like micro-organization of grid cells in the medial entorhinal cortex. *Cell*, 175(3):736–750.e30, Oct 2018.

[19] K.J. Yoon, M.A. Buice, R. Barry, C.and Hayman, N. Burgess, and I.R. Fiete. Specific evidence of low-dimensional continuous attractor dynamics in grid cells. *Nat Neurosci*, 16(8):1077–84, Aug 2013.

[20] S.G. Trettel, J.B. Trimper, E. Hwaun, I.R. Fiete, and L.L. Colgin. Grid cell co-activity patterns during sleep reflect spatial overlap of grid fields during active behaviors. *Nat Neurosci*, 22(4):609–617, 04 2019.

[21] Richard J Gardner, Li Lu, Tanja Wernle, May-Britt Moser, and Edvard I Moser. Correlation structure of grid cells is preserved during sleep. *Nat Neurosci*, 22(4):598–608, 04 2019.

[22] Richard J Gardner, Erik Hermansen, Marius Pachitariu, Yoram Burak, Nils A Baas, Benjamin A Dunn, May-Britt Moser, and Edvard I Moser. Toroidal topology of population activity in grid cells. *Nature*, pages 1–6, 2022.

[23] Mikail Khona and Ila R Fiete. Attractor and integrator networks in the brain. *Nature Reviews Neuroscience*, 23:744–766, 2022.

[24] Yoram Burak and Ila R Fiete. Accurate path integration in continuous attractor network models of grid cells. *PLoS Comput Biol*, 5(2):e1000291, Feb 2009.

[25] Mark C Fuhs and David S Touretzky. A spin glass model of path integration in rat medial entorhinal cortex. *J Neurosci*, 26(16):4266–4276, 2006.

[26] Alexis Guanella, Daniel Kiper, and Paul Verschure. A model of grid cells based on a twisted torus topology. *Int. J. Neural Syst.*, 17(4):231–240, August 2007.

[27] H. Stensola, T. Stensola, T. Solstad, K. Fröland, M. Moser, and E. Moser. The entorhinal grid map is discretized. *Nature*, 492(7427):72–8, Dec 2012.

[28] Michel Kerszberg and Lewis Wolpert. Specifying positional information in the embryo: looking beyond morphogens. *Cell*, 130(2):205–209, 2007.

[29] Gašper Tkačik and Thomas Gregor. The many bits of positional information. *Development*, 148(2):dev176065, 2021.

[30] François Schweisguth and Francis Corson. Self organization in pattern formation. *Dev. Cell*, (49):659–677, Jun 2019.

[31] Thomas Gregor, David W Tank, Eric F Wieschaus, and William Bialek. Probing the limits to positional information. *Cell*, 130(1):153–164, 2007.

[32] Shawn C Little, Mikhail Tikhonov, and Thomas Gregor. Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell*, 154(4):789–800, 2013.

[33] Naama Barkai and Ben-Zion Shilo. Variability and robustness in biomolecular systems. *Molecular cell*, 28(5):755–760, 2007.

[34] Xiao-Jing Wang. Macroscopic gradients of synaptic excitation and inhibition in the neocortex. *Nature Reviews Neuroscience*, 21(3):169–178, 2020.

[35] P. Beed, A. Gundlfinger, S. Schneiderbauer, J. Song, C. Bohm, A. Burgalossi, M. Brecht, I. Vida, and D. Schmitz. Inhibitory gradient along the dorsoventral axis in the medial entorhinal cortex. *Neuron*, 79:1197–1207, 2013.

[36] Andrew Miri, Kayvon Daie, Aristides B Arrenberg, Herwig Baier, Emre Aksay, and David W Tank. Spatial gradients and multidimensional dynamics in a neural integrator circuit. *Nature neuroscience*, 14(9):1150–1159, 2011.

[37] Joshua B Burt, Murat Demirtaş, William J Eckner, Natasha M Navejar, Jie Lisa Ji, William J Martin, Alberto Bernacchia, Alan Anticevic, and John D Murray. Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nature neuroscience*, 21(9):1251–1259, 2018.

[38] Bertha Vázquez-Rodríguez, Laura E Suárez, Ross D Markello, Golia Shafiei, Casey Paquola, Patric Hagmann, Martijn P Van Den Heuvel, Boris C Bernhardt, R Nathan Spreng, and Bratislav Misic. Gradients of structure–function tethering across neocortex. *Proceedings of the National Academy of Sciences*, 116(42):21219–21227, 2019.

[39] Hugh Pastoll, Derek L Garden, Ioannis Papastathopoulos, Gülşen Sürmeli, and Matthew F Nolan. Inter-and intra-animal variation in the integrative properties of stellate cells in the medial entorhinal cortex. *Elife*, 9:e52258, 2020.

[40] Ben D Fulcher, John D Murray, Valerio Zerbi, and Xiao-Jing Wang. Multimodal gradients across mouse cortex. *Proceedings of the National Academy of Sciences*, 116(10):4689–4695, 2019.

[41] Derek L F Garden, Paul D Dodson, Cian O'Donnell, Melanie D White, and Matthew F Nolan. Tuning of synaptic integration in the medial entorhinal cortex to the organization of grid cell firing fields. *Neuron*, 60(5):875–89, Dec 2008.

[42] Katherine W Rogers and Alexander F Schier. Morphogen gradients: from generation to interpretation. *Annu Rev Cell Dev Biol*, 27:377–407, 2011.

[43] François Schweisguth and Francis Corson. Self-organization in pattern formation. *Dev Cell*, 49(5):659–677, 06 2019.

[44] Leigh Wilson and Malcolm Maden. The mechanisms of dorsoventral patterning in the vertebrate neural tube. *Dev Biol*, 282(1):1–13, Jun 2005.

[45] Sabine Grosser, Federico J Barreda, Prateep Beed, Dietmar Schmitz, Sam A Booker, and Imre Vida. Parvalbumin interneurons are differentially connected to principal cells in inhibitory feedback microcircuits along the dorsoventral axis of the medial entorhinal cortex. *Eneuro*, 8(1), 2021.

[46] BRAIN Initiative Cell Census Network (BICCN). A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature*, 598:86–102, 2021.

[47] Meng Zhang, Stephen W Eichhorn, Brian Zingg, Zizhen Yao, Kaelan Cotter, Hongkui Zeng, Hongwei Dong, and Xiaowei Zhuang. Spatially resolved cell atlas of the mouse primary motor cortex by merfish. *Nature*, 598(7879):137–143, 2021.

[48] D St Johnston and C Nüsslein-Volhard. The origin of pattern and polarity in the drosophila embryo. *Cell*, 68(2):201–19, Jan 1992.

[49] Stephen W Paddock, Eric J Hazen, and Peter J DeVries. Methods and applications of three-color confocal imaging. *BioTechniques*, 22(1):120–126, 1997.

[50] Alexis Guanella, Daniel Kiper, and Paul Verschure. A model of grid cells based on a twisted torus topology. *Int J Neural Syst*, 17(4):231–240, 2007.

[51] Jonathan J Couey, Aree Witoelar, Sheng-Jia Zhang, Kang Zheng, Jing Ye, Benjamin Dunn, Rafal Czajkowski, May-Britt Moser, Edvard I Moser, Yasser Roudi, and Menno P Witter. Recurrent inhibitory circuitry as a mechanism for grid formation. *Nat Neurosci*, 16(3):318–24, Mar 2013.

[52] Lisa M Giocomo, Eric A Zilli, Erik Fransén, and Michael E Hasselmo. Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing. *Science*, 315(5819):1719–1722, 2007.

[53] Lisa M Giocomo, Syed A Hussaini, Fan Zheng, Eric R Kandel, May-Britt Moser, and Edvard I Moser. Grid cells use hcn1 channels for spatial scaling. *Cell*, 147(5):1159–1170, 2011.

[54] John J Tukker, Prateep Beed, Michael Brecht, Richard Kempter, Edvard I Moser, and Dietmar Schmitz. Microcircuits for spatial coding in the medial entorhinal cortex. *Physiological reviews*, 102(2):653–688, 2022.

[55] James G Heys, Lisa M Giocomo, and Michael E Hasselmo. Cholinergic modulation of the resonance properties of stellate cells in layer ii of medial entorhinal cortex. *Journal of Neurophysiology*, 104(1):258–270, 2010.

[56] H S Seung. How the brain keeps the eyes still. *Proc Natl Acad Sci U S A*, 93(23):13339–13344, 1996.

[57] Bailu Si, Sandro Romani, and Misha Tsodyks. Continuous attractor network model for conjunctive position-by-velocity tuning of grid cells. *Plos computational biology*, 10(4):e1003558, 2014.

[58] Mark S Goldman, Joseph H Levine, Guy Major, David W Tank, and H Sebastian Seung. Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. *Cerebral cortex*, 13(11):1185–1195, 2003.

[59] Sung Soo Kim, Hervé Rouault, Shaul Druckmann, and Vivek Jayaraman. Ring attractor dynamics in the drosophila central brain. *Science*, 356(6340):849–853, 05 2017.

[60] Philip W Anderson. Absence of diffusion in certain random lattices. *Physical review*, 109(5):1492, 1958.

[61] Xue-Xin Wei, Jason Prentice, and Vijay Balasubramanian. A principle of economy predicts the functional architecture of grid cells. *Elife*, 4:e08362, Sep 2015.

[62] Martin Stemmler, Alexander Mathis, and Andreas VM Herz. Connecting multiple spatial scales to decode the population activity of grid cells. *Science Advances*, 1(11):e1500816, 2015.

[63] K Zhang. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J Neurosci*, 16(6):2112–26, Mar 1996.

[64] Arseny Finkelstein, Dori Derdikman, Alon Rubin, Jakob N Foerster, Liora Las, and Nachum Ulanovsky. Three-dimensional head-direction coding in the bat brain. *Nature*, 517(7533):159–164, 2015.

[65] Yoram Burak and Ila Fiete. Do we understand the emergent dynamics of grid cell activity? *J Neurosci*, 26(37):9352–9354, 2006.

[66] John Widloski, Michael P Marder, and Ila R Fiete. Inferring circuit mechanisms from sparse neural recording and global perturbation in grid cells. *Elife*, 7, 07 2018.

[67] Yi Gu, Sam Lewallen, Amina A Kinkhabwala, Cristina Domnisoru, Kijung Yoon, Jeffrey L Gauthier, Ila R Fiete, and David W Tank. A map-like micro-organization of grid cells in the medial entorhinal cortex. *Cell*, 175(3):736–750.e30, 10 2018.

[68] Louis Kang and Vijay Balasubramanian. A geometric attractor mechanism for self-organization of entorhinal grid modules. *Elife*, 8, 08 2019.

[69] Sally A Keith, Jeffrey A Maynard, Alasdair J Edwards, James R Guest, Andrew G Bauman, Ruben Van Hooidonk, Scott F Heron, Michael L Berumen, Jessica Bouwmeester, Srisakul Piromvaragorn, et al. Coral mass spawning predicted by rapid seasonal rise in ocean temperature. *Proceedings of the Royal Society B: Biological Sciences*, 283(1830):20160011, 2016.

[70] Daniel A Power, Richard A Watson, Eörs Szathmáry, Rob Mills, Simon T Powers, C Patrick Doncaster, and BłaŻej Czapp. What can ecosystems learn? expanding evolutionary ecology with learning theory. *Biology direct*, 10:1–24, 2015.

[71] Giulio Biroli, Guy Bunin, and Chiara Cammarota. Marginally stable equilibria in critical ecosystems. *New Journal of Physics*, 20(8):083051, 2018.

[72] Michael Turelli. Niche overlap and invasion of competitors in random environments i. models without demographic stochasticity. *Theoretical Population Biology*, 20(1):1–56, 1981.

[73] Giulia Garcia Lorenzana and Ada Altieri. Well-mixed lotka-volterra model with random strongly competitive interactions. *Physical Review E*, 105(2):024307, 2022.

[74] John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982.

[75] Jordan G Okie, David J Van Horn, David Storch, John E Barrett, Michael N Gooseff, Lenka Kopsova, and Cristina D Takacs-Vesbach. Niche and metabolic principles explain patterns of diversity and distribution: theory and a case study with soil bacterial communities. *Proceedings of the Royal Society B: Biological Sciences*, 282(1809):20142630, 2015.

[76] Diego Alarcón and Lohengrin A Cavieres. Relationships between ecological niche and expected shifts in elevation and latitude due to climate change in south american temperate forest plants. *Journal of Biogeography*, 45(10):2272–2287, 2018.

[77] David Schellenberger Costa, Friederike Gerschlauer, Ralf Kiese, Markus Fischer, Michael Kleyer, and Andreas Hemp. Plant niche breadths along environmental gradients and their relationship to plant functional traits. *Diversity and Distributions*, 24(12):1869–1882, 2018.

[78] Alex Salas-López, Cyrille Violle, François Munoz, Florian Menzel, and Jérôme Orivel. Effects of habitat and competition on niche partitioning and community structure in neotropical ants. *Insect Communities: Diversity Patterns and their Driving Forces*, 16648714, 2023.

[79] Lu Luan, Yuji Jiang, Francisco Dini-Andreote, Thomas W Crowther, Pengfa Li, Mohammad Bahram, Jie Zheng, Qinsong Xu, Xue-Xian Zhang, and Bo Sun. Integrating ph into the metabolic theory of ecology to predict bacterial diversity in soil. *Proceedings of the National Academy of Sciences*, 120(3):e2207832120, 2023.

[80] Meaghan Conway. *Niche Evolution Along a Gradient of Ecological Specialization*. The University of Maine, 2019.

[81] James H Brown, James F Gillooly, Andrew P Allen, Van M Savage, and Geoffrey B West. Toward a metabolic theory of ecology. *Ecology*, 85(7):1771–1789, 2004.

[82] Aleksej Zelezniak, Sergej Andrejev, Olga Ponomarova, Daniel R Mende, Peer Bork, and Kiran Raosaheb Patil. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proceedings of the National Academy of Sciences*, 112(20):6449–6454, 2015.

[83] Brian J Enquist, Evan P Economo, Travis E Huxman, Andrew P Allen, Danielle D Ignace, and James F Gillooly. Scaling metabolism from organisms to ecosystems. *Nature*, 423(6940):639–642, 2003.

[84] Brian J Enquist, James H Brown, and Geoffrey B West. Allometric scaling of plant energetics and population density. *Nature*, 395(6698):163–165, 1998.

[85] Che-Hung Lin, Shunichi Takahashi, Aziz J Mulla, and Yoko Nozawa. Moonrise timing is key for synchronized spawning in coral dipsastraea speciosa. *Proceedings of the National Academy of Sciences*, 118(34):e2101985118, 2021.

[86] Paulina Kaniewska, Shahar Alon, Sarit Karako-Lampert, Ove Hoegh-Guldberg, and Oren Levy. Signaling cascades and the importance of moonlight in coral broadcast mass spawning. *elife*, 4:e09991, 2015.

[87] L Penland, J Kloulechad, D Idip, and R Van Woesik. Coral spawning in the western pacific ocean is related to solar insolation: evidence of multiple spawning events in palau. *Coral Reefs*, 23:133–140, 2004.

[88] Zizhen Yao, Cindy T J van Velthoven, Michael Kunst, Meng Zhang, Delissa McMillen, Changkyu Lee, Won Jung, Jeff Goldy, Aliya Abdelhak, Matthew Aitken, Katherine Baker, Pamela Baker, Eliza Barkan, Darren Bertagnolli, Ashwin Bhandiwad, Cameron Bielstein, Prajal Bishwakarma, Jazmin Campos, Daniel Carey, Tamara Casper, Anish Bhaswanth Chakka, Rushil Chakrabarty, Sakshi Chavan, Min Chen, Michael Clark, Jennie Close, Kirsten Crichton, Scott Daniel, Peter DiValentin, Tim Dolbeare, Lauren Ellingwood, Elysha Fiabane, Timothy Fliss, James Gee, James Gerstenberger, Alexandra Glandon, Jessica Gloe, Joshua Gould, James Gray, Nathan Guilford, Junitta Guzman, Daniel Hirschstein, Windy Ho, Marcus Hooper, Mike Huang, Madie Hupp, Kelly Jin, Matthew Kroll, Kanan Lathia, Arielle Leon, Su Li, Brian Long, Zach Madigan, Jessica Malloy, Jocelin Malone, Zoe Maltzer, Naomi Martin, Rachel McCue, Ryan McGinty, Nicholas Mei, Jose Melchor, Emma Meyerdierks, Tyler Mollenkopf, Skyler Moonsman, Thuc Nghi Nguyen, Sven Otto, Trangthanh Pham, Christine Rimorin, Augustin Ruiz, Raymond Sanchez, Lane Sawyer, Nadiya Shapovalova, Noah Shepard, Cliff Slaughterbeck, Josef Sulc, Michael Tieu, Amy Torkelson, Herman Tung, Nasmil Valera Cuevas, Shane Vance, Katherine Wadhwani, Katelyn Ward, Boaz Levi, Colin Farrell, Rob Young, Brian Staats, Ming-Qiang Michael Wang, Carol L Thompson, Shoaib Mufti, Chelsea M Pagan, Lauren Kruse, Nick Dee, Susan M Sunkin, Luke Esposito, Michael J Hawrylycz, Jack Waters, Lydia Ng, Kimberly Smith, Bosiljka Tasic, Xiaowei Zhuang, and Hongkui Zeng. A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature*, 624(7991):317–332,

December 2023.

[89] Ben D Fulcher, John D Murray, Valerio Zerbi, and Xiao-Jing Wang. Multimodal gradients across mouse cortex. *Proc. Natl. Acad. Sci. U. S. A.*, 116(10):4689–4695, March 2019.

[90] John Widloski and Ila R Fiete. A model of grid cell development through spatial exploration and spike time-dependent plasticity. *Neuron*, 83(2):481–495, Jul 2014.

[91] Andrea Banino, Caswell Barry, Benigno Uria, Charles Blundell, Timothy Lillicrap, Piotr Mirowski, Alexander Pritzel, Martin J Chadwick, Thomas Degris, Joseph Modayil, Greg Wayne, Hubert Soyer, Fabio Viola, Brian Zhang, Ross Goroshin, Neil Rabinowitz, Razvan Pascanu, Charlie Beattie, Stig Petersen, Amir Sadik, Stephen Gaffney, Helen King, Koray Kavukcuoglu, Demis Hassabis, Raia Hadsell, and Dharshan Kumaran. Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, 05 2018.

[92] Christopher J. Cueva and Xue-Xin Wei. Emergence of grid-like representations by training recurrent neural networks to perform spatial localization. In *International Conference on Learning Representations*, 2018.

[93] Ben Sorscher, Gabriel Mel, Surya Ganguli, and Samuel Ocko. A unified theory for the origin of grid cells through the lens of pattern formation. In *Advances in Neural Information Processing Systems*, pages 10003–10013, 2019.

[94] Rylan Schaeffer, Mikail Khona, and Ila R Fiete. No free lunch from deep learning in neuroscience: A case study through models of the entorhinal-hippocampal circuit. *Advances in neural information processing systems*, 2022.

[95] Rylan Schaeffer, Mikail Khona, Sanmi Koyejo, and Ila Rani Fiete. Disentangling fact from grid cell fiction in trained deep path integrators. *arXiv preprint arXiv:2312.03954*, 2023.

[96] Xiao-Jing Wang. Theory of the multiregional neocortex: large-scale neural dynamics and distributed cognition. *Annual review of neuroscience*, 45:533–560, 2022.

[97] Guang Chen, Byungwoo Kang, Jack Lindsey, Shaul Druckmann, and Nuo Li. Modularity and robustness of frontal cortical networks. *Cell*, 184(14):3717–3730, 2021.

[98] Nadav Kashtan, Avi E Mayo, Tomer Kalisky, and Uri Alon. An analytically solvable model for rapid evolution of modular structure. *PLoS computational biology*, 5(4):e1000355, 2009.

[99] Giambattista Parascandolo, Niki Kilbertus, Mateo Rojas-Carulla, and Bernhard Schölkopf. Learning independent causal mechanisms. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 4036–4044. PMLR, 2018.

[100] Guangyu Robert Yang, Madhura R Joglekar, H Francis Song, William T Newsome, and Xiao-Jing Wang. Task representations in neural networks trained to perform many cognitive tasks. *Nat. Neurosci.*, 22(2):297–306, 2019.

[101] Andrea *et al* Banino. Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, 2018.

[102] Jun Sun and Michael W. Deem. Spontaneous emergence of modularity in a model of evolving individuals. *Physical Review Letters*, 99(228107), 2007.

[103] Jacob Andreas, Dan Klein, and Sergey Levine. Modular multitask reinforcement learning with policy sketches. In *International Conference on Machine Learning*, pages 166–175. PMLR, 2017.

[104] Nadav Kashtan and Uri Alon. Spontaneous evolution of modularity and network motifs. *Proc Natl Acad Sci U S A*, 102(39):13773–8, Sep 2005.

[105] Trevor McCourt, Ila R Fiete, and Isaac L Chuang. Noisy dynamical systems evolve error correcting codes and modularity. *arXiv preprint arXiv:2303.14448*, 2023.

[106] Katarzyna Kuzmicz-Kowalska and Anna Kicheva. Regulation of size and scale in vertebrate spinal cord development. *Wiley Interdisciplinary Reviews: Developmental Biology*, 10(3):e383, 2021.

[107] Donna M Fekete. Gradients in the developing spinal cord–mitogens join forces with morphogens. *TRENDS in Genetics*, 18(8):393, 2002.

[108] Brian A Camley. Collective gradient sensing and chemotaxis: modeling and recent developments. *Journal of Physics: Condensed Matter*, 30(22):223001, 2018.

[109] Carlos S Galvan-Ampudia, Guillaume Cerutti, Jonathan Legrand, Géraldine Brunoud, Raquel Martin-Arevalillo, Romain Azais, Vincent Bayle, Steven Moussu, Christian Wenzl, Yvon Jaillais, et al. Temporal integration of auxin information for the regulation of patterning. *Elife*, 9:e55832, 2020.

[110] Yoram Burak and Ila R Fiete. Fundamental limits on persistent activity in networks of noisy neurons. *Proc Natl Acad Sci U S A*, 109(43):17645–50, Oct 2012.

[111] Avishai Gavish, Arkadi Shwartz, Abraham Weizman, Eyal Schejter, Ben-Zion Shilo, and Naama Barkai. Periodic patterning of the drosophila eye is stabilized by the diffusible activator scabrous.

Nature communications, 7(1):1–10, 2016.
[112] Mark S Goldman, A Compte, and Xiao-Jing Wang. Neural integrator models. *Encyclopedia of neuroscience*, pages 165–178, 2010.
[113] Sophie Deneve, Peter E Latham, and Alexandre Pouget. Reading population codes: a neural implementation of ideal observers. *Nature neuroscience*, 2(8):740–745, 1999.
[114] Nicolas Brunel. Is cortical connectivity optimized for storing information? *Nature neuroscience*, 19(5):749–755, 2016.
[115] KiJung Yoon, Sam Lewallen, Amina A Kinkhabwala, David W Tank, and Ila R Fiete. Grid cell responses in 1d environments assessed as slices through a 2d lattice. *Neuron*, 89(5):1086–99, Mar 2016.
[116] Kijung Yoon, Michael A Buice, Caswell Barry, Robin Hayman, Neil Burgess, and Ila R Fiete. Specific evidence of low-dimensional continuous attractor dynamics in grid cells. *Nat Neurosci*, 16(8):1077–84, Aug 2013.
[117] J. J. Couey, A. Witoelar, S. Zhang, K. Zheng, J. Ye, B. Dunn, R. Czajkowski, M. Moser, E. Moser, Y. Roudi, and M. Witter. Recurrent inhibitory circuitry as a mechanism for grid formation. *Nat Neurosci*, 16(3):318–24, Mar 2013.
[118] M V Tsodyks and M V Feigelman. The enhanced storage capacity in neural networks with low activity level. *Europhys. Lett.*, 6(2):101–105, May 1988.
[119] LM Pismen. Pattern selection at the bifurcation point. *The Journal of Chemical Physics*, 72(3):1900–1907, 1980.
[120] M. C. Cross and P. C. Hohenberg. Pattern formation outside of equilibrium. *Reviews of Modern Physics*, 65(3), 1993.
[121] Michael Cross and Henry Greenside. *Pattern formation and dynamics in nonequilibrium systems*. Cambridge University Press, 2009.
[122] Ben Sorscher, Gabriel C Mel, Samuel A Ocko, Lisa M Giocomo, and Surya Ganguli. A unified theory for the computational and mechanistic origins of grid cells. *Neuron*, 2022.
[123] JE Widloski. Grid cell attractor networks: development and implications. UT Ph.D. Theses and Dissertations.
[124] A. Gierer and H. Meinhardt. A theory of biological pattern formation. *Kybernetik*, 12(1):30–39, dec 1972.
[125] James Dickson Murray. *Mathematical Biology*. Number Ch. 16. Springer, Berlin, 2003.
[126] Ernest Oliver Tuck. On positivity of fourier transforms. *Bulletin of the Australian Mathematical Society*, 74(1):133–138, 2006.
[127] BG Giraud and R Peschanski. On positive functions with positive fourier transforms. *arXiv preprint math-ph/0504015*, 2005.
[128] Bertrand G Giraud and Robi Peschanski. On the positivity of fourier transforms. *arXiv preprint arXiv:1405.3155*, 2014.
[129] Julija Krupic, Marius Bauza, Stephen Burton, Caswell Barry, and John O'Keefe. Grid cell symmetry is shaped by environmental geometry. *Nature*, 518(7538):232–235, 2015.
[130] Tor Stensola, Hanne Stensola, May-Britt Moser, and Edvard I Moser. Shearing-induced asymmetry in entorhinal grid cells. *Nature*, 518(7538):207–212, 2015.
[131] Guifen Chen, Daniel Manson, Francesca Cacucci, and Thomas Joseph Wills. Absence of visual input results in the disruption of grid cell firing in the mouse. *Current Biology*, 26(17):2335–2342, 2016.
[132] John J Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc Natl Acad Sci U S A*, 81:3088–3092, May 1984.
[133] Andrew L Krause, Václav Klika, Thomas E Woolley, and Eamonn A Gaffney. From one pattern into another: analysis of turing patterns in heterogeneous domains via wkbj. *J R Soc Interface*, 17(162):20190621, 01 2020.
[134] Rishidev Chaudhuri, Alberto Bernacchia, and Xiao-Jing Wang. A diversity of localized timescales in network activity. *Elife*, 3:e01239, 2014.

MATERIALS AND METHODS

a. Grid cell continuous attractor network We use a continuous attractor network (CAN) model [112–114] for grid cells [18, 24, 115, 116], with neural dynamics obeying

$$\frac{\partial s(i,t)}{\partial t} + \frac{s(i,t)}{\tau} = \phi\left[\sum_j W_0(i,j)s(j,t) + B(i,t)\right], \tag{8}$$

where $s(i,t)$ represents the synaptic activation of neuron i at time t, $W_0(i,j)$ represents the synaptic strength of the coupling from neuron j to neuron i, $B(i,t)$ represents the feed-forward bias to neuron i, and ϕ is a non-decreasing nonlinearity, for which we use the rectification function ($\phi(z) = [z]_+ = z$ for $z > 0$ and 0 otherwise). Each neuron i has a preferred direction θ_i that is used to perform velocity integration. In the one-dimensional version of our setup, each spatial location $\mathbf{x}$ on the neural sheet has two neurons, with preferred directions $\theta = 0$ and $\theta = \pi$. Correspondingly, in the two-dimensional version of our setup, each location on the neural sheet has four neurons, with preferred directions $\theta = n\pi/4$ for $n \in \{0, 1, 2, 3\}$. The synaptic weights $W_0(i,j)$ are defined via an interaction kernel $W(\Delta x)$ such that

$$W_0(i,j) = W(|\mathbf{x}_i - \mathbf{x}_j - \mathbf{\Delta s}(\theta_j)|), \tag{9}$$

where $\mathbf{x}_i$ represents the spatial location of neuron i, and $\mathbf{l}(\theta)$ is a vector with length Δs oriented parallel to the angle θ. The feed-forward bias $B(i,t;\theta)$ is given by

$$B(i,t) = b + b_{vel}|\mathbf{v}|\cos(\theta_\mathbf{i} - \psi), \tag{10}$$

where ψ is the direction of the input velocity signal and $|\mathbf{v}|$ is the speed. This results in neurons with direction preference θ driving activity in the network towards the direction of their outgoing weight shifts $\mathbf{\Delta s}(\theta)$. This mechanism is responsible for velocity integration by the network [24].

We first described the dynamics under fixed arbitrary kernels, demonstrating that they result in hexagonal pattern formation. These arbitrary kernels were constructed by interpolating between random points via the following protocol: First, we construct 'x-values' by considering $n + n_{zero}$ uniformly spaced points from $-L$ to L, which are then perturbed by the addition of a randomly sampled number from $-L/4n$ to $L/4n$ (this perturbation makes the points less regular, while disallowing consecutive points to be extremely close to each other). Second, we construct n 'y-values' sampled from a uniform distribution from -1 to 1, and define the remaining n_{zero} y-values to be 0 (the n_{zero} values at zero ensure that the interpolated function decays to zero). Then, a cubic spline interpolation (top row of Fig. 10a) or a linear interpolation (bottom row of Fig. 10a) is performed between the y-values and the x-values to generate an arbitrary function $\omega(x)$. This generated function is however not symmetric, as is required for kernel functions — thus, we construct the interaction kernel as $W(\Delta x) = \omega(\Delta x) + \omega(-\Delta x)$. Kernels whose dynamics lead to infinitely diverging firing rates are rejected and resampled. These kernels were simulated on sheets with 256×256 neurons with aperiodic boundary conditions[24]. n was randomly chosen between 2, 3 or 4, and L was scaled as necessary to obtain a large number of activity bumps on the sheet to prevent finite-size effects from distorting the hexagonal lattice of activity.

For the case of module formation through peak selection, the interaction weight kernel W is given by the sum of two components $W = W^g_{n_{DV}} + W^f$. The first, $W^g_{n_{DV}}$ drives local pattern formation, and has a spatial scale $\sigma(n_{DV})$, which varies smoothly in a gradient along the dorso-ventral axis, and the second, W^f has a fixed spatial scale d everywhere on the neural sheet. A variety of functions $W^g_\mathbf{x}$ can drive local pattern formation. For concreteness, we use two specific examples: the Mexican-hat profile[24] (used in Figs. 2a-c,j,k, 11 and SI Fig. 12)

$$W^g_{\text{mexican-hat}}(\Delta x) = \\ \alpha_E \exp\left[-\gamma\frac{(\Delta x)^2}{2\sigma_{mh}(n_{DV})^2}\right] - \alpha_I \exp\left[-\frac{(\Delta x)^2}{2\sigma_{mh}(n_{DV})^2}\right], \tag{11}$$

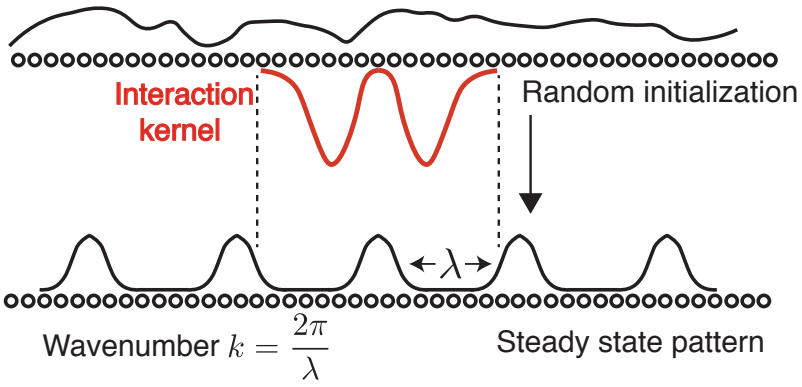


FIG. 9. Local pattern formation in continuous attractor models of grid cells: Through local amplification of random fluctuations, the lateral interaction forms periodic patterns.

and the box-function profile[117] (used in Fig. 2 and SI Fig. 12)

$$W^g_{\text{box}}(\Delta x) = \alpha_0 \times \mathbb{1}_{|\Delta x|<\sigma_b(\mathbf{x})} = \begin{cases} \alpha_0 & \text{if } |\Delta x| < \sigma_b(n_{DV}), \\ 0 & \text{if } |\Delta x| \geq \sigma_b(n_{DV}). \end{cases} \tag{12}$$

For the fixed-width interaction $W^f(\Delta x)$, we implement 3 main types — localized (used in Figs. 2,3 and SI Fig. 12), diffuse (used in Fig. 2 and SI Fig. 12) and decaying (used in SI Fig. 12).

$$\begin{aligned} W^f_{\text{localized}}(\Delta x) &= \alpha_S \exp\left[-\frac{(|\Delta x| - d_{loc})^2}{2\epsilon_S^2}\right], \\ W^f_{\text{diffuse}}(\Delta x) &= \alpha_1 \times \mathbb{1}_{|\Delta x|<d_{dif}}, \\ W^f_{\text{decaying}}(\Delta x) &= \alpha_T \times [d_{dec} - |\Delta x|]_+. \end{aligned}$$

In particular,

- In Figs. 11 we use only a smoothly varying Mexican-hat pattern forming kernel $W = W^g_{\text{mexican-hat}}$
- In Figs. 2a-c,g, j,k we use $W = W^g_{\text{mexican-hat}} + W^f_{\text{localized}}$
- In Figs. 2d-f,h we use a 'Lincoln hat' profile $W = W^g_{\text{box}} + W^f_{\text{diffuse}}$
- and, in SI Fig. 12 we present numerical simulations of other combinations of pattern forming and fixed-scale kernels.

To construct spatially heterogeneous kernels for analyzing the robustness to inhomogeneity in Fig. 6 we use the box function to construct

$$\begin{aligned} W^g_\xi[\mathbf{x}, \mathbf{x}'] &= \alpha_0 \times \mathbb{1}_{|\mathbf{x}-\mathbf{x}'|\times(1+\xi^g_1(\mathbf{x}'))<\sigma^0_b(1+\xi^g_2(\mathbf{x}'))}, \\ W^f_\xi[\mathbf{x}, \mathbf{x}'] &= \alpha_1 \times \mathbb{1}_{|\mathbf{x}-\mathbf{x}'|\times(1+\xi^f_1(\mathbf{x}'))<d_{dif}(1+\xi^f_2(\mathbf{x}'))}, \end{aligned}$$

Parameter	Value
τ	30
dt	0.05
b	$\begin{cases} 70 & \text{in 1D} \\ 1 & \text{in 2D} \end{cases}$
b_{vel}	$\begin{cases} 105 & \text{in 1D} \\ 1 & \text{in 2D} \end{cases}$
Δs	2

TABLE I. Parameters held constant across all numerical simulations

$W^g_{\text{mexican-hat}}$ parameters	Value
α_E	1000
α_I	1000
γ	1.05
N^{1D}	3000
N^{2D}_y	100
N^{2D}_x	1000
$\sigma_{mh}(n_{DV})$	$1/\sqrt{2\beta(n_{DV})}$
$\beta(n_{DV})$	$\beta_0 + (\beta_1 - \beta_0)n_{DV}/N'$
N'	$\begin{cases} N^{1D} & \text{in 1D} \\ N^{2D}_x & \text{in 2D} \end{cases}$
β_0	$\begin{cases} 2.5 \times 10^{-2} & \text{in 1D} \\ 3/676 & \text{in 2D} \end{cases}$
β_1	$\begin{cases} 2.5 \times 10^{-1} & \text{in 1D} \\ 9/338 & \text{in 2D} \end{cases}$
W^g_{box} parameters	Value
N^{1D}	5000
α_0	-40
$\sigma_b(n_{DV})$	$15 + 30n_{DV}/N$

TABLE II. Pattern forming kernel parameters used for numerical simulations

where $\xi^{g,f}_{1,2}(\mathbf{x}')$ are independent random numbers chosen uniformly from $\epsilon^{g,f}_{1,2} \times [-1, 1]$. In the particular case of Fig. 6b, ϵ^g_2 is varied along the x-axis of the plot, and other noise terms are set to zero (In the one-dimensional case ϵ_1 and ϵ_2 have the same effect); for Fig. 6c, $\epsilon^g_1 = 0.2$, $\epsilon^g_2 = 0.3$, $\epsilon^f_1 = 0.05$ and $\epsilon^f_2 = 0$ (See SI Sec. D 9 for more details).

In Table I we present a list of common parameters used across all numerical grid-cell simulations. Then, in Tables II,III we present the parameter values used for the kernels used in our numerical simulations

b. Ecological niche formation For the modular niche formation, we consider the setup as described in Eq. 7, with $N = 1000$ species, each characterized by a random $M = 2000$ dimensional random feature vector indicating resource preference. We numerically simulate our setup on a discrete lattice $x \in \{0, L\}$ for $L = 300$ in Fig. 8i left, and $L = 500$ otherwise.We instantiate the nonlinearity H as a shifted Heaviside function, $H[x] = 1$ for $x \geq 0.5$, and $H[x] = 0$ otherwise, and choose the death rate $\delta_i = 0.1$ for all species. To construct $W(i, j)$ as an interaction matrix that quantifies the cooperation and competition between species, we follow a set up similar to a Hopfield model with $\{0, 1\}$ activity[118]. We first choose a set of random points in N-dimensional species space $\mathbf{s}_q$ for $q \in \{1, \cdots, Q\}$, denoting *potential* niches. We choose Q such that $1 \ll Q \ll N$. Each $\mathbf{s}_q$ vectors consists of a $+1$ at elements corresponding to species that may co-exist, and -1 otherwise. In practice, we draw each element uniformly from the set $\{0, 1\}$, constructing an $N \times Q$

$W^f_{\text{localized}}$ parameters	Value
α_S	4
d_{loc}	$\begin{cases} 84 & \text{in 1D} \\ 50 & \text{in 2D} \end{cases}$
ϵ_S	$\begin{cases} 4.77 & \text{in 1D} \\ 1.6 & \text{in 2D} \end{cases}$

W^f_{diffuse} parameters	Value
α_{dif}	-0.25
d_{dif}	135

W^f_{decaying} parameters	Value
α_T	25
d_{dec}	150

TABLE III. Fixed-scale kernel parameters used for numerical simulations

matrix. The weight matrix $W(i,j)$ is then constructed as

$$W(i,j) = 0.015 \times \begin{cases} \sum_{i=1}^{Q} s_q(i) s_q(j), & \text{for } i \neq j \\ c, & \text{for } i = j \end{cases},$$

where c is a positive constant set to 10.

The spatial interaction kernel $K(x,x') = K(x-x')$ is chosen to be a Gaussian function with standard deviation 1.75 (which is much smaller than the entire spatial extent of the system, L). The end points of the resource gradient are chosen as two random $M = 2000$ dimensional vectors with elements draw independently from i.i.d. Gaussian distributions with zero mean and standard deviation $2/N$, and the preference vectors $\mathbf{b}_i$ are drawn from i.i.d. Gaussian distributions with zero mean and unit standard deviation.

The initial condition for the simulation is set to be the uniform state $s(i) = 0.5$ for all i, and the simulation is run until the dynamics reach a fixed point state. The final formed fixed point state is examined by calculating the correlation matrix

$$C(x,x') = \sum_{i=1}^{N} \cap s^*(i,x) \cap s^*(i,x')$$

where

$$\cap s(i) = [s(i) - \langle s \rangle] / \sqrt{\sum_i [s(i) - \langle s \rangle]^2}$$

where $\langle s \rangle = (1/N) \sum_j s(j)$.

SUPPLEMENTARY TEXT

The supplemental information is structured as follows: First, in SI Sec. A we present the mathematical analysis for pattern formation, and generalize the theory of CAN models of grid cells to show analytically and numerically that an infinite set of local interaction kernels can generate a grid cell network, as shown in Fig. 2 and Fig. 10. Second, we demonstrate analytically and numerically in SI Sec. B that simply introducing a gradient in the pattern forming kernel of the continuous attractor model is *not* sufficient to result in modularization, as demonstrated in

Fig. 1 of the main text. Third, in Sec. B 1 we show how the addition of a Gaussian localized kernel results in self-organized modularization. Fourth, we show in Sec. D that among arbitrary kernels, those with simple shapes result in a simple equation describing the detailed period ratios of the formed grid modules as shown in Fig. 4. Fifth, this will lead to simple estimates for the number of modules and their sizes in terms of other system parameters, which we derive in SI Sec. D 7. Sixth, after having described our results primarily for the case of one-dimensional grid cells, we then demonstrate in Sec. D 8 that our arguments extend naturally to two dimensions, and we present numerical results demonstrating the same. Seventh, in SI Sec. E 1 we then demonstrate that our results and predictions of grid period ratios are consistent with available data sources to a large extent. Finally, we generalize our result to the context of dynamics on a rough energy landscape (SI Sec. F), and provide broader perspectives of our results in the contexts of general loss optimization (Sec. G) and eigenvector localization (SI Sec. H).

Appendix A: Generalization of grid cell CAN dynamics theory: infinite set of interactions produce grid cells

It is known that Mexican hat-like kernels [24] and Lincoln hat-style kernels [51] generate grid patterning. While there are analytical results on why grid patterning emerges from a Mexican hat interaction, the Lincoln hat result is empirical, without theory. Here we seek to explain when grid patterning emerges, and to determine other kernel shapes that are consistent with it.

Consider the standard equations for the dynamics of recurrently connected neurons (expressed for notational simplicity in the continuum or large neural number limit):

$$\frac{\partial s(\mathbf{x},t)}{\partial t} + \frac{s(\mathbf{x},t)}{\tau} = \phi\left[\int_{-\infty}^{+\infty} W(\mathbf{x},\mathbf{x}')s(\mathbf{x}',t)d\mathbf{x}' + B\right], \tag{A1}$$

where $s(\mathbf{x})$ is the synaptic activation of the neuron at the vector position $\mathbf{x}$ on a 2-dimensional neural sheet, $W(\mathbf{x},\mathbf{x}')$ is coupling strength from a neuron at $\mathbf{x}'$ to a neuron at $\mathbf{x}$, τ is the biophysical time-constant of individual neurons, ϕ is a non-negative monotonic transfer function, and B is a uniform feedforward input to all neurons. The neural nonlinearity is any non-odd ($\phi(-x) \neq \phi(-x)$ function, for reasons given below and in [119–122]). For simplicity, we select the rectification function ($\phi(z) = [z]_+ = z$ for $z > 0$ and 0 otherwise.

To obtain conditions for hexagonal pattern formation dynamics, we perform a linear stability analysis about the homogenous steady state[123]. While we will provide the details of this analysis in the more general setting with gradients in kernel widths in Sec. B, we summarize the key result here: for an interaction kernel $W(\mathbf{x},\mathbf{x}') = W(|\mathbf{x}-\mathbf{x}'|)$, the growth rate $\alpha(k)$ of a periodic mode with wave vector k is related to the Fourier transform $\tilde{W}$ as $\alpha(k) = \tilde{W}(k) - 1/\tau$.

From this linear (in)stability analysis, four conditions on the interaction kernel weights $W(\mathbf{x},\mathbf{x}')$ may be sufficient for grid-like patterning: 1) For global stability, let $\int W(\mathbf{x},\mathbf{x}')d\mathbf{x} < 0$ (this is consistent with models of grid cells with negative recurrent coupling [24, 117] and with experiments suggesting that grid pattern formation might dominated by recurrent inhibitory circuitry [117]). 2) Let the interactions be radially and translationally symmetric, $W(\mathbf{x},\mathbf{x}') \equiv W(|\mathbf{x}-\mathbf{x}'|;\sigma)$, which means that the Fourier transform can be written in terms of its radial part: $\tilde{W}(\mathbf{k}) = \tilde{W}(|\mathbf{k}|) = \tilde{W}(k)$. 3-4) To ensure a non-zero wavelength k of pattern emergence, the Fourier transform of W should satisfy that its maximum occurs at a non-zero value of k, $k^* = \arg\max \tilde{W}(k) > 0$, and that this maximum should be positive and sufficiently large, $\tilde{W}(k^*) > 1/\tau$. Note that conditions 3-4) can be easily made true so long as W is not everywhere negative, and we are permitted a global scaling factor to ensure that the positive component is sufficiently large.

The emergent activity pattern will consist of superpositions of waves with period $2\pi/k^*$ [14, 24, 120, 123–125]. This period scales as σ, the characteristic width of the interaction kernel W. The specific geometry of the emergent period-$2\pi/k^*$ pattern depends on the relative strengths and

interactions of the waves of wavenumber k^*. If the interaction kernel is isotropic and the boundary conditions are infinite or isotropic, the formed pattern will be an equally-weighted superposition of all three waves of wavenumber k^*, defining a triangular lattice. The phase of the formed pattern will be set by spontaneous symmetry breaking.

The non-odd nature of the transfer function results in patterns with *hexagonal* rather than other symmetries [119–122]. This can be intuitively seen by noting that upon addition of plane wave solutions $\mathrm{Re}[\sum_n e^{i\mathbf{k}_n\cdot\mathbf{x}}]$, hexagonal patterning formed by adding three plane waves such that $\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3=0$ is the smallest summation that breaks inversion symmetry (i.e., the maxima are more positive than the minima are negative). Thus, hexagonal patterns emerge when the transfer function provides inversion symmetry breaking.

How many kernel functions W satisfy these conditions? Essentially, an infinite set does so (with rare exceptions). First we discuss some of the exceptions to gain some insight. Gaussian and Lorentzian functions, when they are positive, have a single peak in their Fourier transforms at $k=0$ when the functions are positive. When the functions are negative everywhere, they fail to satisfy condition 1). Thus, Gaussian and Lorentzian functions are two special functions that do not satisfy the criteria 1)-4). However, as argued in Sec. C, making small perturbations to functions that do not satisfy 1)-4) results in the conditions 1)-4) being satisfied, suggesting that the functions that do not satisfy 1)-4) are a small and very special set, and that most functions can be scaled to satisfy 1)-4).

We next performed numerical experiments to test the hypothesis that randomly generated functions will generically have Fourier Transforms that are not negative everywhere or only non-negative at 0, and therefore might generate grid-like patterning (see Methods for details of random sampling of kernel functions). We found that indeed randomly constructed kernel functions satisfied the hypothesized property for their Fourier transforms: we generated 10^6 random localized kernel functions, and all of these satisfied the conditions of being not negative everywhere or being non-negative only at $k=0$ (SI Fig.10). We further found that these kernel functions, under the further condition that they did not produce diverging neural activity, generated hexagonal patterns. Some of these are shown in Fig. 2a. In sum, an infinite set of local interaction profiles will generate grid cell-like activation patterns. Such candidate profiles can be generated at random and with very high probability generate grid-like patterning.

Appendix B: Pattern formation with graded kernels

Motivated by the experimental observations described in the main text, we modify the Mexican-hat function to introduce a smooth gradient in the characteristic interaction widths σ_E, σ_I.

$$W^g_{n_{DV}}(\mathbf{\Delta x}) = \alpha_E \exp\left(-\frac{\mathbf{\Delta x}^2}{2\sigma_E(n_{DV})}\right) - \alpha_I \exp\left(-\frac{\mathbf{\Delta x}^2}{2\sigma_I(n_{DV})}\right), \tag{B1}$$

where $\sigma_E(n_{DV})$ and $\sigma_I(n_{DV})$ are now functions that depend on position in the neural sheet, and encode the smoothly varying characteristic scale of the Mexican-hat interaction along the dorso-ventral axis:

$$\sigma_{E/I}(n_{DV}) = \sigma_{E/I} + \sigma'_{E/I}(0)\cdot n_{DV}. \tag{B2}$$

For such graded kernels, we will use $W(\mathbf{x},\mathbf{x}')$ and $W_{n_{DV}}(\mathbf{x}-\mathbf{x}') = W_{n_{DV}}(\mathbf{\Delta x})$ interchangeably. In this case, Eq. A1 then becomes

$$\frac{\partial s(\mathbf{x},t)}{\partial t} + \frac{s(\mathbf{x},t)}{\tau} = \phi\left[\int_{-\infty}^{+\infty} W_{n_{DV}}(\mathbf{x}-\mathbf{x}')s(\mathbf{x}',t)d\mathbf{x}' + B(\mathbf{x})\right], \tag{B3}$$

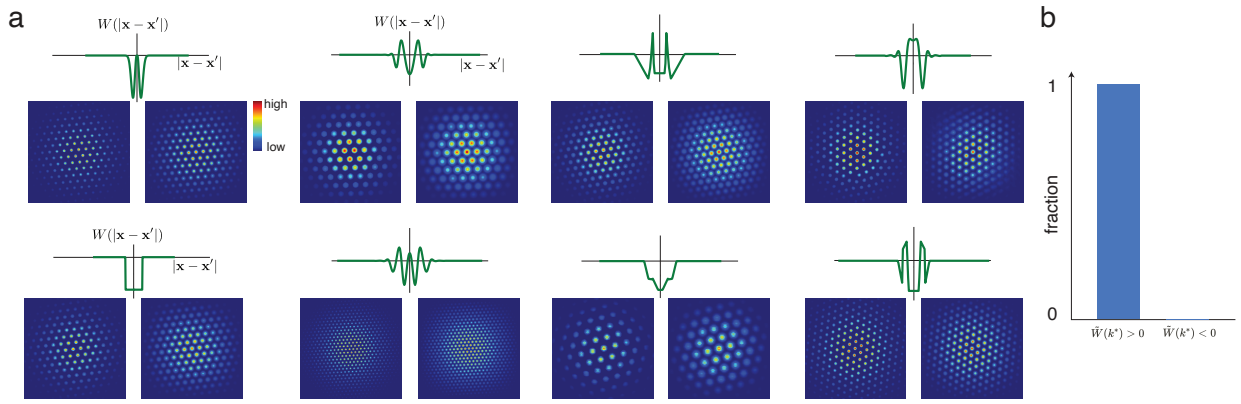


FIG. 10. All kernels satisfying the conditions laid out in the main text can result in pattern formation, with appropriate scaling.

Under this approximation, we perform a linear stability analysis of the neural dynamics, to identify the growing periodic modes locally at the position on the neural sheet n_{DV}.

We first identify an unstable steady-state solution to Eq. (B3), which we denote as $s_0(\mathbf{x})$. This solution satisfies

$$\frac{s_0(n_{DV})}{\tau} = \phi\left[\int_{-\infty}^{+\infty} W_{n_{DV}}(\mathbf{x}-\mathbf{x}')s_0(\mathbf{x}')d\mathbf{x}' + B(\mathbf{x})\right]. \tag{B4}$$

In the limit of very slowly varying changes in $W_{n_{DV}}(\mathbf{\Delta x})$ as a function of n_{DV}, the unstable steady state solution will be

$$s_0(n_{DV}) = \frac{\tau\bar{B}}{1-\tau\bar{W}}, \tag{B5}$$

where $\bar{B} = \int B(\mathbf{x})d\mathbf{x}$ and $\bar{W} = \int W_{n_{DV}}(\mathbf{x}-\mathbf{x}')d\mathbf{x}'$. (For $\tau\bar{W}_{\mathbf{x}} > 1$, the only locally homogeneous steady state is $s_0(n_{DV}) = 0$ due to the rectifying nonlinearity, which as we justify shortly cannot support periodic pattern formation due to being a stable fixed point).

We then consider a perturbative analysis, by examining the evolution of $s(\mathbf{x},t) = s_0(n_{DV}) + \epsilon(\mathbf{x},t)$. We apply our analysis to the early time evolution of this initial condition, such that $\epsilon(\mathbf{x},t) \ll s_0(n_{DV})$. Inserting our form of $s(\mathbf{x},t)$ in Eq. (B3), we obtain

$$\frac{\partial\epsilon(\mathbf{x},t)}{\partial t} + \frac{\epsilon(\mathbf{x},t)}{\tau} = \phi'(\bar{W}s_0(n_{DV}) + \bar{B})\int_{-\infty}^{\infty} W_{n_{DV}}(\mathbf{x}-\mathbf{x}')\epsilon(\mathbf{x}',t)d\mathbf{x}'. \tag{B6}$$

Since $W_{n_{DV}}(\mathbf{x}-\mathbf{x}')$ is a local kernel, we approximate the above integral with one evaluated over the region $\{\mathbf{x}' : |\mathbf{x}-\mathbf{x}'| < l\}$, with l much larger than the length-scale of the kernel $W_{n_{DV}}$ at all $\mathbf{x}$. Over this interval, we posit that $\epsilon(\mathbf{x}',t) = \epsilon e^{i\mathbf{k}\cdot\mathbf{x}'+\alpha(\mathbf{k})t}$, where $\alpha(\mathbf{k})$ denotes the growth rate of this ϵ perturbation. Inserting this form into Eq. (B6) yields,

$$\alpha(\mathbf{k}) + 1/\tau = \phi'(\bar{W}s_0(n_{DV}) + \bar{B})\int_{\infty}^{\infty} W_{n_{DV}}(\mathbf{x}-\mathbf{x}')e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}d\mathbf{x}', \tag{B7}$$

$$= \phi'[\bar{W}s_0(n_{DV}) + \bar{B}]\mathcal{F}[W_{n_{DV}}(\mathbf{x}-\mathbf{x}')], \tag{B8}$$

$$= \phi'[\bar{W}s_0(n_{DV}) + \bar{B}]\mathcal{F}W_{n_{DV}}(\mathbf{k}) \tag{B9}$$

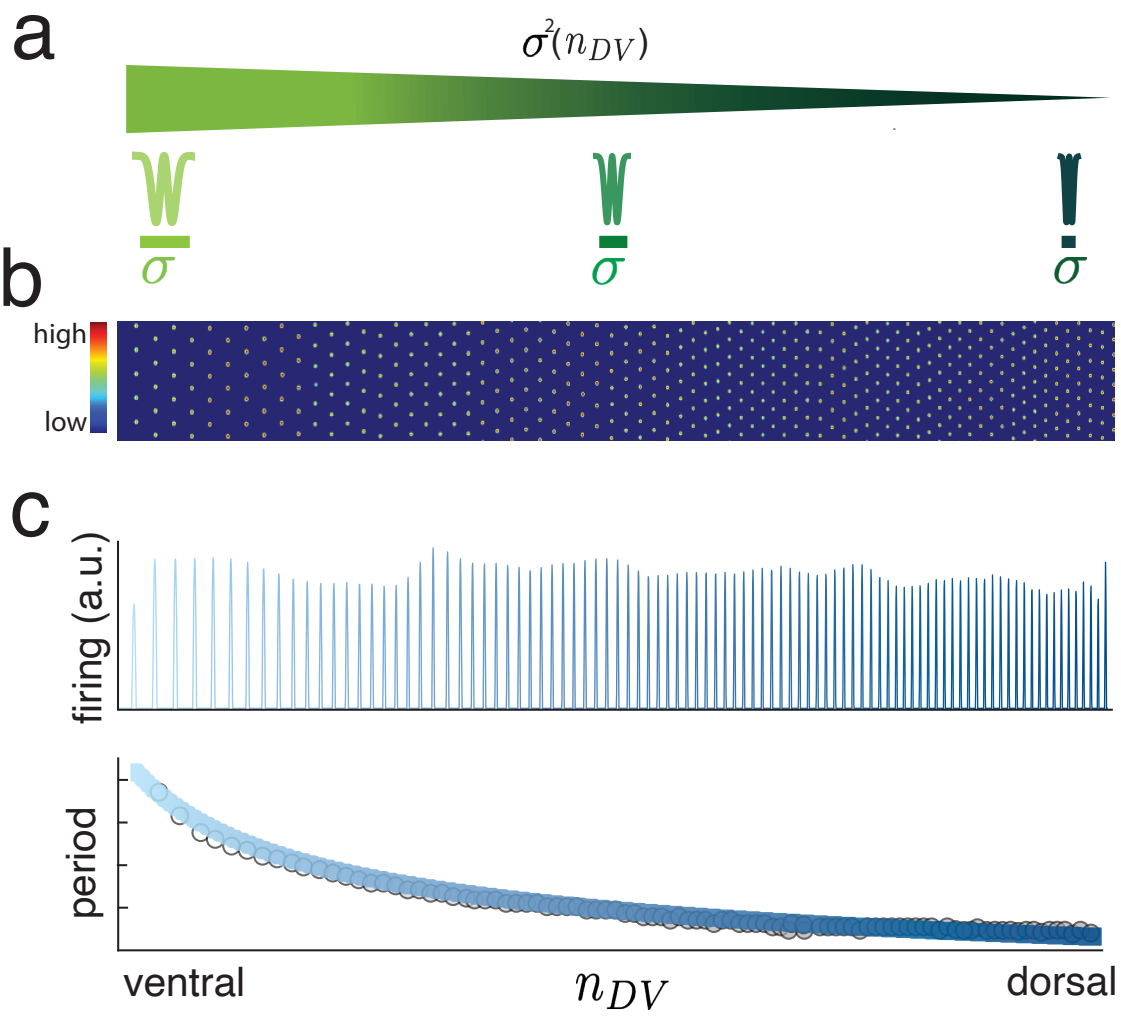


FIG. 11. (a-c) Naive merger of the two mechanisms by smoothly scaling the width of the pattern-forming lateral interaction (j) in the grid cell CAN model [24] does not generate global modularity in 2-dimensional (b) or 1-dimensional (c) grid models: the result is one smoothly varying periodic pattern.

where $\mathcal{F}[W_{n_{DV}}(\mathbf{x}-\mathbf{x}')] = \mathcal{F}W_{n_{DV}}(\mathbf{k})$ is the Fourier transform of the interaction kernel corresponding to position n_{DV} on the neural sheet. For the rectifying nonlinearity $\phi' = 1$, and the requirement for the periodic perturbation to be growing is $\alpha(\mathbf{k}) = \mathcal{F}W_{n_{DV}}(\mathbf{k}) - 1/\tau > 0$.

Note that since $W_{n_{DV}}(\mathbf{\Delta x})$ is a kernel, it is a radially-symmetric real function, and hence the Fourier transform $\mathcal{F}W_{n_{DV}}(\mathbf{k})$ will also be real function that is radially-symmetric in k. Thus, for simplicity, we will only focus on the magnitude of $\mathbf{k}$, which we denote as $k = |\mathbf{k}| \geq 0$ (In this context, for the two-dimensional case, one may re-interpret the radial component of the Fourier transform of $W_{n_{DV}}(\mathbf{\Delta x})$ as the Hankel transform of $W_{n_{DV}}(|\mathbf{\Delta x}|)$).

By definition, the magnitude of the wave vector k^* that corresponds to the fastest growing mode locally around position $\mathbf{x}$ on the neural sheet will be the $\mathbf{k}$ that maximizes $\alpha(\mathbf{k})$. Under the approximation of slow changes in the length-scale of the interaction kernel $W_{n_{DV}}(\mathbf{\Delta x})$, we see from Eq. (B9) that

$$k^*(n_{DV}) = \arg\max_k \mathcal{F}W_{n_{DV}}(k), \tag{B10}$$

since $W_{n_{DV}}(\mathbf{\Delta x})$ (and hence $s_0(n_{DV})$) has been assumed to have a negligible dependence on n_{DV}.

For $W_{n_{DV}}(\mathbf{\Delta x})$ given by Eq. (B1), i.e., without any additional fixed-scale interaction, we obtain from Eq. (B10)

$$[k^*(n_{DV})]^2 = \frac{2}{\sigma_E(n_{DV})^2 - \sigma_I(n_{DV})^2} \log\left(\frac{\alpha_E \sigma_E(n_{DV})^3}{\alpha_I \sigma_I(n_{DV})^3}\right). \tag{B11}$$

If we assume that $\sigma_{E/I}(\mathbf{x}) = \eta_{E/I}\sigma(n_{DV})$, where η_E and η_I are $\mathbf{x}$-independent constants, then we obtain

$$k^*(n_{DV}) \propto 1/\sigma(n_{DV}), \tag{B12}$$

and hence

$$\lambda^*(n_{DV}) \propto \sigma(n_{DV}), \tag{B13}$$

where $\lambda^*(n_{DV})$ is the periodicity of the grid pattern formed locally around position n_{DV}. This results in a smooth change of grid period, corresponding to the observation in Fig. 1g of the main text.

Note that this result is generally true for any pattern forming kernel $W^g_{n_{DV}}(\mathbf{\Delta x})$ that has a Fourier transform with at least one local maximum, and *does not* rely on the specific form of a Mexican-hat interaction. Indeed, Eq. (B13) holds for any kernel $W^g_{n_{DV}}(\Delta x)$ that depends on a length-scale $\sigma(n_{DV})$. As an example, we present the corresponding analysis for the box-shaped kernel employed for pattern formation in Ref. [117].

In this case

$$W^g_{n_{DV}}(\mathbf{\Delta x}) = -W_0 \mathbb{1}_{\mathbf{\Delta x} \leq \sigma(n_{DV})}. \tag{B14}$$

As discussed above, the quantity of interest is $\mathcal{F}W^g_{n_{DV}}(\mathbf{k})$

$$\mathcal{F}W^g_{n_{DV}}(\mathbf{k}) = \int_{-\infty}^{\infty} -W_0 \mathbb{1}_{|\mathbf{x}| \leq \sigma(n_{DV})} e^{i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x} \tag{B15}$$

$$= -W_0 \int_{|\mathbf{x}| \leq \sigma(n_{DV})} e^{i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x}. \tag{B16}$$

The above integral can be calculated in a one-dimensional setup to obtain

$$\mathcal{F}W^g_{n_{DV}}(k) = -2W_0 \frac{\sin(k\sigma(n_{DV}))}{k} \tag{B17}$$

and can be calculated in a two-dimensional setup to obtain

$$\mathcal{F}W^g_{n_{DV}}(k) = -2\pi W_0 \sigma(n_{DV}) \frac{J_1(k\sigma(n_{DV}))}{k}. \tag{B18}$$

In both of the above cases, note that $k^* \propto 1/\sigma(n_{DV})$ since $\sigma(n_{DV})$ is the only length-scale characterizing the kernel $W^g_{n_{DV}}$. In particular, numerical maximization yields

$$k^* \approx \begin{cases} 4.493/\sigma(n_{DV}) & \text{on a one-dimensional sheet, and} \\ 5.136/\sigma(n_{DV}) & \text{on a two-dimensional sheet.} \end{cases} \tag{B19}$$

1. Fixed-scale interactions and modularization

We now claim that the addition of a fixed-scale kernel, $W^f(\mathbf{\Delta x})$ is sufficient to result in modularization of grid periods, with discrete changes in grid period as a function of spatial position along the dorso-ventral axis. This set of interactions can effectively be implemented by two populations of interneurons - one with fixed arborization and weaker synaptic connections and one with varying arborization length and stronger synaptic connections.

For simplicity, we shall present the specific Fourier transform computations for the one-dimensional problem, although we note that all of the qualitative results hold in two dimensions as well, with the Fourier transforms of the relevant functions replaced with their Hankel transforms (as shown in Sec. D 8).

We include an additional weak interaction term W^f that critically does *not* depend on the neural sheet position x. For reasons that will become apparent soon, we choose kernels $W^f(\Delta x)$ such that the Fourier transform changes sign a sufficiently large number of times. We hypothesize that this requirement is not particularly restrictive, and will demonstrate that this holds for most kernels W^f.

The entire interaction profile is then given by

$$W_{n_{DV}}(\Delta x) = W^g_{n_{DV}}(\Delta x) + W^f(\Delta x). \tag{B20}$$

We first demonstrate our result with an example of a simple kernel, to justify how Eq. (B10) leads to the emergence of discrete grid modules. Consider the localized excitatory interaction

$$W^f(\Delta x) = \alpha_S \exp\left(-\frac{(\Delta x - d)^2}{2\epsilon_S^2}\right) + \alpha_S \exp\left(-\frac{(\Delta x + d)^2}{2\epsilon_S^2}\right). \tag{B21}$$

Corresponding to our interpretation of $W^f(\Delta x)$ above being a localized kernel, we choose $\epsilon_S \ll d$.

This choice of $W_{n_{DV}}(\Delta x) = W^g_{n_{DV}}(\Delta x) + W^f(\Delta x)$ leads to the the Fourier transform,

$$\begin{aligned}\mathcal{F}W_{n_{DV}}(k) &= \mathcal{F}W^g_{n_{DV}}(k) + \mathcal{F}W^f(k), &\text{(B22)}\\ &= \sqrt{2\pi}\left[\alpha_E\sigma_E(n_{DV})\exp\left(-\frac{\sigma_E(n_{DV})^2k^2}{2}\right)\right.\\ &\quad - \alpha_I\sigma_I(n_{DV})\exp\left(-\frac{\sigma_I(n_{DV})^2k^2}{2}\right)\\ &\quad \left.+2\alpha_S\epsilon_S\cos(kd)\exp\left(-\frac{\epsilon_S^2k^2}{2}\right)\right]. &\text{(B23)}\end{aligned}$$

In our model, the magnitude of the $W^f(\Delta x)$, i.e., α_S, is chosen to be smaller than the magnitude of the Mexican-hat interaction. Thus we interpret $\mathcal{F}W^f(k)$ in Eq. (B23) as being a small perturbation to the Fourier transform of the usual Mexican-hat interaction, $\mathcal{F}W^g_{n_{DV}}(k)$. Further, since d is assumed to be much larger than the scale of the Mexican-hat, $\sigma_{E/I}$, then the term $\cos(kd)$ in $\mathcal{F}W^f(k)$ oscillates at a k-scale much smaller than the relevant scales of $\mathcal{F}W^g_{n_{DV}}(k)$ (see Fig. 3b-c of the main text). Additionally, since $\epsilon_S \ll d$, the gaussian envelope multiplying the rapidly oscillating term has a scale $1/\epsilon$, which is much larger than the periodicity $1/d$.

Thus, in k-space, the rapidly oscillating term, $\mathcal{F}W^f(k)$ can be thought of as predefining a set $S = \{k_1, k_2, \ldots\}$ of local maxima. Under the approximations made above, the addition of the smoother function $\mathcal{F}W^g_{n_{DV}}(k)$, will not change the position of the local maxima. This results in the *local* maxima of $\mathcal{F}W_{n_{DV}}(k)$ also being the same set S. Importantly, we note that since S was predefined purely via $\mathcal{F}W^f(k)$, *there is no* n_{DV} *dependence on the set* S.

Following Eq. (B10), the wave-vector corresponding to the pattern formation at point x on the neural sheet corresponds to the *global* maxima of $\mathcal{F}W_{n_{DV}}(k)$. Thus, at all points, the pattern formation corresponds to one of the discrete set of choices of wave vectors, $S = \{k_1, k_2 \ldots\}$. As can be seen from Fig. 3c, the smoothly varying gradient in the Mexican-hat term, $\mathcal{F}W^g_{n_{DV}}$ as a function of x picks different choices of k_i depending on the position n_{DV} — the $k \in S$ that is nearest to the maxima of $\mathcal{F}W^g_{n_{DV}}(k)$ will be chosen as the global maxima, and will be the wave vector corresponding to the pattern at n_{DV}. We refer to this mechanism as "peak selection".

For our particular choice of $W^f(x)$ made in Eq. (B21), we obtained

$$\mathcal{F}W^f(k) = 2\alpha_S \epsilon_S \cos(kd) \exp\left(-\frac{\epsilon_S^2 k^2}{2}\right). \tag{B24}$$

We can then approximate the local maxima of $\mathcal{F}W^f(k)$ as occurring at

$$S = \left\{ \frac{2m\pi}{d} \middle| m \in \mathbb{Z}^+ \right\}. \tag{B25}$$

This immediately indicates that the ratios of periods of successive grid modules will be given by

$$\frac{\lambda_{m+1}}{\lambda_m} = \frac{m+1}{m}. \tag{B26}$$

Thus, the addition of a fixed-scale interaction, W^f such as Eq. (B21) results in discrete grid modules. We now show that this peak-selection mechanism, and hence modularization, occurs for arbitrary choices of the fixed-scale interaction kernel $W^f(\Delta x)$.

Appendix C: Kernels that lead to modularization

The peak-selection modularization mechanism described above arises naturally from the presence of the rapidly oscillating term in $\mathcal{F}W^f(k)$. In fact, for discrete grid modules to occur, the only constraints imposed on the fixed-scale kernel W^f are: (a) the Fourier transform $\mathcal{F}W^f(k)$ must have a sufficiently large number of maxima (at least 4 maxima, corresponding to the 4 grid modules observed in experimental observations); and, (b) these maxima must be at scales smaller than $1/\sigma$ in k-space. Here we argue that this is generally true for arbitrary kernels, modulo a single scaling parameter.

We hypothesize and give support, without formal proof, that almost every arbitrarily chosen kernel $W^f(\Delta x)$ will have a Fourier transform with multiple maxima satisfying condition (a). We will then argue that this kernel can always be scaled to satisfy condition (b).

To motivate our hypothesis, we first note that it is actually possible to construct specific kernels $W^f(\Delta x)$ whose Fourier transform does not present multiple maxima. For example, the Gaussian kernel, $W_{gauss}(\Delta x) = \exp[-(\Delta x)^2/2]$, results in a Fourier transform that is unimodal. However, we hypothesize that such functions are rare in the space of all continuous functions in L^2. Indeed, we can construct a function that is arbitrarily close to the Gaussian kernel whose Fourier transform will have an infinite number of maxima: Let $f_0(\Delta x) = \mathbb{1}_{[-1,1]}$ be the box function. Define

$$f_n = f * f_{n-1}$$

for all $n \geq 1$, where $f * g$ represents the convolution of functions f and g. By the central limit theorem, $\sqrt{n} f_n(\sqrt{n}\Delta x)$ will approach $W_{gauss}(\Delta x)$. However,

$$\mathcal{F}f_n(k) = [2\sin(k)/k]^n, \tag{C1}$$

which clearly has an infinite number of maxima. Thus, even though the Gaussian kernel has a unimodal Fourier transform, we can construct a function $g_n(\Delta x) = \sqrt{n} f_n(\sqrt{n}\Delta x)$ that is arbitrarily close to the Gaussian kernel (for sufficiently large n) but has a Fourier transform that presents an infinite number of maxima.

In this context, we claim that almost every arbitrarily chosen kernel $W^f(\Delta x)$ will have a Fourier transform with multiple maxima. This may be intuited as follows: First note that Fourier space is a dual space, and hence instead of considering arbitrary kernels in real space we may equivalently choose arbitrary kernels in Fourier space. Further assuming that $\mathcal{F}W^f(k)$ is a smooth function, we hypothesize that generically smooth functions that are in L^2 will almost always have multiple maxima and minima. Note that this heuristic also applies to the pattern forming kernel as well — we hypothesize that generic L^2 smooth functions will have some maxima and minima with a global maxima that exists at $k > 0$ with probability 1, and will not be always negative (in which case a rescaling will make the maxima larger than the constant specified by requirement 2 for pattern forming kernels in the main text). Thus we expect that kernels will generically result in hexagonal pattern formation, as demonstrated in Fig. 10.

Thus condition (a) may be satisfied for arbitrary kernels $W^f(\Delta x)$.

Next, note that scaling a function in real space results in an inverse scaling of the Fourier transform, i.e., $\mathcal{F}[W^f(a\Delta x) = \mathcal{F}W^f(k/a)$. Hence, we can always scale the function $W^f(\Delta x)$ to obtain a Fourier transform with maxima that are within any desired scale, allowing condition (b) to be satisfied.

In Fig. 12, we show examples of modularization arising from different combinations of graded pattern forming kernels (W^g) and fixed-scale kernels (W^f). In each case, we also present the expected periodicity in each module as a function of spatial position as given by the perturbative analysis Eq. (B10). The analytical result based on linear stability provides an excellent prediction of the pattern periods per module (see also Main text, Fig. 3e). It also predicts the locations of the module boundaries (see also Main text, Fig. 3e) though not as accurately: module boundary predictions tend to be slightly but systematically offset relative to the simulated dynamics, due to the effects of nonlinearity in the later stages of pattern formation.

Appendix D: Simple kernels and period ratios

What kinds of fixed-scale interactions might be present in the medial-entorhinal cortex? As described in the main text, in the context of biology, we might expect *simple* interaction kernels W^f to be relevant i.e., the fixed-scale interaction profile W^f has the following characteristics: (a) there exists a *single* length-scale d that primarily characterizes the shape of W^f; (b) any other length-scales relevant to W^f, say scales $\epsilon_1, \epsilon_2, ...$ are each much smaller than the primary length scale d. Further, we assume that the primary length-scale associated with the fixed-scale interaction is larger than the length-scales of the pattern forming kernel, i.e., $d \gg \sigma_{E/I}(n_{DV})$.

We will demonstrate that *simple* fixed-scaled interaction kernels result in analytic expressions for grid periods that are characterized by a single angular variable ϕ

$$\frac{\lambda_{m+1}}{\lambda_m} = \frac{m + 1 + \phi/(2\pi)}{m + \phi/(2\pi)}. \tag{D1}$$

Before filling in the details of our argument, we present an intuitive explanation of the general idea:

Consider the following basic classes of *simple* kernels that satisfy the above-described criteria corresponding to a length-scale d:

(a) $g(|\Delta x| - d)$, for arbitrary functions $g(\rho)$ that are nonzero only over scales $|\rho| < \epsilon_i$ (a *localized* kernel), and,

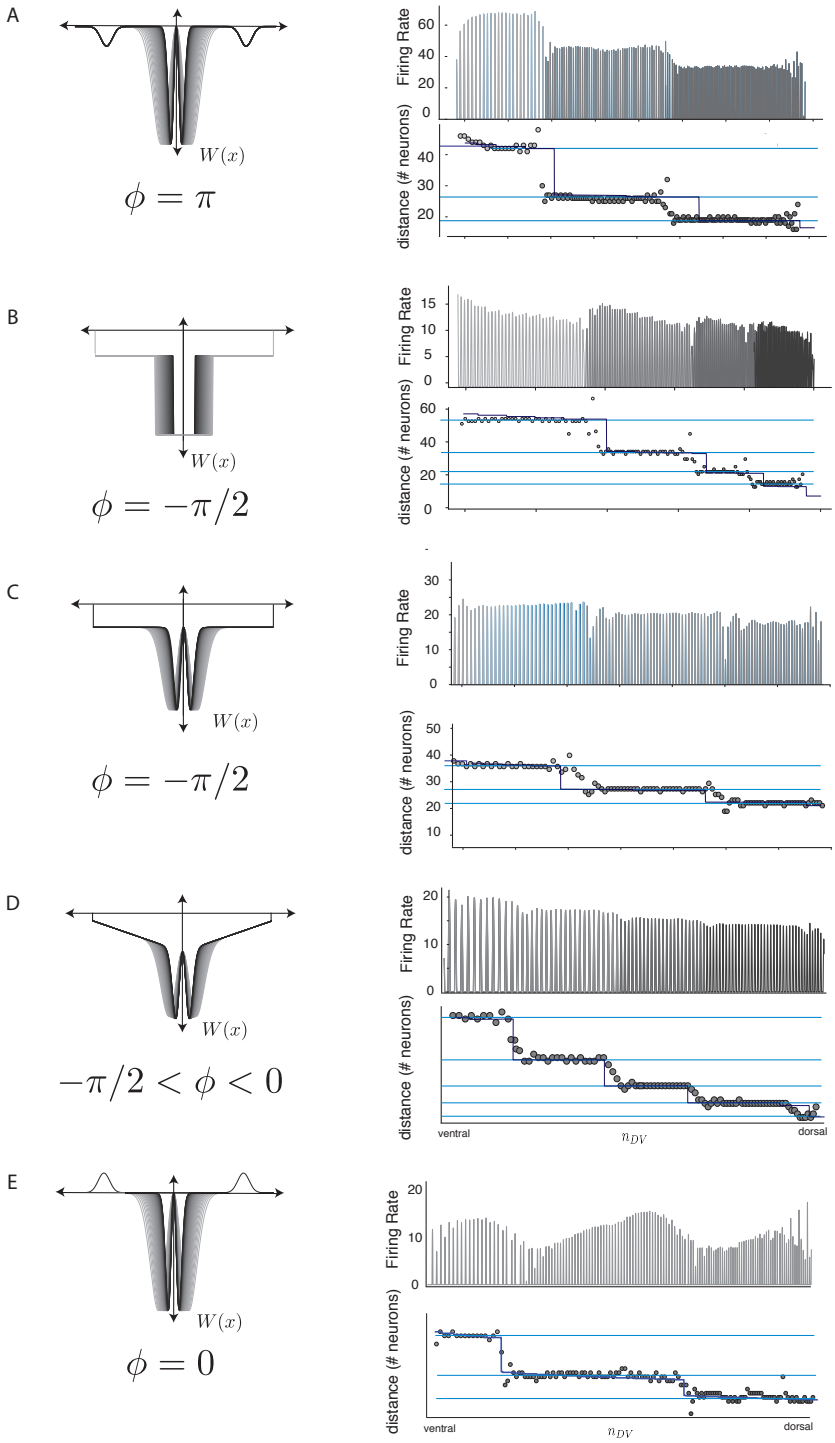


FIG. 12. Examples of modularization and population activity (right column) with various pattern forming and fixed-scale lateral interactions (left column). In each case the dark-blue curve shows the predicted value of the grid period from Eq. (B10), and is in close agreement with the numerical simulation of the population activity. Each of the fixed-scale interactions has a qualitatively different shape, spanning different values of ϕ (see Fig. 3)

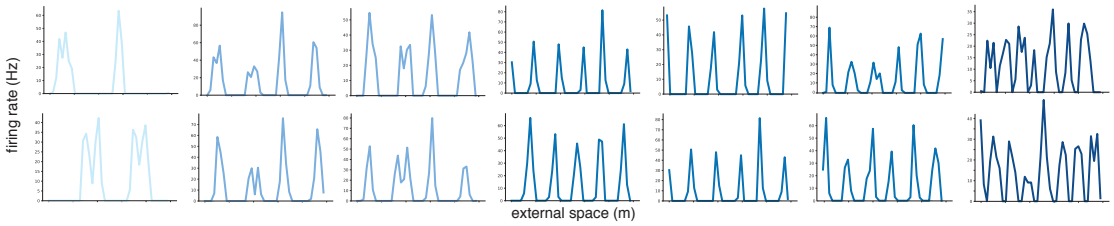


FIG. 13. Sample tuning curves from several neurons in all modules from the network of Fig 2a.

(b) A constant term, that is uniform everywhere up to $\Delta x = d$, after which it falls to zero (a *diffuse* kernel),

(c) A decaying term, that decreases from a constant value at $\Delta x = 0$ to zero at $\Delta x = d$ (a *decaying* kernel).

We also define *short-range* kernels, as any arbitrary function $h(\Delta x)$ that is nonzero only over scales $|x| < \epsilon_i$.

Any *simple* kernel $W^f(\Delta x)$ can be generally constructed as a linear combination of the above basic classes. In addition, *simple* kernels may also contain an added component of a *short-range* kernel.

To see that *simple* kernels will generally result in grid period ratios corresponding to Eq. (D1), we will examine the approximate Fourier transform structure for each component of the linear combination of *simple* kernels corresponding to a given length-scale d. We first demonstrate that each of the basic *simple* kernels will result in Fourier transforms that are sinusoidal functions with phase shifts and decaying envelopes and hence each basic *simple* kernel will satisfy Eq. (D1). We then show that short-range kernels present Fourier transforms that vary only at large scales, and can be ignored in our analyses of *simple* kernels. We then use these results to demonstrate that all *simple* kernels constructed as the above-described linear combination will have sinusoidal Fourier transforms and will satisfy Eq. (D1).

1. Localized kernels

For a general localized kernel $W^f(\Delta x) = g(|\Delta x| - d)$ we obtain

$$\mathcal{F}W^f(k) = \Re[e^{-ikd}\mathcal{F}g(k)]. \tag{D2}$$

Since $g(x)$ is supported over a scale ϵ, the Fourier transform $\mathcal{F}g(k)$ will only vary at scales $k \sim 1/\epsilon \gg 1/d$. Thus for $1/d \ll k \ll 1/\epsilon$, we can approximate Eq. (D2) as

$$\mathcal{F}W^f(k) = |\mathcal{F}g(k)| \cos\left(kd - \psi\right), \tag{D3}$$

where $\psi = \arg[\mathcal{F}g(k)]$. The local maxima of $\mathcal{F}W^f(k)$ will then occur at

$$S = \left\{ \frac{2m\pi + \psi}{d} \middle| m \in \mathbb{Z}^+ \right\}, \tag{D4}$$

resulting in period ratios described by

$$\frac{\lambda_{m+1}}{\lambda_m} = \frac{m + 1 + \psi/(2\pi)}{m + \psi/(2\pi)}, \tag{D5}$$

which is identical to Eq. (D1) for $\phi = \psi$. We also note that we can now ascribe an interpretation to the phase angle ϕ — it is the phase difference between $\mathcal{F}W^f(k)$ and $\cos(kd)$.

2. Diffuse kernels

We model a diffuse interaction kernel $W^f(n_{DV})$ as

$$W^f(n_{DV}) = -W_0 \mathbb{1}_{[-d,d]} = \begin{cases} -W_0 & \text{if } |n_{DV}| \leq d \\ 0 & \text{if } |n_{DV}| > d \end{cases}. \tag{D6}$$

Corresponding to the discussion above, we look at the Fourier transform $\mathcal{F}W^f(k)$

$$\mathcal{F}W^f(k) = \int_{-\infty}^{+\infty} -W_0 \mathbb{1}_{[-d,d]} e^{ikx} dx = \int_{-d}^{+d} -W_0 e^{ikx} dx \tag{D7}$$

$$= -2W_0 \frac{\sin(kd)}{k} = -2W_0 d \ \mathrm{sinc}(kd). \tag{D8}$$

Note that once again, similar to Eqn. (B24), we obtain a functional form consisting of a periodic function ($\sin(kd)$) that is multiplied by a decaying envelope $1/(kd)$. Ignoring the effects of the envelope function, the maxima of this function occur at

$$S \approx \left\{ \frac{2m\pi - \pi/2}{d} \middle| m \in \mathbb{Z}^+ \right\}, \tag{D9}$$

which immediately results in period ratios of the form

$$\frac{\lambda_{m+1}}{\lambda_m} \approx \frac{m + 1 - 1/4}{m - 1/4}, \tag{D10}$$

which corresponds to the result in Eq. (D1) for $\phi = \pi/2$.

More precisely, the extrema of $\mathcal{F}W^f(k)$ occur at $k_m d = q - 1/q - 2/3q^3 + O(q^{-5})$ where $q = \left(m + \frac{1}{2}\right)\pi$. Notably, the errors decay approximately as $1/(\pi m)$, and thus for modules generated corresponding to $m \gtrsim 2$ will result in period ratios that approximate Eq. (D1) closely.

3. Decaying kernels

Decaying kernels with a scale d may be modeled as any monotonically decreasing function that decays from some constant W_0 at $\Delta x = 0$, to zero, at $\Delta x = d$. For simplicity, we consider the simplest linear approximation to such a kernel, modeled as a triangular kernel. For additional subtleties in the treatment of other decaying kernels, see D 5 a The triangular kernel can be written as:

$$W^f(\Delta n_{DV}) = \begin{cases} W_0(\Delta n_{DV} - d)/d & \text{if } \Delta n_{DV} < d \\ 0 & \text{if } \Delta n_{DV} \geq d \end{cases} \tag{D11}$$

This function can be written as the convolution of 2 diffuse box functions:

$$W^f(\Delta n_{DV}) = (-W_0 \mathbb{1}_{[-d/2,d/2]}) * (W_0 \mathbb{1}_{[-d/2,d/2]}).$$

Thus, its Fourier transform is:

$$\mathcal{F}W^f(k) = -W_0^2 d^2 \left(\frac{\sin(kd/2)}{(kd/2)} \right)^2$$

$$= -\frac{2W_0^2}{k^2}[1 - \cos(kd)].$$

Once again, we obtain a simple trigonometric function, with maxima at

$$S \approx \left\{ \frac{2m\pi}{d} \middle| m \in \mathbb{Z}^+ \right\}, \tag{D12}$$

which immediately results in period ratios of the form

$$\frac{\lambda_{m+1}}{\lambda_m} \approx \frac{m+1}{m}, \tag{D13}$$

which corresponds to the result in Eq. (D1) for $\phi = 0$.

4. Short-range kernels

For the case of a short-range kernel $W^f(\Delta x)$ that extends upto a scale ϵ, we note from the Fourier uncertainty principle that the characteristic k-scales of $\mathcal{F}W^f(k)$ will $\sim 1/\epsilon \gg 1/d$. Thus, unlike the three other types of simple kernels discussed above, short range kernels do not have structure at the scale of $1/d$. Since all relevant scales are much larger than $1/d$, adding short range kernels to any of the other types of *simple* kernels will *not* change the structure of local maxima at scales of $1/d$.

5. Arbitrary simple kernels

We now consider a general form for *simple* kernels, by constructing linear combinations of the above described three basic classes of *simple* kernels each corresponding to the same length scale d and additional short-range kernels.

$$W^f = a_{\text{local}} W^f_{\text{local}} + a_{\text{diffuse}} W^f_{\text{diffuse}} + a_{\text{decaying}} W^f_{\text{decaying}} + a_{\text{short}} W^f_{\text{short}}. \tag{D14}$$

As demonstrated in the preceding sections, the Fourier transform $\mathcal{F}W^f(k)$ will be given as

$$\mathcal{F}W^f(k) = a_{\text{local}} |\mathcal{F}g(k)| \cos(kd - \psi) - 2W_0 a_{\text{diffuse}} \sin(kd)/k - 2W_0^2 a_{\text{decaying}} (1 - \cos(kd))/k + \mathcal{F}h(k) \tag{D15}$$

$$= H_0(k) + \sum_{i=0}^{3} H_i(k) \cos(kd + \phi_i) \tag{D16}$$

for some constants ϕ_i, and some envelope functions $H_i(k)$ for $i = 0, 1, 2, 3$ that are slowly varying for $kd \gtrsim \mathcal{O}(1)$. Under this approximation, $\mathcal{F}W^f(k)$ is simply the sum of multiple sinusoidal waves with different phases and identical frequencies. Thus,

$$\mathcal{F}W^f(k) \approx \cos(kd - \phi) \tag{D17}$$

for some ϕ and $kd \gtrsim \mathcal{O}(1)$. Hence, the maxima of $\mathcal{F}W^f(k)$ occur at

$$S \approx \left\{ \frac{2n\pi + \phi}{d} \middle| n \in \mathbb{Z}^+ \right\}, \tag{D18}$$

which immediately results in period ratios of the form Eq. (D1). Note that the approximations made above imply that there may be deviations from our results for the maxima corresponding to small k values — this may manifest as deviations in the largest period grid module away from Eq. D1.

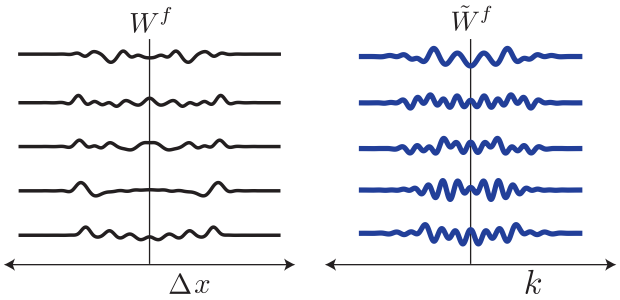


FIG. 14. Randomly constructed fixed-scale interactions (left column) and their Fourier transforms (right column), in addition to the hand-designed ones in Fig.3, that give $\phi = 0$.

a. *Caveats*

Clearly there exist *simple* kernels with Fourier transforms that are not given by $\mathcal{F}W^f(k) \approx \cos(kd - \phi)$. For example the Gaussian kernel, $W^f(\Delta x) = \exp[-\Delta x^2/(2d^2)]/(d\sqrt{2\pi})$ is a *simple* decaying kernel (since it has only a single scale d). Yet, its Fourier transform is simply $\mathcal{F}W^f(k) = \exp[-k^2d^2/2]$, which has only a single maximum! However, as we have shown earlier, there exist kernels that are arbitrarily close to the Gaussian kernel, whose Fourier transforms are given by powers of trigonometric functions, and hence have multiple regularly-spaced maxima with a spacing of $\sim 1/d$. Similarly, there exist additional *simple* functions[126–128], $f(\Delta x)$, (like the Gaussian kernel) whose Fourier transforms $\mathcal{F}f(k)$ have a small number of maxima. We hypothesize that for all such functions $f(\Delta x)$ there exist *simple* kernels $g(\Delta x)$ that are arbitrarily close to $f(\Delta x)$ and possess regularly spaced maxima.

6. Period ratios

Having demonstrated analytically that *simple* kernels result in a sequence of period ratios given by Eq. (D1), we now address the question of the mean period ratio over the sequence and over different values of ϕ. In the main text we have demonstrated that setting $\phi = 0$ results in a detailed period ratio sequence that is in close agreement with the sequence of experimentally observed values. Here we consider the period ratios obtained for other values of ϕ, to demonstrate that the experimental observation of mean period ratios being approximated by 1.4 [27] emerges naturally from our setup.

From Eq. (D1), we obtained that the period ratio, $r_m = \lambda_{m+1}/\lambda_m$ can be written as

$$r_m = 1 + 1/(m + f), \tag{D19}$$

where $f = \phi/(2\pi)$. We ignore $m = 1$, since that results in a period ratio close to 2, which does not correspond to experimental observations. Averaging the period ratio over the next 4 modules (corresponding to r_m for $m \in \{2 \dots 4\}$) results in

$$\langle r_m \rangle_m = 1 + \frac{1}{3}\left(\frac{1}{f+2} + \frac{1}{f+3} + \frac{1}{f+4}\right) \tag{D20}$$

As can be seen in Fig. 15, this mean period ratio lies in the range [1.3,1.45], indicating that at all values of ϕ, the period ratio obtained from Eq. (D1) matches well with experimental observations. The average of these period ratios over all values of ϕ can also be calculated as

$$\langle r_m \rangle_{\phi,m} = 1 + \frac{1}{3}\left[\log\left(\frac{5}{3}\right) + \log\left(\frac{7}{5}\right) + \log\left(\frac{9}{7}\right)\right] \tag{D21}$$

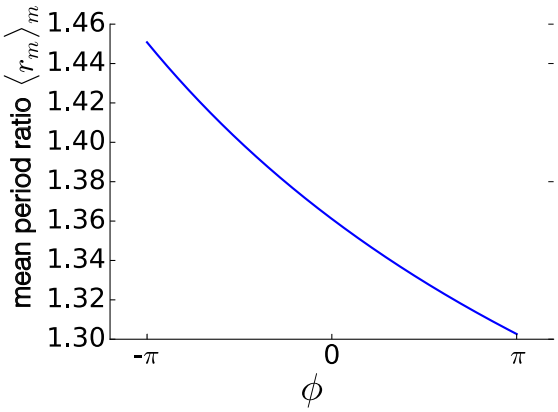


FIG. 15. **Mean grid-period ratios** Ratios of grid periods averaged over 4 modules as a function of the phase shift ϕ in Eq. (D1)

which is approximately equal to 1.37.

7. Module size; number of modules as a topological quantity

As discussed in the main text, peak-selection for modularization is a highly robust mechanism that is largely indifferent to system parameters such as the the particular forms of the fixed-scale interaction and the shape of the gradient. Here we provide an analysis of the number of modules, the scaling of module sizes, and the positions of module boundaries, which also exhibit the same robustness. Further, we also describe how this robustness may be interpreted as arising from a topological origin, similar to topological robustness in other physical systems like the quantum hall effect.

Recall that for the continuously graded kernel $W^g_{n_{DV}}(\Delta x)$ with characteristic spatial scale $\sigma(n_{DV})$ at position n_{DV}, the wave-vector of the formed pattern was proportional to $1/\sigma(n_{DV})$:

$$k^{*g}(n_{DV}) = \eta/\sigma(n_{DV}), \tag{D22}$$

where η is an n_{DV}-independent constant that depends on only the particular form of the graded kernel. Let the spatial extent of the system be $n_{DV} \in [0, L]$, with $\sigma(n_{DV})$ monotonic such that $\sigma_{\min} = \sigma(0) \leq \sigma(n_{DV}) \leq \sigma(L) = \sigma_{\max}$.

We assume for simplicity that the fixed-scale lateral interaction is a *simple* kernel, such that $\mathcal{F}W^f(k) \sim \cos(kd - \phi)$. Thus, the local maxima generated by $\mathcal{F}W^f(k)$ occur at $k_n \approx (2n\pi + \phi)/d$, where n are the natural numbers. As discussed in the main text, each of these local maxima is 'selected' in turn by the moving broad peak of the Fourier transform of the graded kernel, whose position according to Eq. D22 occurs at $k^{*g}(n_{DV}) = \eta/\sigma(n_{DV})$.

Notably, the selected maximum k_m will be robust to small perturbations in the selection function $\mathcal{F}W^g_{n_{DV}}(k)$, since k_m will remain quantized to one of the discrete values prespecified by the set $\{k_n \| n \in \mathbb{N}\}$. In this sense, the chosen maximum k_m (and hence the corresponding module) presents the hallmarks of a topologically protected state[1]. The topological number corresponding to a given module is the module number m, which is a topological invariant similar to a winding number[1](Note that in our convention the module number m is ordered such that the largest grid period module is the first module. This is opposite to the numbering usually used in the literature, such as in [27]).

The set of modules expressed through the length of the system corresponds to the set of local maxima k_n that lie within the range $[\eta/\sigma_{\max}, \eta/\sigma_{\min}]$ that is delineated by the range of peak

positions of the graded interaction. It follows that the maxima selected by the graded interaction obey:

$$\frac{\eta}{\sigma_{\max}} \leq \frac{2n\pi + \phi}{d} \leq \frac{\eta}{\sigma_{\min}}. \tag{D23}$$

Thus, the set of formed modules are determined by the set of integers n that fit in the following interval:

$$\frac{-\phi + \eta d/\sigma_{\max}}{2\pi} \leq n \leq \frac{-\phi + \eta d/\sigma_{\min}}{2\pi} \tag{D24}$$

and hence the number of modules N_{mod} is:

$$\begin{aligned} \#\text{ modules} &\equiv N_{mod} \\ &= \left\lfloor \frac{-\phi + \eta d/\sigma_{\min}}{2\pi} \right\rfloor - \left\lceil \frac{-\phi + \eta d/\sigma_{\max}}{2\pi} \right\rceil \\ &= \left\lfloor \frac{-\phi + k^{*g}(0)d}{2\pi} \right\rfloor - \left\lceil \frac{-\phi + k^{*g}(L)d}{2\pi} \right\rceil \end{aligned} \tag{D25}$$

where $\lfloor\ \rfloor$, $\lceil\ \rceil$ indicate the floor and ceiling operations, respectively.

The above result leads to the following observations: First, the central quantity essential for determining the number of modules is the difference in the integer ratios of the fixed-scale interaction width to the extremal lateral interaction widths, $d/\sigma_{\min}, d/\sigma_{\max}$. Second, the number of modules depends only on the end-point values $\sigma_{\min}, \sigma_{\max}$ of the smoothly varying width $\sigma(n_{DV})$ the graded interaction; notably, it does not depend on the detailed shape of $\sigma(n_{DV})$. Moreover, if $\sigma_{\min}, \sigma_{\max}$ are varied smoothly (while d is held fixed), or if d is varied smoothly (while $\sigma_{\min}, \sigma_{\max}$ are held fixed), the number of modules will remain fixed, until the change becomes large enough to accommodate one additional or one less module. Thus, the number of modules is also a topological invariant of the system, through the module number m. Third, the number of modules does not depend on the system size L, or the number of neurons n_{DV} the system is discretized into (cf. Fig. 3f). Fourth, since the average module size will be L/N_{mod}, the module sizes are extensive in L. Thus, for sufficiently large L, the module sizes can be orders of magnitude larger than the scales of the lateral interaction d and σ.

Note that the above argument on topological robustness of the modularization of the system is not restricted to the case of *simple* fixed-scale kernels. Indeed, for any fixed-scale interaction W^f, the topological number m for any given expressed module will correspond to selecting the m^{th} maximum of $\mathcal{F}W^f(k)$, for $k > 0$.

a. Module boundary locations

Following the peak-selection arguments made earlier, the module boundaries will occur at spatial locations that have $k^{*g}(n_{DV})$ in between k_n and k_{n+1} (the specific location will depend on the particular forms of the kernels). As a zeroth order approximation, we can assume that the module boundaries will occur near $(k_n + k_{n+1})/2$,

$$k^{*g}(n_{\text{boundary}}) \approx \frac{(2n+1)\pi + \phi}{d} \tag{D26}$$

and thus

$$n_{\text{boundary}} \approx \sigma^{-1}\left(\frac{\eta d}{(2n+1)\pi}\right). \tag{D27}$$

where σ^{-1} is the inverse function of $\sigma(n_{DV})$, $\sigma^{-1} \circ \sigma(x) = x$. Thus, while the specific positions of the module boundaries are dependent on the shape of the gradient $\sigma(n_{DV})$, qualitative features such as the number of modules, module periods and module sizes are indifferent to the particular forms of the gradient (cf. Fig. 3f).

In (Fig. 4d), we vary the width of the $\sigma(x)$ in two different ways: linearly along and in a square root along n_{DV}. This leads to a shift in the module boundary locations that is predicted by fourier theory.

8. 2D analysis

We have presented a majority of the above analysis for the case of one-dimensional grid cells. Here we briefly present the analogous computations for the Fourier transforms in two dimensions. We first demonstrate a classical result relating the Fourier transform of radially symmetric functions to the Hankel transform, which we shall then use to compute the relevant transforms. Consider the Fourier transform of a function $f(\mathbf{x}) = f(x, y)$

$$\mathcal{F}f(\mathbf{k}) = \int f(\mathbf{x}) e^{i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x}$$
$$\mathcal{F}f(k_x, k_y) = \int f(x, y) e^{ik_x x + ik_y y} dx dy.$$

Define polar coordinates in real and Fourier space such that:

$$x = r\cos\theta$$
$$y = r\sin\theta$$
$$k_x = k\cos\phi$$
$$k_y = k\sin\phi$$

This leads to the dot product $\mathbf{k}\cdot\mathbf{x}$ to be simplified as

$$\begin{aligned} k_x x + k_y y &= rk(\cos\theta\cos\phi + \sin\theta\sin\phi) \\ &= rk\cos(\theta - \phi) \end{aligned}$$

Thus,

$$\mathcal{F}f(k_x, k_y) = \mathcal{F}f(k, \phi) = \int_0^\infty \int_0^{2\pi} r dr d\theta f(r, \theta) e^{ikr\cos(\theta-\phi)}$$

In all cases of interest, the function f is a kernel, and is hence a radially-symmetric real function $f(r, \theta) = f(r)$. Similarly, the Fourier transform $\mathcal{F}f$ will also be a real radially-symmetric function $\mathcal{F}f(k, \phi) = \mathcal{F}f(k)$. Thus

$$\mathcal{F}f(k) = \int_0^\infty \int_0^{2\pi} r dr d\theta f(r) e^{ikr\cos(\theta-\phi)}, \tag{D28}$$
$$= \int_0^\infty r dr f(r) \int_0^{2\pi} e^{ikr\cos(\theta-\phi)} d\theta, \tag{D29}$$
$$= 2\pi \int_0^\infty r f(r) J_0(kr) dr, \tag{D30}$$

where J_0 is the Bessel function of the first kind, defined by

$$J_0(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{ix\cos(\theta-\phi)} d\theta.$$

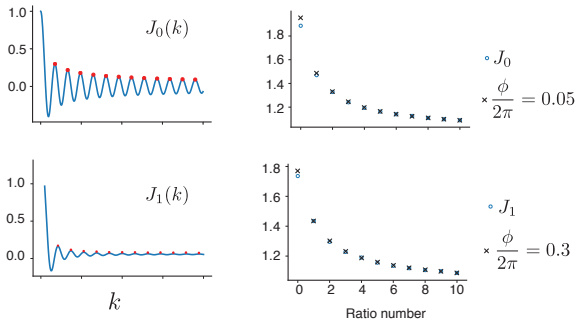


FIG. 16. Bessel functions (left column) and period ratios for Bessel function maxima (right column) with their best-fit values of ϕ for the period ratios corresponding to Eq. (D1)

Equation (D30) defines the Hankel transform (of order zero) of $f(r)$ — the radial component of the Fourier transform of the kernel $f(\mathbf{x})$ is simply the Hankel transform of $f(|\mathbf{x}|)$.

For the localized gaussian secondary interaction, we can calculate the Fourier transform analytically.

$$\begin{aligned}\mathcal{F}W_{\text{local}}(k) &= 2\pi \int_0^\infty r \left[\alpha_E e^{-r^2/2\sigma_E^2} - \alpha_I e^{-r^2/2\sigma_I^2} + \alpha_S e^{-(r-d)^2/2\sigma_S^2}\right] J_0(kr)dr \\ &= 2\pi \left[\alpha_E \sigma_E^2 e^{-k^2\sigma_E^2/2} - \alpha_I \sigma_I^2 e^{-k^2\sigma_I^2/2} + \alpha_S J_0(kd)\sigma_S^2 e^{-k^2\sigma_S^2/2}\right]\end{aligned}$$

We can also analytically calculate the Fourier transform for a box-like interaction:

$$\begin{aligned}\mathcal{F}W_{\text{diffuse}}(k) &= 2\pi W \int_0^d r J_0(kr)dr \\ &= \frac{2\pi W}{k^2} \int_0^{kd} \rho J_0(\rho)dr \\ &= \frac{2\pi W}{k^2} \left[kdJ_1(kd)\right] \\ &= \frac{2\pi W d^2 J_1(kd)}{kd}\end{aligned}$$

We can similarly also define a two-dimensional equivalent of the decaying kernel, as the convolution of the half-sized circular box kernel with itself. Thus, by applying convolution theorem to the result on diffuse kernels we obtain

$$\mathcal{F}W_{\text{decaying}}(k) = \left[\frac{\pi W d J_1(kd/2)}{k}\right]^2.$$

Note that $J_0(x)$ and $J_1(x)$ display qualitatively similar behavior to $\cos(x)$ and $\sin(x)$ respectively, apart from an amplitude modulation of the peaks — particularly, we note that the Bessel functions display approximately periodic maxima, which was the central property required for all of our results on modularization and peak selection to apply. We demonstrate this in Fig.16, where we show that the maxima of the Bessel functions are approximately periodic, and fit the form of Eq. (D1) well. In particular, note that the best-fit value of ϕ for $J_0(k)$ is approximately 0, which is similar to $\cos(k)$, and the best-fit value of ϕ for $J_1(k)$ is approximately $\pi/4$, which is similar to $\sin(k)$.

We implemented a 2d simulation that generates 3 discrete modules as shown in Figure 18. For computational feasibility, the simulation was performed in 2 parts: one with $x \in [0, 0.6N_x^{2d}]$ and the

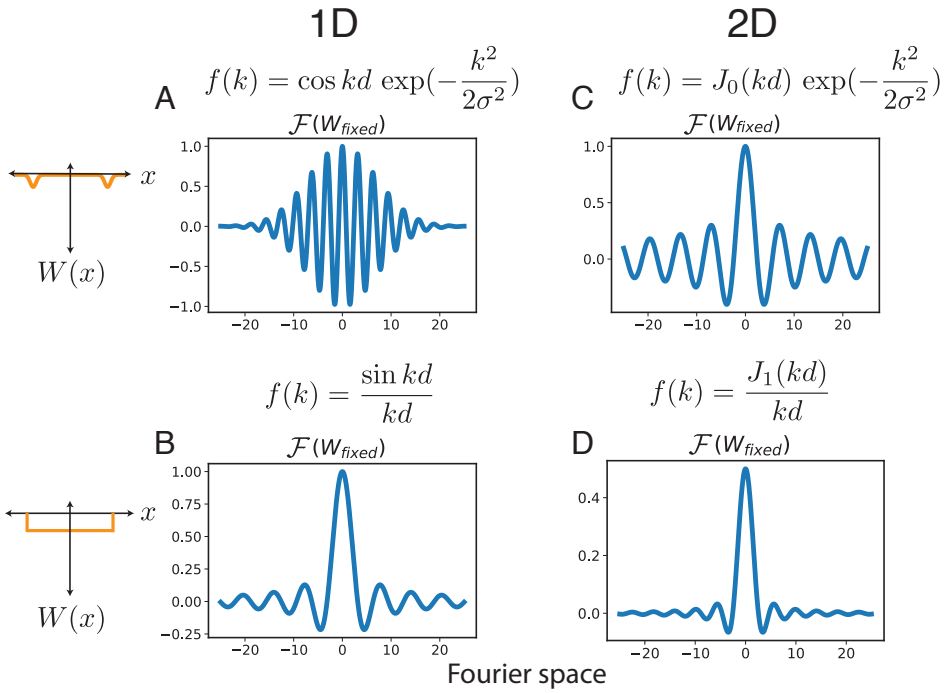


FIG. 17. Fixed interactions(left, in orange) and their oscillatory Fourier transforms in 1D (left column) and 2D (right column).

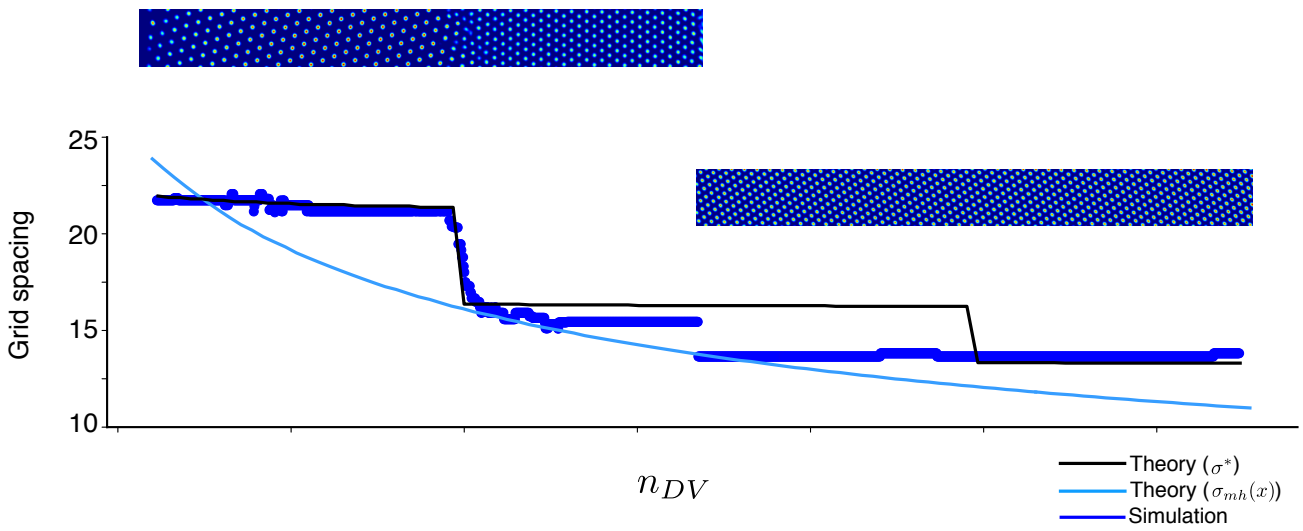


FIG. 18. **2d simulation with 3 modules:** (top) Snapshots of population activity showing 3 discrete 2d grid modules, (bottom) plot of grid spacing and comparision with Hankel transform predictions. Grid spacing determined by calculating the (neural) spatial auto-correlation of the population firing activity.

other with $x \in [0.6N_x^{2d}, N_x^{2d}]$. The weight matrices for each network were of size 100x1000 each. The weight matrix for a single large 100x2000 network would have contained $4x10^{10}$ elements, which we found prohibitively difficult and slow to run.

Fig 19(a) shows another instance of a modular 2d network, the only difference being the value of d_{loc}, which changed from 50 to 45. Fig 19(b) shows the same simulation with 2 distinct random initializations. The pair of resulting modules in each simulation have different relative orientations.

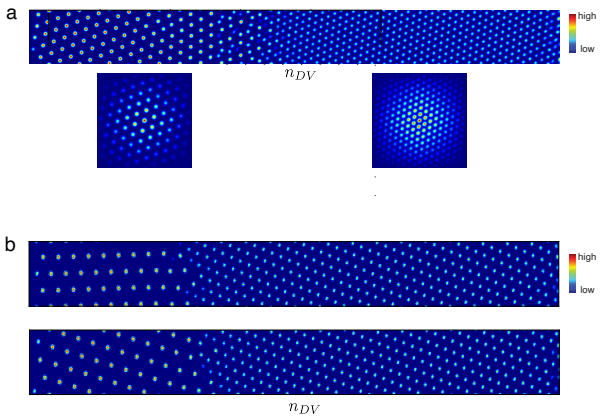


FIG. 19. (a) *Top*: Another instance of a spontaneously formed modularized two dimensional network with parameters given in Table 4. *Bottom*: autocorrelation plots of regions within the two formed modules, demonstrating the six-fold symmetry of the pattern formation. (b) Two different random initializations of the network from Fig 2h show different relative orientations between the 2 formed modules.

Because finite size effects from our simulations also partially constrain the orientations of the modules (data not shown), we cannot make predictions about the relative orientations of the grid modules found in experiments [27].

9. Robustness to spatial noise

In the main text, we discussed how the topological robustness properties of peak selection result in the formed modules being stable to several forms of noise. Particularly, here we focus on the robustness to spatial heterogeneities in the lateral interaction kernels.

We first examine the robustness to spatial heterogeneities in the pattern forming kernel W^g. To construct such an inhomogeneous pattern-forming interaction, we construct the noisy kernel at location $\mathbf{x}$, by replacing the spatially homogeneous kernel $W^g[\mathbf{x}, \mathbf{x}'] = W^g[\mathbf{x} - \mathbf{x}']$, with a spatially heterogeneous kernel $W^g_\xi[\mathbf{x}, \mathbf{x}'] = W^g[|\mathbf{x} - \mathbf{x}'| + \xi(\mathbf{x}')]$, where $\xi(\mathbf{x}')$ is a random number sampled independently for each spatial location $\mathbf{x}'$ with mean zero and variance ϵ^2. In Fig. 20d we present examples of such kernels for the case of $W^g[\mathbf{x}, \mathbf{x}']$ described by the box function Eq. (12). Note how the independent sampling of $\xi(\mathbf{x}')$ at each location results in a heterogenous kernel W^g_ξ that varies in scale at different $\mathbf{x}$, and is no longer radially symmetric.

Recall that peak selection entails that the grid period at any location n_{DV} is dependent on the set of potential maxima defined by $\mathcal{F}W^f(k)$, with a selection between these maxima performed by the broader peak of $\mathcal{F}W^g(k)$. If noise in the form of spatial heterogeneities are only introduced in W^g (and hence introduced in $\mathcal{F}W^g$) this results in a noisy selection function. However, since the same maxima will be chosen for a range of selection functions (See Fig. 20a-b), the heterogeneity in W^g will not be manifested in the emergent grid period.

We next consider the addition of similar heterogeneities in the fixed-scale interaction as well, W^f (such as in Fig. 20c). Note that maxima induced by *simple* W^f are at $k_n \approx (2n\pi + \phi)/d$, where n are the natural numbers, and hence the grid periodicity of the n^{th} module is given by $\lambda_n \approx d/(n+\phi/2\pi)$. If we consider $\mathcal{O}(\epsilon)$ noise added to W^f in the form of spatial heterogeneities, this would result in an $\mathcal{O}(\epsilon)$ error in the effective fixed-scale d. However, since λ_n is approximately d/n, thus the effective noise in periodicity of the n^{th} module, λ_n, will be $\mathcal{O}(\epsilon/n)$. Thus, higher module numbers (corresponding to modules with smaller grid periods) have additional error correction beyond the robustness conferred by the topological nature of the peak selection process. This

results in clean hexagonal firing fields despite inhomogeneities introduced in all lateral interactions as shown in Fig. 6.

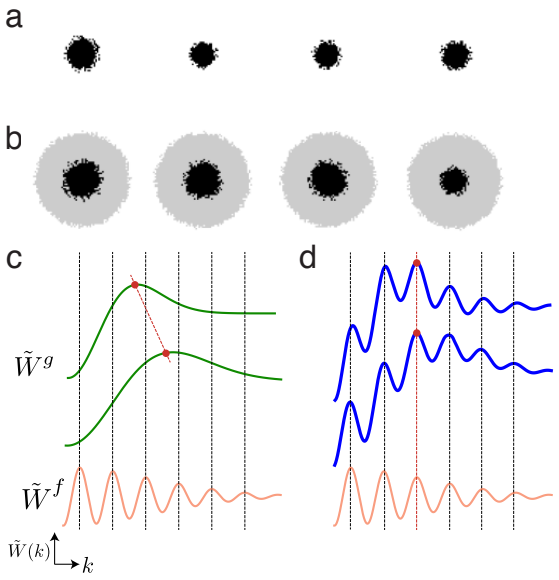


FIG. 20. **Noise robustness in peak selection process** demonstrating how the additional of the smaller oscillatory fourier transform of the fixed interaction leads to no change in maxima despite smooth movement of the primary peak. (a) Example pattern forming interaction kernels from 4 neurons without a secondary fixed scale interaction (b) Example composite kernels from 4 neurons showing both the pattern forming interaction (black) and fixed scale interaction (grey)(c-d) The movement of the pattern forming interaction leads to a shift in the location of the gloabl maxima in the absence of a secondary interaction. This secondary interaction prevents any shift in the location of the global maxima when defined by the sum of the pattern forming interaction and the fixed scale interaction.

a. Peak selection stabilizes against finite-neuron-number effects

This robustness to spatial noise discussed in the above section, through the addition of a secondary length-scale kernel, also manifests itself in terms of added stability towards finite-neuron-

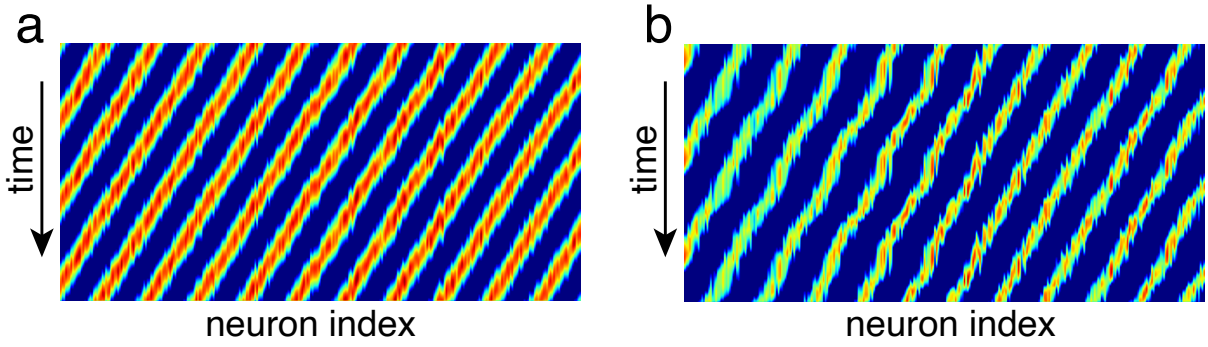


FIG. 21. **Addition of secondary length-scale interaction provides robustness to path integration** (a) With a noisy kernel and the inclusion of fixed length interaction, the model can perfectly path integrate. (b) With a noisy kernel but without the fixed length interaction, the model's path integration performance is noisy.

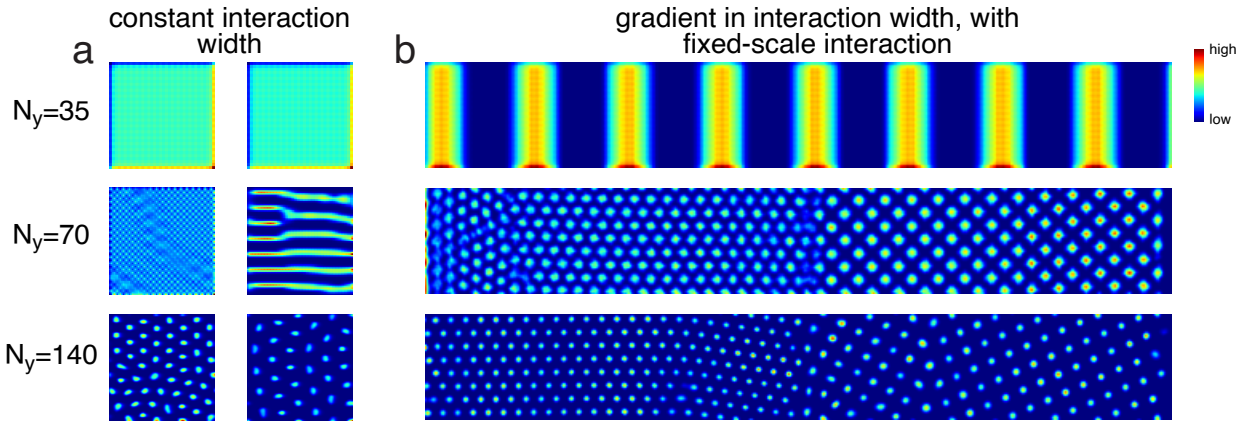


FIG. 22. **Addition of secondary fixed-scale interaction stabilizes against finite-neuron-number effects** (a) Simulations of 2d grid cell network with a single length scale in the interaction kernel, with system sizes 35×35, 70×70 and 140×140 neurons. The largest system results in pattern formation, while the smaller ones do not, due to finite size effects. *Left* and *right* correspond to two different interaction kernel widths. (b) Simulations of a 2d module forming grid cell network with system size along the shorter, non-gradient dimension being 35, 70 and 140 respectively. The gradient in interaction widths along the longer dimension linearly changes from between the two widths considered in the corresponding row in (a). The secondary interaction encourages the formation of module even at sizes where the single module simulation does not.

number effects. To examine this effect, we first compare the grid pattern formation on neural sheets with $N = 35 \times 35$, $N = 70 \times 70$ and $N = 140 \times 140$ neurons, Fig. 22a by using a single homogeneous pattern forming kernel. Here, we note that the neural sheet with 140×140 neurons is large enough for stable pattern formation, whereas in the smaller sheets (with scaled kernel interaction widths) the patterned dynamics is washed out due to finite-neuron-number effects. Then, we considered rectangular sheets whose shorter axes had 35, 70 and 140 neurons, and longer axes scaled sufficiently to fit two modules. On these sheets, we simulated dynamics with an interaction kernel given by the sum of a graded W^g and a fixed W^f, similar to Fig. 2. For a neural sheet with only 35 neurons along the short axis we could not observe stable hexagonal pattern formation. Remarkably, even when the short axis has only 70 neurons (a neuron number that was too small to permit pattern formation in the single pattern forming kernel case), we continue to obtain hexagonal pattern formation, Fig. 22b, *top*. At this neural sheet size, when simulating with different random initial conditions (not shown here), a small fraction of simulations also failed to show pattern formation.

Thus, we empirically observe that the addition of a secondary scale as W^f can stabilize against finite neuron-number effects, implying that the formation of K modules through peak selection may require marginally fewer than K times as many neurons as would be necessary for a single grid cell module.

Appendix E: Alternative biophysical gradients

For introducing a gradient in time constant, we introduce two populations of neurons with different time constants. In this model, we consider two interacting cell types such that one of the cell types has a fixed time constant along the DV axis, and the other whose time constant is smoothly graded along that axis. The time-constant affects the time-constant of dendritic integration. Though biologically distinct, the effect of modulating the dendritic integration time-constant of one cell type is mathematically similar to the effect of modulating the width of connectivity of

one cell type while keeping the other fixed:

$$\dot{s_1} = -s_1/\tau_1 + \text{ReLU}[\int W_1(x,x')s_1(x')dx' + \int W_2(x,x')s_2(x')dx']$$
$$\dot{s_2} = -s_2/\tau_2 + \text{ReLU}[\int W_1(x,x')s_1(x')dx' + \int W_2(x,x')s_2(x')dx']$$

Here, W_2 is parametrized as the fixed-scale interaction described in the main text, i.e., W_f; and W_1 incorporates the effect of dendritic attenuation as

$$W_1(x,x') = W_{n_{DV}}(x-x') = W_{pf}(x-x') * \exp(-|x-x'|/\sigma_\tau(n_{DV}))$$

where W_{pf} is the pattern forming kernel (similar to W_g in the main text, but without any explicit gradient in widths), and $/\sigma_\tau$ captures the gradient in time constant via an exponential attenuation. Thus, in combination with the dendritic attenuation, the W_1 interaction plays the role of W^g in the main text, providing an effective gradient in spatial lengthscale though only a gradient in the neural time constant.

To obtain multiscale pattern formation, we also examine a gradient in synaptic strength. Here, a nonlinearity is imposed on the synaptic strength that thresholds the maximum possible coupling strength. The summed graded plus fixed-scale interaction W is thus changed to:

$$W_{n_DV}(\mathbf{\Delta x}) = \Phi[\gamma(n_{DV})W(\mathbf{\Delta x})], \tag{E1}$$

where $\gamma(n_{DV})$ is the gradient in synaptic strength, and Φ is a thresholding function, $\Phi[x] = -\Theta$ if $x < -\Theta$, $\Phi[x] = \Theta$ if $x > \Theta$ and $\Phi[x] = x$ otherwise. This nonlinearity results in an interplay between synaptic strength and an effective interaction width. For example, the effective lengthscale $\sigma(n_{DV})$ will scale as $\sim \sqrt{\log \gamma(n_{DV})}$ for a mexican-hat W^g (assuming that $|W^f| < \Theta$ as would be the case for weak W^f).

Note that in each of the two cases described above, there is an effective lengthscale that varies in a gradient in the main pattern-forming interaction, akin to W^g in the main text. Further, in both cases the secondary interaction, W^f remains unaffected. Thus, following the arguments presented in the main text, the length scale of the pattern formation will be governed by the local maxima of the Fourier transform of W^f, exactly equivalent to the pattern formation dynamics described for the case of a gradient in the explicit interaction width. As a result, the period ratio prediction will have the exact same form as Eq. (5), and similarly all other derived results follow through for these alternative biophysical gradients in time constant or synaptic strength.

1. Comparison of experimental observations with predicted period ratios

The general mechanism of peak-selection presented above describes how discrete modules can spontaneously arise in the presence of continuous gradients, by consideration of an additional fixed-scale lateral interaction W^f. However, this mechanism does not provide any testable predictions for the ratio of grid periods unless additional assumptions are made. If indeed we assume that W^f is a *simple* kernel, i.e., W^f is primarily defined by a single spatial scale, then we demonstrated in SI Sec. D that the period ratios will be given by the simple formula, Eq. D1. In this section, we show that experimental observations of grid periods largely appear to match our predicted period ratios for *simple* kernels with $\phi = 0$.

For verification of our main results on the predicted form of period ratios, we examine the literature for grid period measurements for multiple simultaneously measured grid modules in rats[27, 129–131]. We note that a large fraction of experimental observations of grid cells with

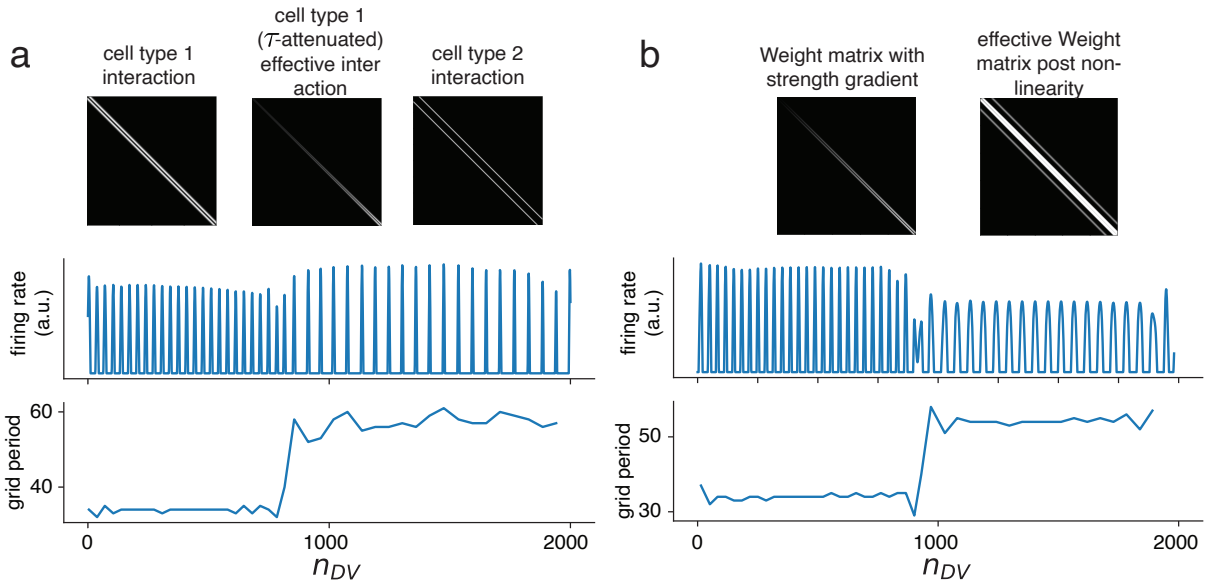


FIG. 23. **Alternative biophysical gradients:** (a) Neuronal time-constant: We introduce 2 populations of cells, one with fixed neuronal time constant and one with a gradient in the neuronal time constant. (b) Synaptic strength: We introduce a gradient in the inhibitory strength along with a thresholding non-linearity that couples the strength of interaction with the width. Shown in the figure panels are weight matrices, steady state activity pattern showing 2 discrete modules and grid periods measured as distance between activity bumps in number of neurons (top-bottom)

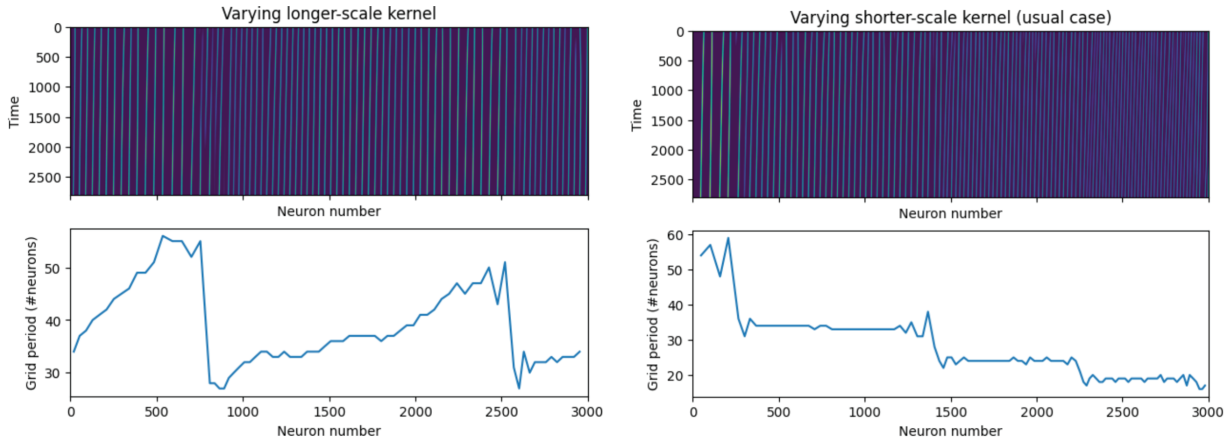
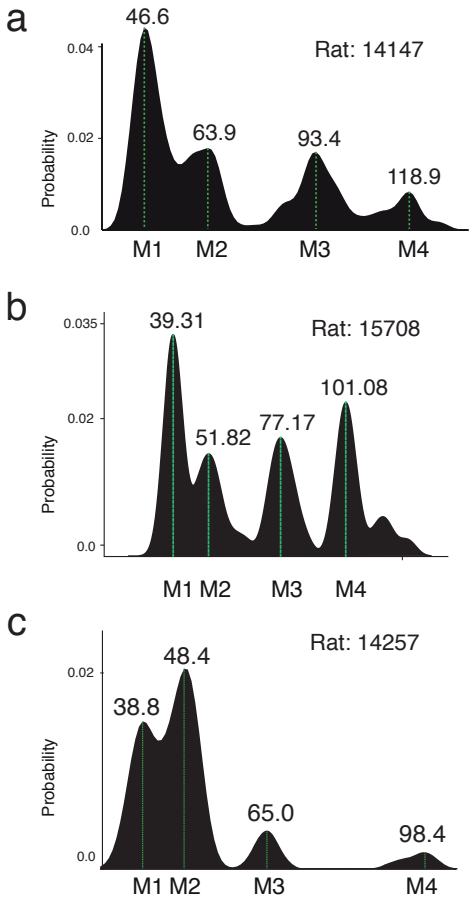


FIG. 24. **Adding a gradient in the longer-scale interaction does not lead to modules:** Steady state activity pattern and grid periods measured as distance between activity bumps with a gradient in longer-scale interaction kernel (left) and gradient in the shorter-scale interaction kernel (right).

more than one module measure only two modules. For a single pair of grid periods λ_1 and $\lambda_2 > \lambda_1$, we can always explicitly solve for ϕ and m in Eq. (D1), to obtain

$$\frac{\phi}{2\pi} = \left\{ \frac{\lambda_2}{\lambda_1 - \lambda_2} \right\}, m = \left\lfloor \frac{\lambda_2}{\lambda_1 - \lambda_2} \right\rfloor, \tag{E2}$$

where $\{x\}$ represents that fractional part of x, and $\lfloor x \rfloor = x - \{x\}$ represents the integer part of x. Thus, a single ratio, because it can always be fit by Eq. (D1), imposes no constraints on the accuracy of the expression.



observation	$r_m(1.26)$ prediction	deviation
1.273	1.83	30.%
1.462	1.46	0.13%
1.371	1.31	4.65%

observation	$r_m(0.7)$ prediction	deviation
1.310	1.90	31.%
1.489	1.48	0.61%
1.318	1.32	0.15%

observation	$r_m(0)$ prediction	deviation
1.514	3/2 = 1.5	0.93%
1.343	4/3 = 1.33	0.75%
1.247	5/4 = 1.25	0.24%

FIG. 25. The 3 rats from Stensola *et al.* with 4 modules and their corresponding periods.

It is possible to obtain a value of ϕ from Eq. (D1) and a single pair of periods; however, the estimate obtained from a single pair is not robust: r_m depends too sensitively on ϕ. For example, in [27], Rat 13388 exhibits grid periods of ≈ 53.24 cm and ≈ 43.00 cm (as estimated from SI Fig. 12b in [27]); Eq. (D1) then yields $\phi/(2\pi) = 0.199$. Assuming a very small measurement error of $\sim 0.5cm$ in the larger period, such that if it were 53.75 cm instead of 53.24, would yield ϕ exactly equal to zero. A simple sensitivity analysis of the magnitude of error in estimating ϕ can be performed from Eq. (E2):

$$\delta\phi = 3\epsilon\frac{\lambda_2}{\lambda_1 - \lambda_2} \approx 3\epsilon m, \tag{E3}$$

where ϵ represents the fractional error in the estimate of grid period. Thus, particularly for smaller grid periods (corresponding to larger m), even small errors in grid period estimation can result in a large error in ϕ, making the errorbars in the estimation of ϕ from a single pair of periods large.

To obtain results with significant statistical certainty, we focus our analysis on published experimental studies that measure at least 50 grid cells per animal, spanning at least 3 distinct modules. This restriction results in grid period data sets for three rats — we present kernel density estimates of the module periods for each of them in Fig. 25 (Fig. 25c corresponds to the data presented in the main text in Fig. 5).

We have already demonstrated in Fig. 5 that Rat 14257 presents an extremely accurate match to the period ratio prediction for $\phi = 0$ (i.e., predicted period ratios of 2, 3/2, 4/3, 5/4,...); in addition, Rat 14147 (observed period ratios of 1.27 , $1.46 \approx 3/2$, $1.37 \approx 4/3$) and Rat 15708

(observed period ratios of 1.31, 1.49 $\approx 3/2$, 1.32 $\approx 4/3$) also match $\phi = 0$ very well (R^2 values of 0.999, 0.979, and 0.968 for Rats 14257, 15708, 14147 resp.) for all grid modules except for the module with the largest period.

Why is there an observed discrepancy for the grid module with the largest period? We propose four possible reasons for this discrepancy: Firstly, this discrepancy may be a result of the approximation made in arriving at Eq. (D17) — since the approximation is particularly accurate for $kd \gtrsim \mathcal{O}(1)$, the potentialy mismatch would primarily affect only the largest grid period module. Secondly, as demonstrated in Sec. D 9, the grid module corresponding to the largest grid period will have the least robustness to noise in the fixed-scale interaction, potentially introducing a large variance in the grid period for that module. Thirdly, as can be seen in Fig. 3h and Eq. (E3), the error in estimating the grid period for the first module ($m = 1$) is the most susceptible to errors in the value of ϕ Lastly, our predictions for grid period ratios Eq. (D1) are for the case of *simple* kernels that have a single spatial scale. A discrepancy at only the largest grid module may thus be suggestive of fixed-scale interactions that are primarily described by a single scale, with an additional low frequency perturbation at a larger spatial scale.

However, note that (particularly for Rats 14147 and 14257) there are relatively few grid cells observed from this largest period module, and the resulting uncertainty in period estimation may instead contribute to the error. In sum, apart from the possibility of some additional low frequency perturbations, the experimental data for rats with several simultaneously observed grid modules is largely consistent with the predicted period ratios for simple kernels with $\phi = 0$.

Skipped modules: Sometimes, neural recordings can miss a module. This can cause a large deviation from our predictions. For example, for a set of 5 modules following period ratios M4/M5 = 1.20, M3/M4 = 1.25, M2/M3 = 1.33, M1/M2 = 1.5. If recordings had missed module M4, the measured ratios would be M1/M2 = 1.5, M2/M3 = 1.33, M3/M5 = 1.5.

However, we do note that available data on multiple modules with a statistically large number of grid cells per module are quite sparse. To obtain further verification of our theoretical results, including the prediction of Eq. (D1) and even more specifically the hypothesis that ϕ is close to zero, additional data with multiple simultaneously observed grid modules will be important.

Appendix F: Lyapunov Function

The energy function of continuous time neural networks can be written as [132]:

$$E(\mathbf{s}) = -\frac{1}{2}\sum_{ij} s(i) W_{ij} s(j) + \sum_i \int_0^{s(i)} \phi^{-1}(s)\, ds - \sum_i I_i s(i), \tag{F1}$$

where $\mathbf{s}$ represents a vector of the synaptic activation at each neuron in the network, and I_i is the input bias to neuron i. For simplicity and since linear analysis does a remarkably good job in predicting the formed modules, let us restrict ourselves to the case of $\phi(x) = x$. Also, since the system is locally translationally invariant, we know that the dominant modes are going to be periodic. Hence, we may evaluate the energy function of the network dynamics (in the linearized regime) by assessing the energy of the periodic neural activity modes:

$$\mathbf{s}_k(\mathbf{x}) = A \sin(\mathbf{k} \cdot \mathbf{x} + \delta) + B, \tag{F2}$$

where $\mathbf{k} = k\hat{\mathbf{k}}$ is an arbitrary Fourier space vector, and A,B and δ are arbitrary constants. For these modes, we can write the energy function in the continuum limit as:

$$E[s_{\mathbf{k}}(\mathbf{x})] = -\frac{1}{2}\int d\mathbf{x} d\mathbf{x}' W(\mathbf{x}, \mathbf{x}') s_{\mathbf{k}}(\mathbf{x}) s_{\mathbf{k}}(\mathbf{x}') + \frac{1}{2}\int d\mathbf{x} s_{\mathbf{k}}(\mathbf{x})^2$$

Assuming that the system size L is large,

$$\begin{aligned}
2E[s_{\mathbf{k}}(\mathbf{x})] &= -\int W(\mathbf{x}-\mathbf{x}')[A\sin(\mathbf{k}\cdot\mathbf{x}+\delta)+B][A\sin(\mathbf{k}\cdot\mathbf{x}'+\delta)+B]d\mathbf{x}d\mathbf{x}' + \int [A\sin(\mathbf{k}\cdot\mathbf{x}+\Delta)+B]^2 d\mathbf{x} \\
&= -A^2\int d\mathbf{u}d\mathbf{v}W(\mathbf{u})\cos(\mathbf{k}\cdot\mathbf{u}) + A^2\int d\mathbf{u}d\mathbf{v}W(\mathbf{u})\cos(2\mathbf{k}\cdot\mathbf{v}+\delta) + B^2\int d\mathbf{x}d\mathbf{x}'W(\mathbf{x}-\mathbf{x}') + L(A^2/2+ \\
&= -A^2L\int d\mathbf{u}e^{i\mathbf{k}\cdot\mathbf{u}}W(\mathbf{u}) + A^2\int d\mathbf{u}W(\mathbf{u})\int d\mathbf{v}\cos(2\mathbf{k}\cdot\mathbf{v}+\delta) + B^2\int d\mathbf{u}d\mathbf{v}W(\mathbf{u}) + L(A^2/2+B^2)/2 \\
&= -A^2L\tilde{W}(k) + LB^2\bar{W} + L(A^2/2+B^2), \\
&= -\text{constant}_1\times\tilde{W}(k) + \text{constant}_2
\end{aligned}$$

where have used the simple trigonometric identity, $2\sin(C)\sin(D) = \cos(C-D) - \cos(C+D)$, and a change of variables, $\int d\mathbf{x}d\mathbf{x}' = (1/2)\int d(\mathbf{x}-\mathbf{x}')d(\mathbf{x}+\mathbf{x}') = \int d\mathbf{u}d\mathbf{v}$, with $\mathbf{u} = \mathbf{x}-\mathbf{x}'$ and $\mathbf{v} = \dfrac{1}{2}(\mathbf{x}+\mathbf{x}')$.

Thus, we obtain that the energy function $E[s_{\mathbf{k}}]$ is a simple linear function of the Fourier transform $\tilde{W}(k)$ of the recurrent weight matrix. The minimum energy solution corresponds to the Fourier mode that maximizes $\tilde{W}(k)$. In other words, the dynamics is dominated by the k^* that maximizes $\tilde{W}(k)$. This result, derived from an energy landscape perspective, is equivalent to the result in Eq. (B10), which we obtained earlier via perturbation analysis.

Appendix G: General formulation of module formation dynamics: Discrete peak selection via loss minimization

In Sec. F, we demonstrated how the pattern formation on the neural sheet can be derived via an energy minimization approach. Here, we use an energy landscape view to describe how loss function minimization results in modular solutions.

The key components for spatially modular solutions to arise from energy minimization are as follows: 1) A spatially-independent loss function $f(\theta)$ with multiple local maxima and minima; 2) A gradient in a spatially-dependent variable, $\theta_0(x)$; and 3) A coupling between the system parameters θ and θ_0, that results in a combined loss function

$$L(\theta,\theta_0(x)) = (1-\alpha)f(\theta) + \alpha\|\theta-\theta_0(x)\|^2 \tag{G1}$$

Under appropriate constraints on $f(\theta)$, solving the following optimization at each x

$$\theta^*(x) = \arg\max_\theta L(\theta,x) \tag{G2}$$

will produce discrete, step-like changes as a function of x. This happens because the smooth minimum given by the $||\theta-\theta_0(x)||^2$ term effectively selects one of the local minima in $f(\theta)$ as the global minimum. As the function $||\theta-\theta_0(x)||^2$ slides smoothly along with x, the peak of $f(\theta)$ selected as the global minimum remains the same for some time, then jumps abruptly. These step-like changes are modular solutions to the global optimization problem. The energy function defined in Eq. (G1) can be viewed as a regularized optimization problem, with the spatially-dependent regularizer $||\theta-\theta_0(x)||^2$ acting as a prior that selects one of the minima of $f(\theta)$ at each location (Fig. 26).

The correspondence of this general picture with the peak selection mechanism described in the main text follows directly with the following identifications: the spatially independent nonlinear loss function $f(\theta)$ with the fixed-scale interaction W^f; the spatially varying parameter prior $\theta_0(x)$ with the graded scale $\sigma(n_{DV})$ of the pattern-forming kernel; the combined loss $L(\theta,x)$ with the full

kernel $W_{n_{DV}}$; and the spatially-varying, multi-step-like set of optima $\theta^*(x)$ with the grid periods $\lambda^*(x)$, respectively. Similar to peak selection for grid cells, the formed modules in this generalized setting will also inherit topological robustness and stability.

We demonstrate a numerical example of this in Fig. 7, where we construct $f(\theta)$ as a random sample from a Gaussian process with a radial basis function kernel, and simulate gradient descent dynamics on the loss function $L(\theta, \theta_0(x))$. To prevent the dynamics from getting stuck in local minima of L, we simulate the gradient descent first purely on the regularization term, with gradually increasing strength of the rugged loss function, through gradually decreasing α with increasing time.

Although we primarily focused on the peak selection process in Fourier space for multi-periodic patterning in grid cells, we also showed that it has a general formulation in terms of dynamics on an energy landscape: One (spatially invariant) interaction sets up an optimization problem with multiple local minima, while a second (spatially graded) interaction defines a locally shallow single-optimum landscape, with a smoothly shifting optimum as a function of space. Thus, the shallow optimum selects one of the narrow local optima as the global optimum, with discontinuous jumps to the next local minimum even as the parameters vary smoothly. This analytical formulation provides a simplifying mathematical perspective on how smooth gradients could lead to discrete patterning and modular specialization in the brain and body [34, 43, 133].

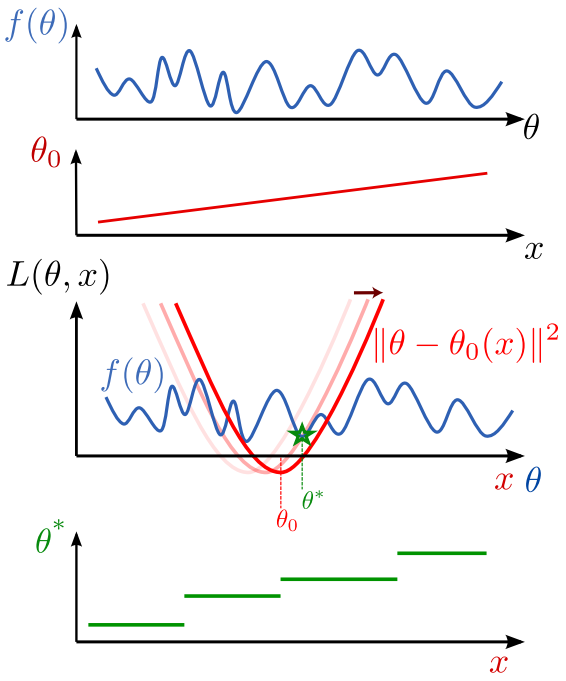


FIG. 26. **A general setting for peak-selection** Assuming a loss function $f(\theta)$ (blue) and a spatially dependent quantity θ_0 (red), a combined loss function $L(\theta, x)$ can be constructed such that the x-dependent optimizer of $L(\theta, x)$ will be modular (green), since it will be constrained to correspond to one of the minima of $f(\theta)$.

Appendix H: The emergence of modules corresponds to the formation of localized eigenvectors

As has been observed before [134], a neural network endowed with slowly varying local interactions shows diverse timescales that are spatially localized: different parts of the network respond with disparate temporal dynamics. We also find a localization of eigenvectors in our multi-module grid network, Fig. 27A. Similar to [134], our interaction matrix has a locally circulant form (due to the slowly varying gradient in lateral inhibition width). This is a signature of a phase transition, similar to the Anderson localization transition in condensed matter physics [60]. The eigenvectors for a regular pattern forming interaction in traditional continuous attractor models are delocalized fourier waves which are then transformed into localized fixed-wavelength gaussian wavepackets with the addition of the gradient and fixed scale interaction.

We find that in the resulting set of localized eigenvectors, each has a different but constant period, Fig. 27B. These periods exactly match the spatial periods of the modules formed in steady state. In sum, the locally circulant matrix gives rise to eigenvector localization, and the localized eigenvectors correspond to the modules.

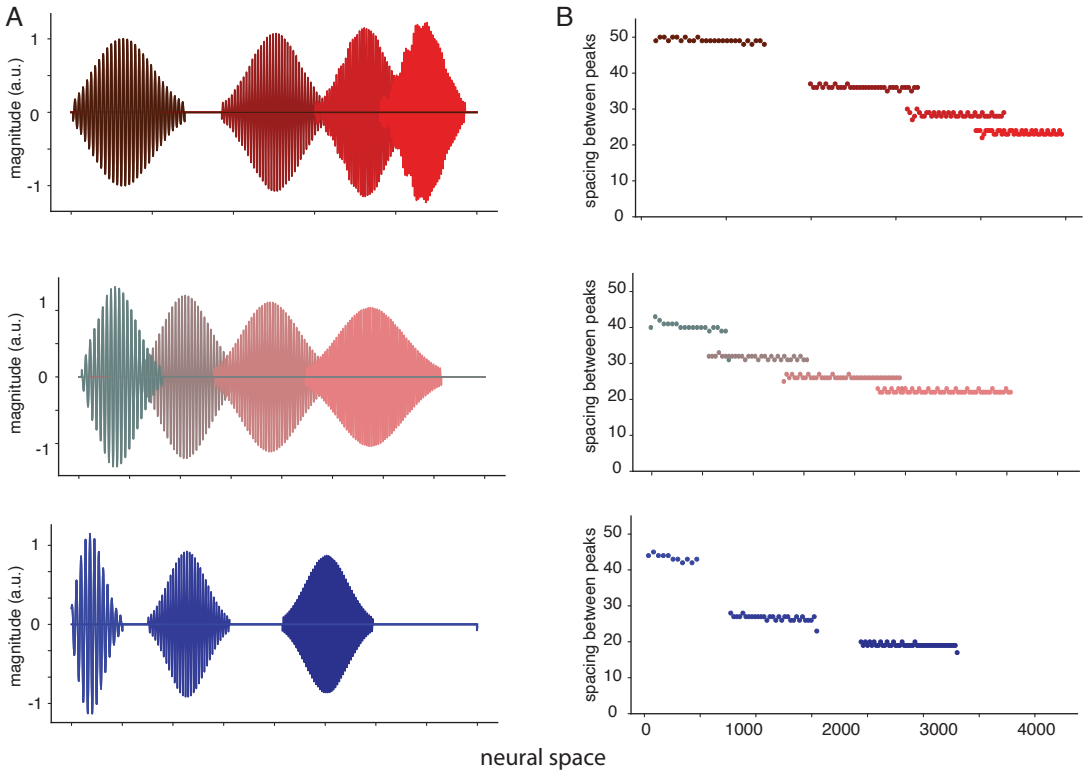


FIG. 27. **Localization of eigenvectors**: A) Eigenvectors of various one-dimensional interaction weight matrices along with the corresponding inter-peak spacings are localized, B) The periodicity within an eigenvector is constant.