A geometric attractor mechanism for self-organization of entorhinal grid modules

Louis Kang[1,2,*] and Vijay Balasubramanian[1]

[1] *David Rittenhouse Laboratories, University of Pennsylvania, Philadelphia, Pennsylvania, USA*
[2] *Redwood Center for Theoretical Neuroscience, University of California, Berkeley, Berkeley, California, USA*

(Dated: March 8, 2019)

Grid cells in the medial entorhinal cortex (MEC) respond when an animal occupies a periodic lattice of "grid fields" in the environment. The grids are organized in modules with spatial periods, or scales, clustered around discrete values separated by ratios in the range 1.2–2.0. We propose a mechanism that produces this modular structure through dynamical self-organization in the MEC. In attractor network models of grid formation, the grid scale of a single module is set by the distance of recurrent inhibition between neurons. We show that the MEC forms a hierarchy of discrete modules if a smooth increase in inhibition distance along its dorso-ventral axis is accompanied by excitatory interactions along this axis. Moreover, constant scale ratios between successive modules arise through geometric relationships between triangular grids and have values that fall within the observed range. We discuss how interactions required by our model might be tested experimentally.

A grid cell has a spatially modulated firing rate that peaks when an animal reaches certain locations in its environment [1]. These locations of high activity form a regular triangular grid with a particular length scale and orientation in space. Every animal has many grid cells that collectively span a wide range of scales, with smaller scales enriched dorsally and larger scales ventrally along the longitudinal axis of the MEC [2]. Instead of being smoothly distributed, grid scales cluster around particular values and thus grid cells are partitioned into modules [2]. Consecutive pairs of modules have scale ratios in the range 1.2–2.0 [2–4]. Across animals, the average scale ratio is constant from one pair of modules to the next [2]. These observations underlie the possibility that the grid system favors a universal scale ratio in the range 1.4 [2] to 1.6 [3, 4].

Encoding spatial information through grid cells with constant scale ratios is thought to provide animals with an efficient way of representing their position within an environment [5–11]. Moreover, periodic representations of space permit a novel mechanism for precise error correction against neural noise [12] and are learned by machines seeking to navigate open environments [13, 14]. These findings provide motivation for forming a modular grid system with a constant scale ratio, but a mechanism for doing so is unknown. Continuous attractor networks [15, 16], a leading model for producing grid cells, would currently require discrete changes in scales to be directly imposed as sharp changes in parameters, as would the oscillatory interference model [17, 18] or hybrid models [19]. In contrast, many sensory and behavioral systems have smooth tuning distributions, such as preferred orientation in visual cortex [20] and preferred head direction in the MEC [21]. A self-organizing map model with stripe cell inputs [22] and a firing rate adaptation model with place cell inputs [23] can generate discrete grid scales, but their ratios are not constant or constant-on-average unless explicitly tuned.

* Corresponding author: louis.kang@berkeley.edu

Here, we present a simple extension of the continuous attractor model that adds excitatory connections between a series of attractor networks along the dorso-ventral axis of the MEC, accompanied by an increase in the distance of inhibition. The inhibition gradient drives an increase in grid scale along the MEC axis. Meanwhile, the excitatory coupling discourages changes in grid scale and orientation unless they occur through geometric relationships with defined scale ratios and orientation differences. Competition between the effects of longitudinal excitation and lateral inhibition self-organizes the complete network into a discrete hierarchy of modules. Certain grid relationships are geometrically stable, which makes them, and their associated scale ratios, insensitive to perturbations. The precise ratios that appear depend on the balance between excitation and inhibition and how it varies along the MEC axis. We show that sampling across a range of these parameters leads to a distribution of scale ratios that matches experiment and is, on average, constant from the smallest to the largest pair of modules.

Continuous attractors are a powerful general method for self-organizing neural dynamics. To our knowledge, our results are the first demonstration of a mechanism for producing a discrete hierarchy of modules in a continuous attractor system.

RESULTS

Standard grid cell attractors are not modular

We assemble a series of networks along the longitudinal MEC axis, numbering them $z = 1, 2, \ldots, 12$ from dorsal to ventral (**Fig. 1a**). Each network contains the standard 2D continuous attractor architecture of the Burak-Fiete model [16]. Namely, neurons are arranged in a 2D sheet with positions (x, y), receive broad excitatory drive (Ref. 24 and **Fig. 1b**), and inhibit one another at a characteristic separation on the neural sheet (**Fig. 1c**; see

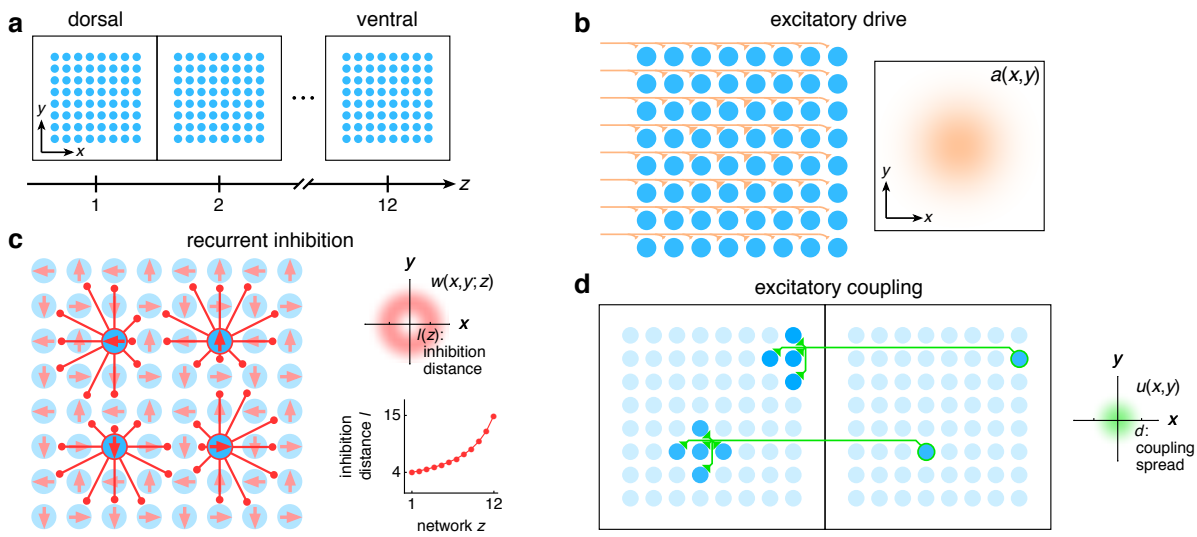

Figure 1. The entorhinal grid system as coupled 2D continuous attractor networks (**Methods**). (**a**) Each network z corresponds to a region along the dorso-ventral MEC axis and contains a 2D sheet of neurons with positions (x, y). (**b**) Neurons receive excitatory drive $a(x, y)$ that is greatest at the network center and decays toward the edges. (**c**) Neurons inhibit neighbors within the same network with a weight $w(x, y; z)$ that peaks at a distance of $l(z)$ neurons, which increases as a function of z. Each neuron has its inhibitory outputs shifted slightly in one of four preferred network directions and receives slightly more drive when the animal moves along its preferred spatial direction. (**d**) Each neuron at position (x, y) in network z excites neurons located within a spread d of (x, y) in network $z - 1$.

Methods for a complete description). This inhibition distance l is constant within each network but increases from one network to the next along the longitudinal axis of the MEC. With these features alone, the population activity in each network self-organizes into a triangular grid whose lattice points correspond to peaks in neural activity (first row of **Fig. 2a**). Importantly, the scale of each network's grid, which we call $\lambda(z)$, is proportional to that network's inhibition distance $l(z)$ ("uncoupled" simulations in **Fig. 2c**). Also, network grid orientations θ show no consistent pattern across scales and among replicate simulations with different random initial firing rates.

Following the standard attractor model [16], the inhibitory connections in each network are slightly modulated by the animal's velocity such that the population activity pattern of each network translates proportionally to animal motion at all times (**Methods**). This modulation allows each network to encode the animal's displacement through a process known as path-integration, and projects the network grid pattern onto spatial rate maps of single neurons. That is, a recording of a single neuron over the course of an animal trajectory would show high activity in spatial locations that form a triangular grid with scale Λ (third row of **Fig. 2a**). Moreover, $\Lambda(z)$ for a neuron from network z is proportional to that network's population grid scale $\lambda(z)$, and thus also proportional to its inhibition distance $l(z)$ (uncoupled simulations in **Fig. 2d**). To be clear, we call Λ the "spatial scale"; it corresponds to a single neuron's activity over the course of a simulation and has units of physical distance in space. By contrast, λ, the "network scale" described above, corresponds to the population activity at a single time and has units of separation on the neural sheet. Similarly, $\Theta(z)$ describes the orientation of the spatial grid of a single neuron in the network z; we call Θ the "spatial orientation". Like the network orientations θ discussed above, spatial orientations of grids show no clustering (uncoupled simulations in **Fig. 2d**).

With an inhibition distance $l(z)$ that increases gradually from one network to the next (**Fig. 1c**), proportional changes in network and spatial scales $\lambda(z)$ and $\Lambda(z)$ lead to a smooth distribution of grid scales (uncoupled simulations in **Fig. 2c**, **d**). To reproduce the experimentally observed jumps in grid scale between modules, the inhibition length would also have to undergo discrete, sharp jumps between certain adjacent networks. A further mechanism would be needed to enforce the preferred orientation differences that are observed between modules. In summary, a grid system created by disjoint attractor networks will not self-organize into modules.

Coupled attractor networks produce modules

Module self-organization can be achieved with one addition to the established features listed above: we introduce excitatory connections from each neuron to those in the preceding network with approximately corresponding neural sheet positions (**Fig. 1d**; see **Methods** for a complete description). That is, a neuron in network z (more ventral) with position (x, y) will excite neurons in net-

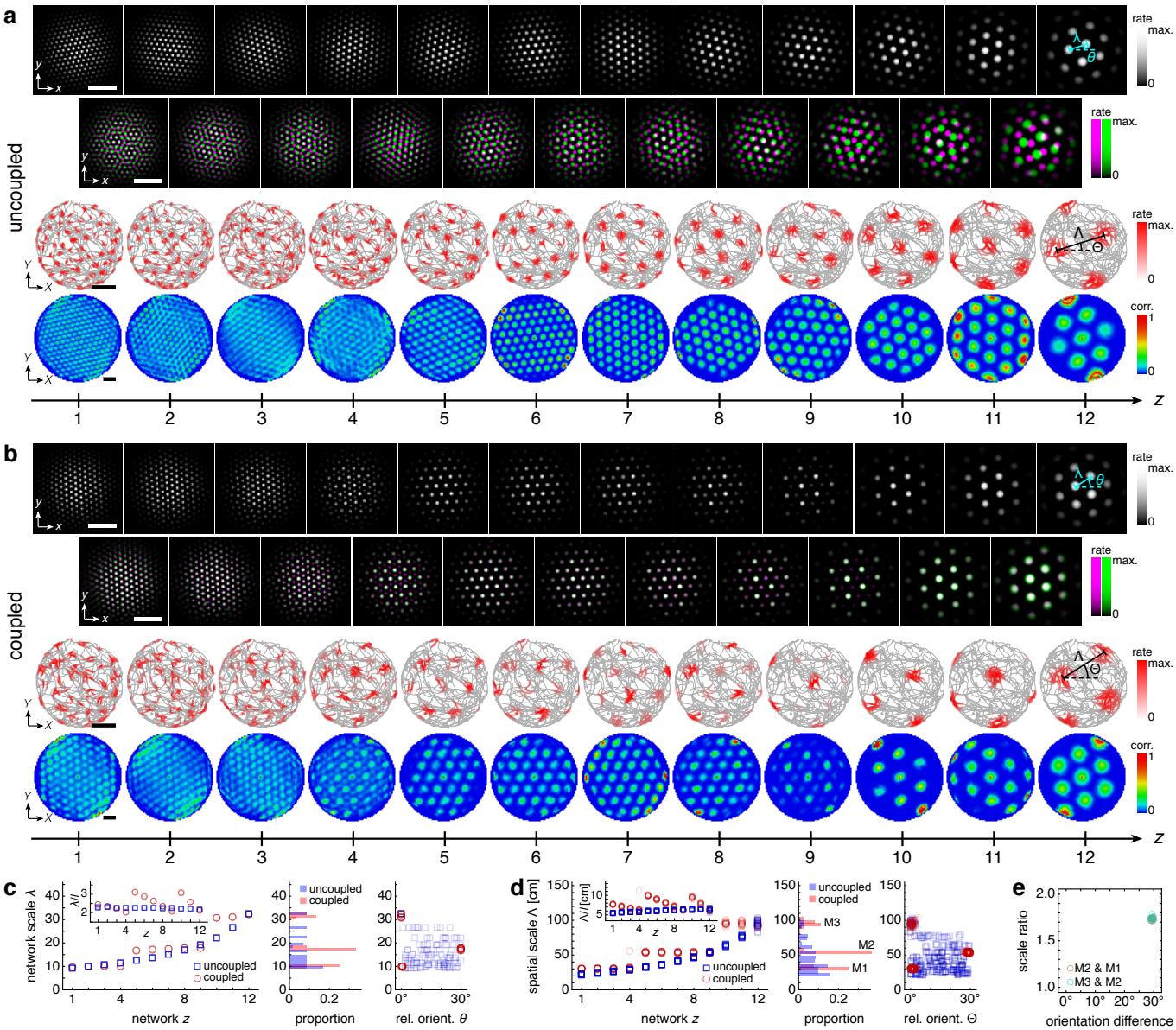

Figure 2. Coupling can induce modularity with fixed scale ratios and orientation differences. (**a**) A representative simulation without coupling. Top row: network activities at the end of the simulation. Second row: activity overlays between adjacent networks depicted in the top row. In each panel, the network with smaller (larger) z is depicted in magenta (green), so white indicates activity in both networks. Third row: spatial rate map of a single neuron for each z superimposed on the animal's trajectory. Bottom row: spatial autocorrelations of the rate maps depicted in the third row. White scale bars, 50 neurons. Black scale bars, 50 cm. (**b**) Same as **a** but for a representative simulation with coupling. (**c**–**e**) Data from 10 replicate uncoupled and coupled simulations. (**c**) Left: network grid scales $\lambda(z)$. For each network, there are 10 closely spaced red circles and 10 closely spaced blue squares corresponding to replicate simulations. Inset: $\lambda(z)$ divided by the inhibition distance $l(z)$. Middle: histogram for λ collected across all networks. Right: network grid orientations θ relative to the network in the same simulation with largest scale. (**d**) Left: spatial grid scales $\Lambda(z)$. For each z, there are up to 30 red circles and 30 blue squares corresponding to 3 neurons recorded during each simulation. Inset: $\Lambda(z)$ divided by the inhibition distance $l(z)$. Middle: histogram for Λ collected across all networks. In the coupled model, grid cells are clustered into three modules. Right: spatial grid orientations Θ relative to the grid cell in the same simulation with largest scale. (**e**) Spatial scale ratios and orientation differences between adjacent modules for the coupled model. Standard parameter values provided in **Table 1**.

work $z-1$ (more dorsal) with positions that are within a distance d of position (x, y). In other words, the distance d is the "spread" of excitatory connections, and we choose a constant value across all networks comparable to the inhibition distance $l(z)$. Similar results are obtained with dorsal-to-ventral or bidirectional excitatory coupling (below) or with a spread $d(z)$ that increases with the inhibition distance $l(z)$ (**Supp. Fig. 1**).

The self-organization of triangular grids in the neural sheet and the faithful path-integration that projects these grids onto single neuron spatial rate maps persist after introduction of inter-network coupling (**Fig. 2b**). Network and spatial scales $\lambda(z)$ and $\Lambda(z)$ still increase from network $z = 1$ (dorsal) to network $z = 12$ (ventral). Yet, **Fig. 2c**, **d** shows that for the coupled model, these scales exhibit plateaus that are interrupted by large jumps, disrupting their proportionality to inhibition distance $l(z)$, which is kept identical to that of the uncoupled system (**Fig. 1c**). Collecting scales across all networks illustrates that they cluster around certain values in the coupled system while they are smoothly distributed in the uncoupled system. We identify these clusters with modules M1, M2, and M3 of increasing scale. Note that multiple networks at various depths z can belong to the same module. Moreover, coupling causes grid cells that cluster around a certain scale to also cluster around a certain orientation (**Fig. 2c**, **d**), as seen in experiment [2]. The uncoupled system does not demonstrate co-modularity of orientation with scale, i.e., two networks with similar grid scales need not have similar orientations unless this is imposed by an external constraint.

In summary, excitatory coupling between grid attractor networks dynamically induces discreteness in grid scales that is co-modular with grid orientation, as observed experimentally [2], and as needed for even coverage of space by the grid map [10].

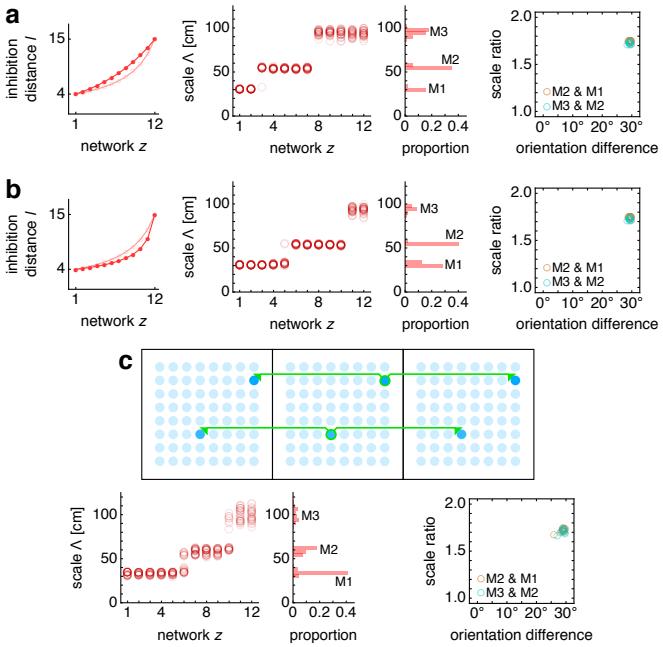

Figure 3. Modules produced by commensurate lattices are robust to parameter perturbations. Data from 10 replicate simulations in each subfigure. (**a**) Left: we use a less concave inhibition distance profile $l(z)$ (dark filled circles) compared to **Fig. 1c** (light empty circles). Middle: spatial grid scales exhibit modules when collected in a histogram across networks. Right: modules have the same scale ratios and orientation differences as in **Fig. 2e**. (**b**) Same as **a**, but with a more concave $l(z)$. (**c**) Simulations with bidirectional point-to-point coupling instead of the unidirectional spread coupling in **Fig. 1d**. Top: schematic of the neuron at position (x, y) in network z exciting only the neuron at (x, y) in networks $z-1$ and $z+1$. Bottom left/right: same as middle/right in **a**. In **a**, inhibition distance exponent $l_{\text{exp}} = 0$. In **b**, $l_{\text{exp}} = -2$. In **c**, coupling spread $d = 1$ and coupling strength $u_{\text{mag}} = 0.4$ in both directions. Other parameter values are in **Table 1**.

Modular geometry is determined by lattice geometry

Not only does excitatory coupling produce modules, it can do so with consistent scale ratios and orientation differences. For the coupled system depicted in **Fig. 2**, scale ratios and orientation differences between pairs of adjacent modules consistently take values 1.74 ± 0.02 and $29.5 \pm 0.4°$, respectively (mean $\pm$ s.d.; **Fig. 2e**). If we perturb the inhibition distance profile $l(z)$ by making it less or more concave, these scale ratios and orientation differences are unchanged (**Fig. 3a**, **b**). Concavity only affects the number of grid cells in each module, which can be tuned to match experimental observations. The same scale ratios and orientation differences also persist after changes to the directionality and spread of excitatory connections. For example, we replace the ventral-to-dorsal connections with bidirectional coupling and decrease the coupling spread d such that a neuron in network z excites only a single neuron in both networks $z-1$ and $z+1$; scale ratios and orientation differences remain at 1.7 and 30°, respectively (**Fig. 3c**). Representative network activities and single neuron rate maps for these simulations are provided in **Supp. Fig. 2**. Data for simulations with only dorsal-to-ventral connections are provided in **Supp. Fig. 3**; they also exhibit the same scale ratios and orientation differences.

We can intuitively understand this precise modularity through the competition between lateral inhibition within networks and longitudinal excitation across networks. In the uncoupled system, grid scales decrease proportionally as the inhibition distance $l(z)$ decreases from $z = 12$ to $z = 1$. However, coupling causes areas of high activity in network z to preferentially excite corresponding areas in network $z-1$, which encourages adjacent networks to share the same grid pattern. Thus, coupling adds rigidity to the system and provides an opposing "force" against the changing inhibition distance that attempts to drive changes in grid scale. This rigidity produces the plateaus in network and spatial scales $\lambda(z)$ and

$\Lambda(z)$ that delineate modules across multiple networks.

At interfaces between modules, coupling can no longer fully oppose the changing inhibition distance, and the grid pattern changes. However, the rigidity fixes a geometric relationship between the grid patterns of the two networks spanning the interface. In the coupled system of **Fig. 2**, module interfaces occur between networks $z = 4$ and 5 and between $z = 9$ and 10. The network population activity overlays of **Fig. 2b** reveal overlap of many activity peaks at these interfaces. However, the more dorsal network (with smaller z) at each interface contains additional small peaks between the shared peaks. In this way, adjacent networks still share many corresponding areas of high activity, as favored by coupling, but the grid scale changes, as favored by a changing inhibition distance. Pairs of grids whose lattice points demonstrate regular registry are called *commensurate* lattices [25] and have precise scale ratios and orientation differences, here respectively $\sqrt{3} \approx 1.7$ and 30°, which match the results in **Figs. 2e** and **3**.

In summary, excitatory coupling can compete against a changing inhibition distance to produce a rigid grid system whose "fractures" exhibit stereotyped commensurate lattice relationships. These robust geometric relationships lead to discrete modules with fixed scale ratios and orientation differences.

In our model, commensurate lattice relationships naturally lead to field-to-field firing rate variability in single neuron spatial rate maps (for example, $z = 8$ in the third row of **Fig. 2b**), another experimentally observed feature of the grid system [26, 27]. At interfaces between two commensurate lattices, only a subset of population activity peaks in the grid of smaller scale overlap with, and thus receive excitation from, those in the grid of larger scale. The network with smaller grid scale will contain activity peaks of different magnitudes; this heterogeneity is then projected onto the spatial rate maps of its neurons.

Excitation-inhibition balance sets lattice geometry

Adjusting the balance between excitatory coupling and a changing inhibition distance produces other commensurate lattice relationships, each of which enforces a certain scale ratio and orientation difference. To explore this competition systematically, we use a smaller coupled model with just two networks, $z = 1$ and 2, and vary three parameters: the coupling spread d, the coupling strength u_{mag}, and the ratio of inhibition distances between the two networks $l(2)/l(1)$ (**Supp. Info.**). For each set of parameters, we measure network scale ratios and orientation differences produced by multiple replicate simulations (**Supp. Fig. 4**). We find that as the excitation-inhibition balance is varied by changing u_{mag} and $l(2)/l(1)$, a number of discretely different relationships appear, which can be summarized in "phase diagrams" (**Fig. 4a**, **b**).

In many regions of the phase diagrams, these lattice relationships are commensurate, each with a characteristic scale ratio and orientation difference (**Fig. 4c**). When parameters are chosen near a boundary between two regions, replicate simulations may adopt either lattice relationship or occasionally be trapped in other metastable relationships due to variations in random initial conditions (**Supp. Fig. 4**). At larger u_{mag} in both phase diagrams, there are fewer regions as $l(2)/l(1)$ varies because a higher excitatory coupling strength provides more rigidity against gradients in inhibition distance (**Fig. 4a**, **b**). However, a larger coupling spread d would cause network $z = 2$ to excite a broader set of neurons in network $z = 1$, softening the rigidity imposed by coupling and producing a wider variety of lattices in **Fig. 4b** than **Fig. 4a**. Also in **Fig. 4b**, when excitation is weak and approaching the uncoupled limit, there is a noticeable region dominated by *incommensurate* lattices, in which the two grids lack consistent registry or relative orientation, and grid scale is largely determined by inhibition distance (**Supp. Fig. 4**).

Figure 4b also contains a larger region of *discommensurate* lattices (although strictly speaking, in condensed matter physics, they would be termed commensurate lattices with discommensurations [25]). Discommensurate networks have closely overlapping activities in certain areas that are separated by a mesh of regions lacking overlap called discommensurations (**Fig. 4d**). They exhibit ranges of scale ratios 1.1–1.4 and orientation differences 0°–10° that ultimately arise from a single source: the density of discommensurations, whose properties can also be explained through excitation-inhibition competition. Stronger coupling drives more activity overlap, which favors sparser discommensurations and lowers the scale ratio and orientation difference. However, a larger inhibition distance ratio drives the two networks to differ more in grid scale, which favors denser discommensurations. To better accommodate the discommensurations, grids rotate slightly as observed previously in a crystal system [28]. **Figure 4e** confirms that scale ratios and orientation differences vary together as the discommensuration density changes.

Thus, by changing the balance between excitation and inhibition, a two-network model yields geometric lattice relationships with various scale ratios and corresponding orientation differences. All of the commensurate relationships (**Fig. 4c**) and almost the entire range of discommensurate relationships (**Fig. 4d**) have scale ratios that fall in the range of experimental measurements, which is roughly 1.2–2.0 [2–4].

Discommensurate lattices produce distinct modular geometries but with more variation

As mentioned above, discommensurate lattices have a range of allowed geometries (**Fig. 4d**, **e**), but they still produce modules in a full 12-network grid system with

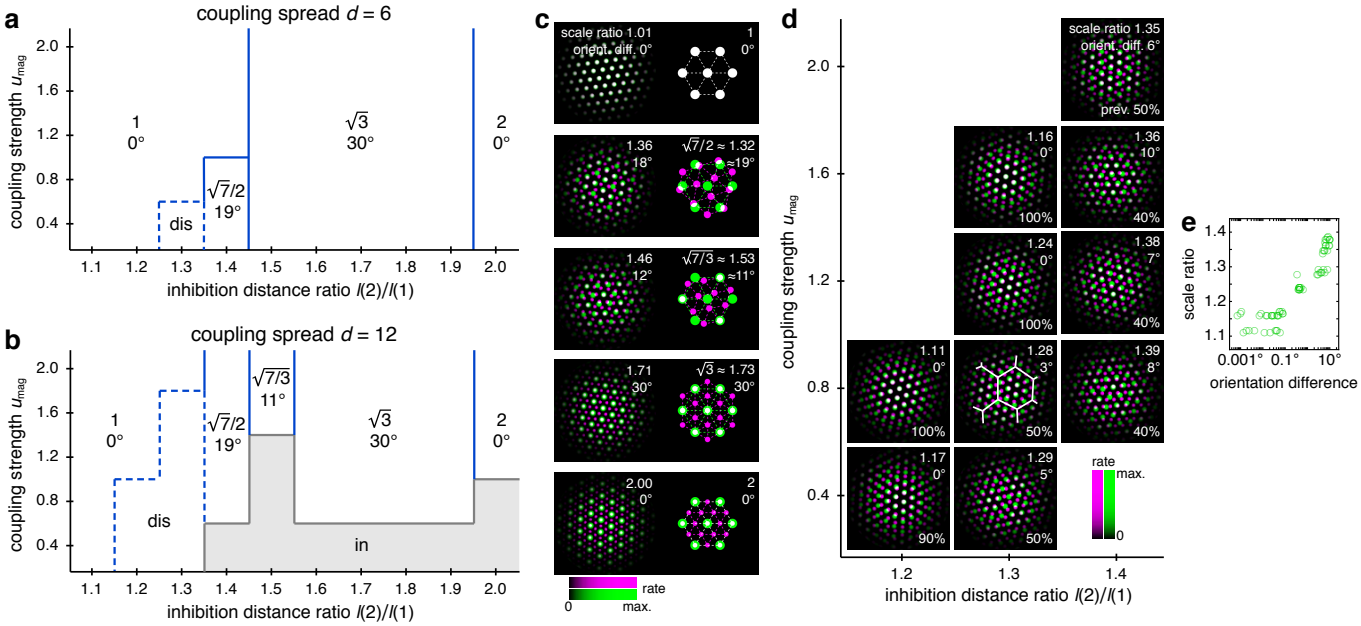

Figure 4. Diverse lattice relationships emerge over wide ranges in simulation parameters. In models with only two networks $z = 1$ and 2, we vary the coupling strength u_{mag} and the ratio of inhibition distances $l(2)/l(1)$ for two different coupling spreads d. (**a**, **b**) Approximate phase diagrams based on 10 replicate simulations for each set of parameters, with the mean of $l(1)$ and $l(2)$ fixed to be 9. The most frequently occurring scale ratio and orientation difference are indicated for each region; coexistence between multiple lattice relationships may exist at drawn boundaries. (**a**) Phase diagram for small coupling spread $d = 6$. Solid lines separate four regions with different commensurate lattice relationships labeled by scale ratio and orientation difference, and dotted lines mark one region of discommensurate lattice relationships. (**b**) Phase diagram for large coupling spread $d = 12$. There are five different commensurate regions, a discommensurate region, as well as a region containing incommensurate lattices (gray). (**c**) Network activity overlays for representative observed (left) and idealized (right) commensurate relationships. Numbers at the top right of each image indicate network scale ratios $\lambda(2)/\lambda(1)$ and orientation differences $\theta(2) - \theta(1)$. Networks $z = 1$ and 2 in magenta and green, respectively, so white indicates activity in both networks. (**d**) Expanded region of **b** displaying discommensurate lattice statistics. For each set of parameters, a representative overlay for the most prevalent discommensurate lattice relationship is shown. The number in the lower right indicates the proportion of replicate simulations with scale ratio within 0.01 and orientation difference within 3° of the values shown at top right. In one overlay, discommensurations are outlined by white lines. (**e**) The discommensurate relationships described in **d** demonstrate positive correlation between scale ratio and the logarithm of orientation difference (Pearson's $\rho = 0.91$). Parameter values provided in **Supp. Info.**

a preferred scale ratio and orientation difference. However, these values do not cluster as strongly as they do for a commensurate relationship, which is geometrically precise.

The phase diagrams of **Fig. 4** provide guidance for modifying a 12-network system that exhibits a $[\sqrt{3}, 30°]$ relationship to produce discommensurate relationships instead. We make the inhibition distance profile $l(z)$ shallower (**Fig. 5a**) and increase the coupling spread d by 50%. Network activity overlays of these new simulations reveal grids obeying discommensurate relationships (**Fig. 5b**), which are projected onto single neuron spatial rate maps through faithful path-integration (**Supp. Fig. 5**). Across replicate simulations with identical parameter values but different random initial firing rates, the discommensurate system demonstrates greater variation in scale and orientation (**Fig. 5c**) than the commensurate systems of **Figs. 2** and **3**. Nevertheless, analysis of each replicate simulation reveals clustering with well-defined modules (**Fig. 5d** and **Supp. Fig. 5**). These modules have scale ratio 1.39 ± 0.10 and orientation difference $6.7 \pm 3.5°$ (mean $\pm$ s.d.; **Fig. 5e**). The preferred scale ratio agrees well with the mean value observed experimentally in [2].

Conceptually, we can interpret the greater spread of scales and orientations in terms of coupling rigidity. Excitatory coupling, especially when the spread is larger, provides enough rigidity in the discommensurate system to cluster scale ratios and orientation differences but not enough to prevent variations in these values. The degree of variability observed in **Fig. 5c**, **d** appears consistent with experimental measurements, which also demonstrate spread [2, 3].

A few module pairs in **Fig. 5e** exhibit a large orientation difference >10°. This is not expected from a discommensurate relationship, and indeed, inspecting the network activities reveals adjacent networks trapped in a relationship with low activity overlap and large ori-

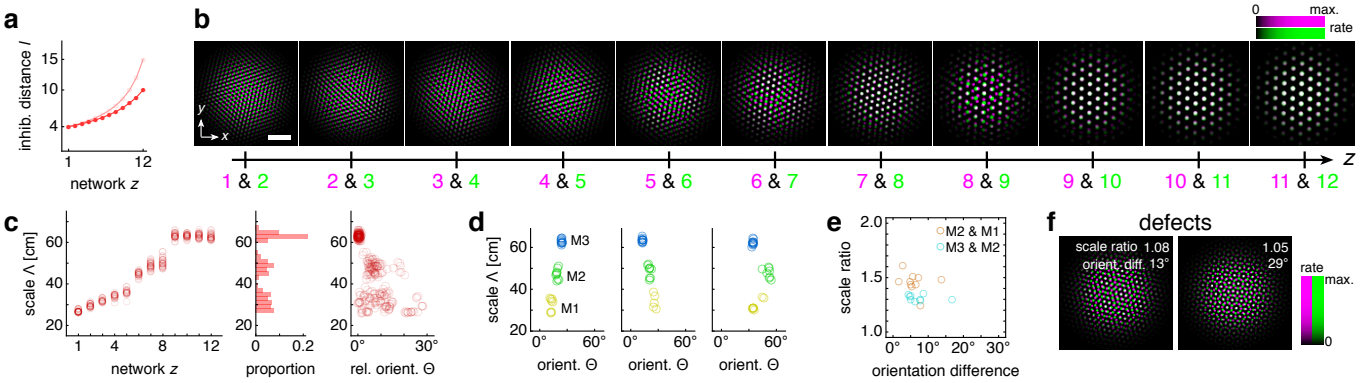

Figure 5. Discommensurate lattice relationships can produce realistic modules. (**a**) We use a shallower inhibition distance profile $l(z)$ (dark filled circles) compared to **Fig. 1c** (light empty circles). (**b**) Representative activity overlays between adjacent networks z in magenta and green, so white indicates activity in both networks. Scale bar, 50 neurons. (**c**–**e**) Data from 10 replicate simulations. (**c**) Left: spatial grid scales $\Lambda(z)$. For each network, there are up to 30 red circles corresponding to 3 neurons recorded during each simulation. Middle: histogram for Λ collected across all networks. Right: spatial orientations Θ relative to the grid cell in the same simulation with largest scale. (**d**) Clustering of spatial scales and orientations for 3 representative simulations. Due to 6-fold lattice symmetry, orientation is a periodic variable modulo 60°. Different colors indicate separate modules. (**e**) Spatial scale ratios and orientation differences between adjacent modules. (**f**) Representative activity overlays demonstrating defects with low activity overlap. Maximum inhibition distance $l_{\max} = 10$, coupling spread $d = 12$. We use larger network size $n \times n = 230 \times 230$ to allow for discommensurate relationships whose periodicities span longer distances on the neural sheets. Other parameter values are in **Table 1**.

entation difference (**Fig. 5f**). In the context of a grid system that otherwise obeys commensurate or discommensurate geometries containing more overlap, we call this less common relationship a "defect." We distinguish between these relationships and the incommensurate lattices discussed above, which also have low activity overlap. Defects arise when the excitatory coupling is strong, and incommensurate lattices arise when this coupling is weak. Also, defects have smaller scale ratios <1.1 and larger orientation differences >10°, whereas incommensurate lattices have larger scale ratios >1.3 and any orientation difference (**Fig. 4b** and **Supp. Fig. 4**).

Thus, networks governed by discommensurate relationships also cluster into modules with a preferred scale ratio and orientation difference within the experimental range [2, 4]. Due to lower coupling rigidity compared to commensurate grid systems, they exhibit increased variability and occasional defects across replicate simulations.

A diversity of lattice geometries maintains constant-on-average scale ratios

So far, each set of 12-network simulations contained replicates with identical parameter values and exhibited a single dominant lattice relationship. We now present results with different parameter values to imitate biological network variability across animals. This procedure leads to modules with different commensurate and discommensurate relationships (**Fig. 6a** and **Supp. Fig. 6**). There is no longer a single preferred scale ratio or orientation difference (**Fig. 6b**), but patterns emerge due to the predominance of discommensurate and commensurate relationships. Recall from **Fig. 5e** that discommensurate module pairs exhibit scale ratios ≈1.4 and orientation differences ≈7°. Combined with $[\sqrt{3} \approx 1.7, 30°]$ module pairs we find a bimodal distribution of orientation differences around 7° and 30°, consistent with experimental data [4], and positive correlation between scale ratio and orientation difference. Modules with low scale ratio but high orientation difference decrease this correlation; they arise from defects (**Fig. 5f**).

Scale ratios across the network variations span a range of values, but their averages are constant across module pairs. That is, the median scale ratio does not change between the pair of modules with smaller scales and the larger pair (**Fig. 6c**). Similarly, mean values are respectively 1.52 ± 0.05 and 1.53 ± 0.05 (mean ± s.e.m.) for module pairs M2 & M1 and M3 & M2. Combining data from both module pairs gives scale ratio 1.52 ± 0.03 (mean ± s.e.m.), which agrees well with the mean value of 1.56 from Ref. 4. Reference 2 reports a slightly smaller mean value of 1.42 ± 0.17 (mean ± s.d.; re-analyzed by Ref. 8), but its broad distribution of scale ratios overlaps considerably with ours. Moreover, we find that the normalized scale *difference* does change its median across module pairs (**Fig. 6d**). This result that scale ratios are constant on average but scale differences are not matches experiment [2].

Thus, although our model can produce modules with fixed scale ratios, allowing for a range of network parameters also produces modules with a range of scale ratios. Nevertheless, the scale ratio averaged over these parameters is still constant across module pairs, a key feature of the grid system that holds even if scales are not governed

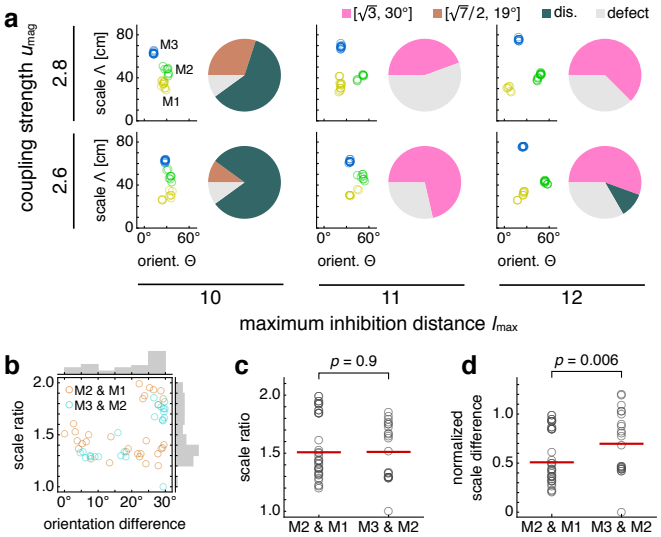

Figure 6. Simulations spanning different parameters contain diversity in lattice relationships, but average scale ratios are still constant between module pairs. Data from 5 replicate simulations for each set of parameters. (**a**) Clustering of spatial scales and orientations for one representative simulation (left) and lattice relationship distribution across all pairs of adjacent modules (right) for each set of parameters. (**b**) Spatial scale ratios and orientation differences between adjacent modules with respective histograms to the right and above. Scale ratios and orientation differences exhibit positive rank correlation (Spearman's $\rho = 0.44$, $p = 0.001$). (**c**) Spatial scale ratios. Means indicated by lines. Medians compared through the Mann-Whitney U test with reported p-value. (**d**) Spatial scale differences normalized by the scale of the first module (M1) in each simulation. Same interpretation of lines and p-value as in **c**. The $u_{\mathrm{mag}} = 2.6$ and $l_{\max} = 10$ data are taken from simulations in **Fig. 5**. Some simulations produced only two modules M1 and M2; one simulation produced four modules, and M4 was excluded from further analysis (**Supp. Fig. 6**). Coupling spread $d = 12$ and network size $n \times n = 230 \times 230$. Other parameter values are in **Table 1**.

by a universal ratio [2].

Testing for coupling with a mock lesion experiment

Excitatory coupling locks networks into scales and orientations imposed by more ventral networks. Disrupting the coupling frees networks from this rigidity, which can change scales and orientations far from the disruption. We demonstrate this effect by inactivating one network $z = 7$ midway through the simulation (**Fig. 7a**). This corresponds experimentally to disrupting all excitatory connections at one location along the dorsoventral MEC axis.

After the lesion, grid cells ventral to the lesion location ($z \geq 8$) are unaffected, but those dorsal to the lesion location ($z \leq 6$) change scale and orientation and form a single module (**Fig. 7b–d**). Network $z = 6$ is no longer constrained by larger grids of more ventral networks, so its scale decreases. The coupling that remains from $z = 6$ to 1 then rigidly propagates the new grid down to network $z = 1$. This post-lesion module M1 has larger scale and 30° orientation difference compared to the pre-lesion M1; these changes also appear as corresponding changes in the scale ratio and orientation difference between modules M2 and M1 (**Fig. 7e**).

Immediate changes in grid scale and/or orientation observed at one location along the longitudinal MEC axis due to a lesion at another location would strongly support the presence of the excitatory coupling predicted by our model. Moreover, the anatomical distribution of the changes would indicate the directionality of coupling; those in grid cells dorsal to the lesion would indicate ventral-to-dorsal coupling and those ventral to the lesion would indicate dorsal-to-ventral coupling.

DISCUSSION

We propose that the hierarchy of grid modules in the MEC is self-organized by competition in attractor networks between excitation along the longitudinal MEC axis and lateral inhibition. We showed that such an architecture, with an inhibition length scale that increases smoothly along the MEC axis, reproduces a central experimental finding: grid cells form modules with scales clustered around discrete values [2–4].

The distribution of scales across modules in our model quantitatively matches experiments. Different groups have reported mean scale ratios of 1.64 (6 module pairs), 1.42 (24 module pairs), and 1.56 (11 module pairs) [2–4]. These data could be interpreted as an indication that the grid system has a preferred scale ratio roughly in range of 1.4–1.7. As we showed, our model naturally produces a hierarchy of modules with scale ratios in this range; its network parameters lead to both commensurate and discommensurate grids (**Fig. 4**). On the other hand, the data on scale ratios between individual pairs of modules actually span a range of values in the different experiments: 1.6–1.9, 1.1–1.8, and 1.2–2.0 [2–4]. This suggests that the underlying mechanism that produces grid modules must be capable of producing different scale ratios as its parameters vary. This is indeed the case for our model, in which variation of network parameters produces a realistic range of scale ratios (**Fig. 6**). Despite variability across individual scale ratios, experiments strikingly reveal that the average scale ratio is the same from the smallest pair of modules to the largest pair, whereas the average scale *difference* changes across the hierarchy [2]. Our model robustly reproduces this observation (**Fig. 6c**, **d**) because its fundamental mechanism of geometric coordination between grids enforces constant-on-average scale ratios even with variation in parameters among individual networks.

Our model requires that grid orientation be co-modular with scale, as observed in experiment [2]. Stud-

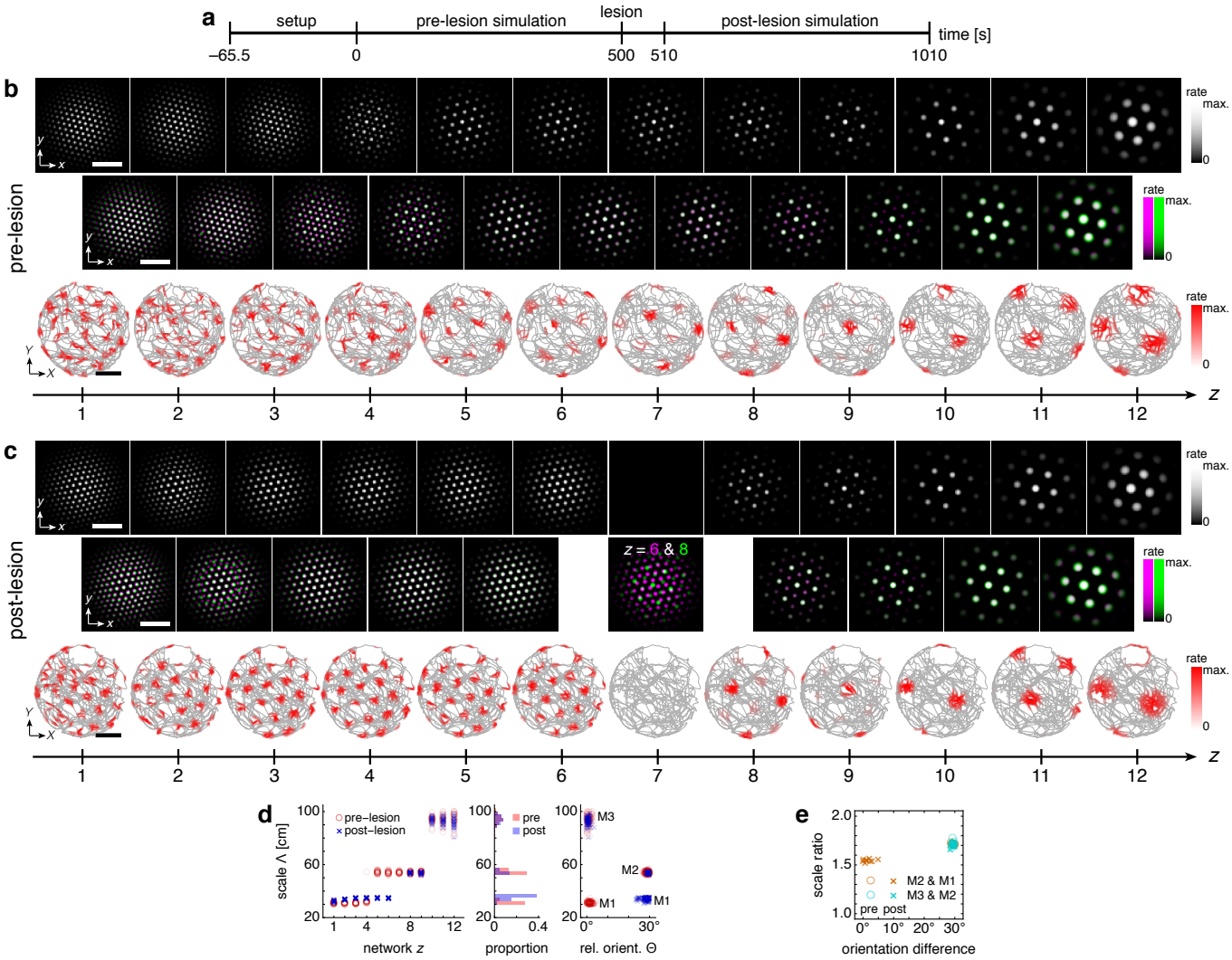

Figure 7. Lesioning a network changes grid scales and orientations of more dorsal networks. (**a**) Lesion protocol. (**b**) A representative simulation before the lesion. Top row: network activities at the end of the pre-lesion simulation. Second row: activity overlays between adjacent networks depicted in the top row. In each panel, the network with smaller (larger) z is depicted in magenta (green), so white indicates activity in both networks. Third row: spatial rate map of a single neuron for each z superimposed on the animal's trajectory. White scale bars, 50 neurons. Black scale bars, 50 cm. (**c**) Same as **b** but after the lesion. Spatial rate maps are recorded from the same neurons as in **b**. (**d**, **e**) Data from 10 replicate simulations. (**d**) Left: spatial grid scales $\Lambda(z)$ before and after the lesion. Middle: histogram for Λ collected across all networks. Right: spatial orientations Θ relative to the grid cell in the same simulation with largest scale. (**e**) Spatial scale ratios and orientation differences between adjacent modules. Standard parameter values provided in **Table 1**.

ies characterizing the statistics of orientation differences between modules are limited, but values seem to span the entire range 0°–30°, with some preference for values at the low and high ends of this range [4]. Our model can capture the entire range of orientation differences with discommensurate relationships favoring small differences and commensurate relationships favoring large differences (**Fig. 4**). Overall, our model predicts a positive correlation between scale ratio and orientation difference (**Figs. 4e** and **6b**), which can be tested experimentally. Existing datasets [2, 4] have a confound—animals are tested in square and rectangular enclosures which have distinguishable orientations marked by the corners. Grid orientations can anchor to such features [29], either through the integration of visual and external cues [30, 31], or through interaction with boundaries [19, 32–37]. Experiments in circular or other non-rectangular environments may help disambiguate the effects of such anchoring. Our model also predicts that orientation differences between modules will be preserved between environments with different geometries since the differences are internally generated by the dynamics of the network. This effect has been observed [4].

Our model produces consistent differences in firing rate

from one grid field to another for some grid cells. This variability arises at module interfaces from the selective excitation of some network activity peaks in the smaller-scale grid by the overlapping activity peaks of the larger-scale grid. Such an explanation for firing rate variability is suggested by Ref. 26 and would be further supported by observing spatial periodicity in the variability corresponding to the scale of the larger grids. An alternative model, in which field-to-field firing rate variability arises from place cell inputs [27], would not lead to such periodicity.

Our model requires excitatory coupling between grid cells at different locations along the longitudinal MEC axis, either through direct excitation or disinhibition [38]. As a result, it predicts that destruction of grid cells, or inactivation of excitatory coupling [39], at a given location along the axis will change grid scales and/or orientations at other locations (**Fig. 7**). The presence of noise correlations across modules, as previously investigated but not fully characterized [40, 41], would suggest connections between modules. Such correlations, and perhaps even lattice relationships, could be observed via calcium imaging of the MEC [42, 43]. A direct test for coupling would involve patch clamp experiments akin to those used to identify local inhibition and excitation and interhemispheric excitation between principal cells in superficial MEC layers [38, 44, 45].

Since spatial grid scales are both proportional to inhibition length scale l and inversely proportional to velocity gain α (Ref. 16 and **Methods**), we also simulated excitatorily coupled networks with a depth-dependent velocity gain $\alpha(z)$ and a fixed inhibition distance l (**Supp. Info.**). In contrast to simulations in one dimension [46], while we observed module self-organization, the system gave inconsistent results among replicate simulations and lacked fixed scale ratios (**Supp. Figs. 7** and **8** and **Supp. Video**). Moreover, recent calcium imaging experiments suggest that activity on the MEC is arranged a deformed triangular lattice [43], as predicted by the continuous attractor model [16], and that regions with activity separated by larger anatomic distances contain grid cells of larger spatial scale. These observations support a changing inhibition length scale over a changing velocity gain as a mechanism for producing different grid scales, under the assumption that anatomic and network distances correspond to each other.

Our results differ from previous work on mechanisms for forming grid modules. Grossberg and Pilly hypothesize that grid cells arise from stripe cells in parasubiculum, and that discreteness in the spatial period of stripe cells leads to modularity of grid cells [22]. However, stripe cells have only been observed once [47, 48], and the origin of discrete periods with constant-on-average ratios in stripe cells would then need to be addressed. Urdapilleta, Si, and Treves propose a model in which discrete modules self-organize from smooth gradients in parameters in a model where grid formation is driven by firing rate adaptation in single cells [23]. They also utilize excitatory coupling among grid cells along the longitudinal MEC axis. However, this model does not have a mechanism to dynamically enforce the average constancy of grid scale ratios, which is a feature of the grid system [2]. The model also does not demonstrate modules with orientation differences near 30° [4]. Our model naturally reproduces these features of the grid system. Furthermore, over the past few years, multiple reports have provided independent experimental support for the importance of recurrent connections among grid cells [38, 39, 44, 49] and for the continuous attractor model in particular [42, 43, 50]. Our work establishes that continuous attractor networks can produce a discrete hierarchy of modules with a constant-on-average scale ratio.

The competition generated between excitatory and inhibitory connections bears a strong resemblance to the Frenkel-Kontorova model of condensed matter physics, in which a periodic potential of one scale acts on particles that prefer to form a lattice of a different, competing scale [51]. This model has a rich literature with many deep theoretical results, including the calculation of complicated phase diagrams involving "devil's staircases" [25, 52] which mirror those of our model (**Fig. 4**). Under certain conditions, our model produces networks with quasicrystalline approximant grids that are driven by networks with standard triangular grids at other scales (**Supp. Fig. 9**). Quasicrystalline order lacks periodicity, but contains more nuanced positional order [53]. This phenomenon wherein quasicrystalline structure is driven by crystalline order in a coupled system was recently observed for the first time in thin-film materials that contain Frenkel-Kontorova-like interactions [54–56].

Commensurate and discommensurate lattice relationships are a robust and versatile mechanism for self-organizing a grid system whose scale ratios are constant or constant on average across a hierarchy of modules. We demonstrated this mechanism in a basic extension of the continuous attractor model with excitatory connections between networks. This model is amenable to extensions that capture other features of the grid system, such as spiking dynamics, learning of synaptic weights [57], the union of our separate networks into a single network spanning the entire MEC, and the addition of border cell inputs or recurrent coupling between modules to correct path-integration errors or react to environmental deformations [36, 37, 58–60].

ACKNOWLEDGMENTS

We are grateful to Xue-Xin Wei, Tom Lubensky, Ila Fiete, and John Widloski for their thoughtful ideas and suggestions, and to the Honda Research Institute and the NSF (grant PHY-1734030) for research support. L.K. is also supported by the Miller Institute for Basic Research in Science. Work on this project at the Aspen Center for Physics was supported by NSF grant PHY-1607611.

Table 1. Main model parameters and their values unless otherwise noted.

Parameter	Variable	Value
Number of networks	h	12
Number of neurons per network	$n \times n$	160×160
Neurons recorded per network		3
Animal speed	$\lvert\mathbf{V}\rvert$	0–1 m/s
Diameter of enclosure		180 cm
Simulation time		500 s
Simulation timestep	Δt	1 ms
Neural relaxation time	τ	10 ms
Hippocampal input strength	a_{mag}	1
Hippocampal input falloff	a_{fall}	4
Inhibition distance minimum	l_{min}	4
Inhibition distance maximum	l_{max}	15
Inhibition distance exponent	l_{exp}	-1
Inhibition strength	w_{mag}	2.4
Subpopulation shift	ξ	1
Coupling spread	d	8
Coupling strength	u_{mag}	2.6
Velocity gain	α	0.3 s/m

METHODS

Model setup and dynamics

We implemented the Burak-Fiete model as follows [16]. Networks $z = 1, \ldots, h$ each contain a 2D sheet of neurons with indices $\mathbf{r} = (x, y)$, where $x = 1, \ldots, n$ and $y = 1, \ldots, n$. Neurons receive broad excitatory input $a(\mathbf{r})$ from the hippocampus, and, to prevent edge effects, those toward the center of the networks receive more excitation than those toward the edges. Each neuron also inhibits others that lie around a length scale of $l(z)$ neurons away in the same network z Moreover, every neuron belongs to one of four subpopulations that evenly tile the neural sheet. Each subpopulation is associated with both a preferred direction $\hat{\mathbf{e}}$ along one of the network axes $\pm\hat{\mathbf{x}}$ or $\pm\hat{\mathbf{y}}$ and a corresponding preferred direction $\hat{\mathbf{E}}$ along an axis $\pm\hat{\mathbf{X}}$ or $\pm\hat{\mathbf{Y}}$ in its spatial environment. A neuron at position $\mathbf{r}$ in network z has its inhibitory outputs $w(\mathbf{r}; z)$ shifted slightly by ξ neurons in the $\hat{\mathbf{e}}(\mathbf{r})$ direction and its hippocampal excitation modulated by a small amount proportional to $\hat{\mathbf{E}}(\mathbf{r}) \cdot \mathbf{V}$, where $\mathbf{V}$ is the spatial velocity of the animal. Note that lowercase letters refer to attractor networks at each depth z in which distances have units of neurons, and uppercase letters refer to the animal's spatial environment in which distances have physical units, such as centimeters.

In addition to these established features [16], we introduce excitatory connections $u(\mathbf{r})$ from every neuron $\mathbf{r}$ in network z to neurons located within a spread d of the same $\mathbf{r}$ but in the preceding network with depth $z - 1$. $u(\mathbf{r})$ is constant for all networks except for the last one $z = h$, which has $u(\mathbf{r}) = 0$. These components lead to the following dynamical equation for the dimensionless neural firing rates $s(\mathbf{r}, z, t)$:

$$\begin{aligned}\tau \frac{s(\mathbf{r}, z, t + \Delta t) - s(\mathbf{r}, z, t)}{\Delta t} + s(\mathbf{r}, z, t) \\ = \Bigg\{ \sum_{\mathbf{r}'} w(\mathbf{r} - \mathbf{r}' + \xi \hat{\mathbf{e}}(\mathbf{r}'); z) s(\mathbf{r}', z, t) \\ + \sum_{\mathbf{r}'} u(\mathbf{r} - \mathbf{r}') s(\mathbf{r}', z + 1, t) \\ + a(\mathbf{r}) \left[1 + \alpha \hat{\mathbf{E}}(\mathbf{r}) \cdot \mathbf{V}(t) \right] \Bigg\}_+ . \end{aligned} \tag{1}$$

Inputs to each neuron are rectified by $\{c\}_+ = 0$ for $c < 0$, c for $c \geq 0$. Δt is the simulation time increment, τ is the neural relaxation time, and α is the velocity gain that describes how much the animal's velocity $\mathbf{V}$ modulates the hippocampal inputs $a(\mathbf{r})$. Note that s can be treated as a dimensionless variable because **Eq. 1** is invariant to scaling of s and a by the same factor.

We use velocities $\mathbf{V}(t)$ corresponding to a real rat trajectory [1, 16]. Details are provided in **Supp. Info.**

Inhibitory and excitatory connections

The hippocampal input is

$$a(\mathbf{r}) = \begin{cases} a_{\mathrm{mag}} \mathrm{e}^{-a_{\mathrm{fall}} r_{\mathrm{scaled}}^2} & r_{\mathrm{scaled}} < 1 \\ 0 & r_{\mathrm{scaled}} \geq 1, \end{cases} \tag{2}$$

where $r_{\mathrm{scaled}} = \sqrt{\left(x - \frac{n+1}{2}\right)^2 + \left(y - \frac{n+1}{2}\right)^2} / \frac{n}{2}$ is a scaled radial distance for the neuron at $\mathbf{r} = (x, y)$, a_{mag} is the magnitude of the input, and a_{fall} is a falloff parameter. The inhibition distance for network z is

$$l(z) = \left[l_{\mathrm{min}}^{l_{\mathrm{exp}}} + \left(l_{\mathrm{max}}^{l_{\mathrm{exp}}} - l_{\mathrm{min}}^{l_{\mathrm{exp}}} \right) \frac{z - 1}{h - 1} \right]^{1/l_{\mathrm{exp}}}, \tag{3}$$

which ranges from $l_{\mathrm{min}} = l(1)$ to $l_{\mathrm{max}} = l(h)$ with concavity tuned by l_{exp}. More negative values of l_{exp} lead to greater concavity; for $l_{\mathrm{exp}} = 0$, we use the limiting expression $l(z) = l_{\mathrm{min}}^{(h-z)/(h-1)} l_{\mathrm{max}}^{(z-1)/(h-1)}$. The recurrent inhibition profile for network z is

$$w(\mathbf{r}; z) = \begin{cases} -\frac{w_{\mathrm{mag}}}{l(z)^2} \frac{1 - \cos[\pi r / l(z)]}{2} & r < 2l(z) \\ 0 & r \geq 2l(z), \end{cases} \tag{4}$$

where w_{mag} is the magnitude of inhibition. We scale this magnitude by $l(z)^{-2}$ to make the integrated inhibition constant across z. The excitatory coupling is

$$u(\mathbf{r}) = \begin{cases} \frac{u_{\mathrm{mag}}}{d^2} \frac{1 + \cos[\pi r / d]}{2} & r < d \\ 0 & r \geq d, \end{cases} \tag{5}$$

where u_{mag} and d are the magnitude and spread of coupling, respectively. In analogy to w_{mag}, we scale u_{mag} by d^{-2}.

Overview of data analysis techniques

To determine spatial grid scales, orientations, and gridness, we consider an annular region of the spatial autocorrelation map that contains the 6 peaks closest to the origin. Grid scale is the radius with highest value, averaging over angles. Grid orientation and gridness are determined by first averaging over radial distance and analyzing the sixth component of the Fourier series with respect to angle [61]. The power of this component divided by the total Fourier power measures "gridness" and its complex phase measures the orientation. Grid cells are subject to a gridness cutoff of 0.6. For each replicate simulation, we cluster its grid cells with respect to scale and orientation using a k-means procedure with k determined by kernel smoothed densities [2]. See **Supp. Info.** for full details.

[1] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser, Nature **436**, 801 (2005).
[2] H. Stensola, T. Stensola, T. Solstad, K. Frøland, M.-B. Moser, and E. I. Moser, Nature **492**, 72 (2012).
[3] C. Barry, R. Hayman, N. Burgess, and K. J. Jeffery, Nat. Neurosci. **10**, 682 (2007).
[4] J. Krupic, M. Bauza, S. Burton, C. Barry, and J. O'Keefe, Nature **518**, 232 (2015).
[5] E. I. Moser, E. Kropff, and M.-B. Moser, Annu. Rev. Neurosci. **31**, 69 (2008).
[6] I. R. Fiete, Y. Burak, and T. Brookings, J. Neurosci. **28**, 6858 (2008).
[7] A. Mathis, A. V. M. Herz, and M. Stemmler, Neural Computation **24**, 2280 (2012).
[8] X.-X. Wei, J. Prentice, and V. Balasubramanian, eLife (2015).
[9] M. Stemmler, A. Mathis, and A. V. M. Herz, Science Advances **1**, e1500816 (2015).
[10] A. Sanzeni, V. Balasubramanian, G. Tiana, and M. Vergassola, Phys. Rev. E **94**, 599 (2016).
[11] N. Mosheiff, H. Agmon, A. Moriel, and Y. Burak, PLOS Comp. Biol. **13**, e1005597 (2017).
[12] S. Sreenivasan and I. Fiete, Nat. Neurosci. **14**, 1330 (2011).
[13] C. J. Cueva and X.-X. Wei, International Conference on Learning Representations (2018).
[14] A. Banino, C. Barry, B. Uria, C. Blundell, T. Lillicrap, P. Mirowski, A. Pritzel, M. J. Chadwick, T. Degris, J. Modayil, *et al.*, Nature , 1 (2018).
[15] M. C. Fuhs and D. S. Touretzky, J. Neurosci. **26**, 4266 (2006).
[16] Y. Burak and I. R. Fiete, PLOS Comp. Biol. **5**, e1000291 (2009).
[17] N. Burgess, C. Barry, and J. O'Keefe, Hippocampus **17**, 801 (2007).
[18] M. E. Hasselmo, L. M. Giocomo, and E. A. Zilli, Hippocampus **17**, 1252 (2007).
[19] D. Bush and N. Burgess, J. Neurosci. **34**, 5065 (2014).
[20] N. P. Issa, A. Rosenberg, and T. R. Husson, J. Neurophysiol. **99**, 2745 (2008).
[21] J. S. Taube, R. U. Muller, and J. B. Ranck, J. Neurosci. **10**, 420 (1990).
[22] S. Grossberg and P. K. Pilly, PLOS Comp. Biol. **8**, e1002648 (2012).
[23] E. Urdapilleta, B. Si, and A. Treves, Hippocampus **77**, 137 (2017).
[24] T. Bonnevie, B. Dunn, M. Fyhn, T. Hafting, D. Derdikman, J. L. Kubie, Y. Roudi, E. I. Moser, and M.-B. Moser, Nat. Neurosci. **16**, 309 (2013).
[25] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, 1995) pp. 601–620.
[26] R. Ismakov, O. Barak, K. Jeffery, and D. Derdikman, Curr. Biol. **27**, 2337 (2017).
[27] B. Dunn, D. Wennberg, Z. Huang, and Y. Roudi, bioRxiv 10.1101/101899 (2017).
[28] J. A. Wilson, J. Phys.: Condens. Matter **2**, 1683 (1990).
[29] T. Stensola, H. Stensola, M.-B. Moser, and E. I. Moser, Nature **518**, 207 (2015).
[30] F. Raudies and M. E. Hasselmo, PLOS Comp. Biol. **11**, e1004596 (2015).
[31] F. Savelli, J. D. Luck, and J. J. Knierim, eLife **6** (2017).
[32] J. Krupic, M. Bauza, S. Burton, and J. O'Keefe, J. Physiol. **594**, 6489 (2016).
[33] L. M. Giocomo, J. Physiol. **594**, 6501 (2016).
[34] T. Evans, A. Bicanski, D. Bush, and N. Burgess, J. Physiol. **594**, 6535 (2016).
[35] K. Hardcastle, N. Maheswaranathan, S. Ganguli, and L. M. Giocomo, Neuron **94**, 375 (2017).
[36] A. T. Keinath, R. A. Epstein, and V. Balasubramanian, eLife **7**, 71 (2018).
[37] S. A. Ocko, K. Hardcastle, L. M. Giocomo, and S. Ganguli, Proc. Natl. Acad. Sci. U. S. A. **115**, E11798 (2018).
[38] E. C. Fuchs, A. Neitz, R. Pinna, S. Melzer, A. Caputi, and H. Monyer, Neuron **89**, 194 (2016).
[39] I. Zutshi, M. L. Fu, V. Lilascharoen, J. K. Leutgeb, B. K. Lim, and S. Leutgeb, Nat. Commun. **9**, 3701 (2018).
[40] A. Mathis, A. V. M. Herz, and M. B. Stemmler, Physical Review E **88**, 022713 (2013).
[41] G. Tocker, O. Barak, and D. Derdikman, Hippocampus **25**, 1599 (2015).
[42] J. G. Heys, K. V. Rangarajan, and D. A. Dombeck, Neuron **84**, 1079 (2014).
[43] Y. Gu, S. Lewallen, A. A. Kinkhabwala, C. Domnisoru, K. Yoon, J. L. Gauthier, I. R. Fiete, and D. W. Tank, Cell **175**, 736 (2018).
[44] J. J. Couey, A. Witoelar, S.-J. Zhang, K. Zheng, J. Ye, B. Dunn, R. Czajkowski, M.-B. Moser, E. I. Moser, Y. Roudi, and M. P. Witter, Nat. Neurosci. **16**, 318 (2013).
[45] J. Winterer, N. Maier, C. Wozny, P. Beed, J. Breustedt, R. Evangelista, Y. Peng, T. D'Albis, R. Kempter, and D. Schmitz, Cell Reports **19**, 1110 (2017).
[46] J. Widloski and I. R. Fiete, Personal communication (2017).
[47] J. Krupic, N. Burgess, and J. O'Keefe, Science **337**, 853 (2012).

[48] Z. Navratilova, K. B. Godfrey, and B. L. McNaughton, J. Neurophysiol. **115**, 992 (2016).
[49] B. Dunn, M. Mørreaunet, and Y. Roudi, PLOS Comp. Biol. **11**, e1004052 (2015).
[50] K. Yoon, M. A. Buice, C. Barry, R. Hayman, N. Burgess, and I. R. Fiete, Nat. Neurosci. **16**, 1077 (2013).
[51] T. Kontorova and J. Frenkel, Zh. Eksp. Teor. Fiz. **8**, 1340 (1938).
[52] P. Bak, Rep. Prog. Phys. **45**, 587 (1982).
[53] D. Levine and P. J. Steinhardt, Phys. Rev. B **34**, 596 (1986).
[54] S. Förster, K. Meinel, R. Hammer, M. Trautmann, and W. Widdra, Nature **502**, 215 (2013).
[55] S. Förster, M. Trautmann, S. Roy, W. A. Adeagbo, E. M. Zollner, R. Hammer, F. O. Schumann, K. Meinel, S. K. Nayak, K. Mohseni, W. Hergert, H. L. Meyerheim, and W. Widdra, Phys. Rev. Lett. **117**, 1260 (2016).
[56] M. Paßens, V. Caciuc, N. Atodiresei, M. Feuerbacher, M. Moors, R. E. Dunin-Borkowski, S. Blügel, R. Waser, and S. Karthäuser, Nat. Commun. **8**, 15367 (2017).
[57] J. Widloski and I. R. Fiete, Neuron **83**, 481 (2014).
[58] K. Hardcastle, S. Ganguli, and L. M. Giocomo, Neuron **86**, 827 (2015).
[59] E. Pollock, N. Desai, X. Wei, and V. Balasubramanian, in *Cosyne Abstracts 2017* (Salt Lake City, UT, USA, 2017).
[60] N. Mosheiff and Y. Burak, in *Cosyne Abstracts 2018* (Denver, CO, USA, 2018).
[61] S. N. Weber and H. Sprekeler, PLOS Comp. Biol. **15**, e1006804 (2019).