Adult dentate gyrus neurogenesis: a functional model

Olivia Gozel* and Wulfram Gerstner

Laboratory of Computational Neuroscience,
School of Life Sciences and School of Computer and Communication Sciences,
Brain-Mind Institute,
Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

1 Abstract

In adult dentate gyrus neurogenesis, the link between maturation of newborn neurons and their function, such as behavioral pattern separation, has remained puzzling. By analyzing a theoretical model, we show that the switch from excitation to inhibition of the GABAergic input onto maturing newborn cells is crucial for their proper functional integration. When the GABAergic input is excitatory, cooperativity drives the growth of synapses such that newborn cells become sensitive to stimuli similar to those that activate mature cells. When GABAergic input switches to inhibitory, competition pushes the configuration of synapses onto newborn cells towards representations that are away from previously stored memories. This enables the maturing newborn cells to code for concepts that are novel, yet similar to familiar ones. Our theory of newborn cell maturation explains how adult newborn cells integrate into the preexisting network and why they promote separation of similar patterns.

2 Introduction

In the adult mammalian brain, neurogenesis, the production of new neurons, is restricted to a few brain areas, such as the olfactory bulb and the dentate gyrus [1]. The dentate gyrus is the entry point of input from cortex, primarily entorhinal cortex (EC), to the hippocampus [2], which is believed to be a substrate of learning and memory. Adult-born cells in dentate gyrus mostly develop into dentate granule cells (DGCs), the main excitatory cells that project to area CA3 [1].

The properties of rodent adult-born DGCs change as a function of their maturation stage, until they become indistinguishable from other mature DGCs at approximately 8 weeks [1, 3] (Fig. 1b). Many of them die before they fully mature [4]. Their survival is experience-dependent, and relies upon NMDA receptor activation [5]. Initially, newborn DGCs have enhanced excitability [6, 7] and stronger synaptic plasticity than mature DGCs, reflected by a larger LTP amplitude and a lower threshold for induction of LTP [6, 8]. Furthermore, after 4 weeks of maturation adult-born DGCs have only weak connections to interneurons, while at 7 weeks of age their activity causes indirect inhibition of mature DGCs [9].

Newborn DGCs receive no direct connections from mature DGCs [10], but are indirectly activated via interneurons [10, 11]. During maturation, the γ-aminobutyric acid (GABAergic) input from interneurons to adult-born DGCs switches from excitatory in the early phase to inhibitory in the late phase of maturation [1] (Fig. 1b). Analogous to a similar transition during embryonic and early postnatal stages [12], this transition is caused by a change in the expression profile of chloride cotransporters, from NKCC1 in the early phase to KCC2 in the late phase [13, 14]. Importantly, it has been shown that GABAergic inputs are crucial for the integration of newborn DGCs into the preexisting circuit [10, 11, 14].

Adult-born DGCs are preferentially reactivated by stimuli similar to the ones they experienced during their early phase of maturation (up to 3 weeks after cell birth) [15]. Even though the amount of newly generated cells per month is rather low (3 to 6% of the total DGCs population [16, 17]), adult-born DGCs are critical for behavioral pattern separation [18–20], in particular in tasks where similar stimuli or contexts have to be discriminated [18, 19]. However, the functional role of adult-born DGCs is controversial [21, 22]. One view is that newborn DGCs contribute to pattern separation through a modulatory role [21]. Another view suggests that newborn DGCs act as encoding units that become sensitive to features of the environment which they encounter during a critical window of maturation [15, 23]. Some authors have even challenged the role of newborn DGCs in pattern separation in the classical sense and have proposed a pattern integration effect instead [22].

To address this controversy, we present a model of how newborn DGCs integrate into the preexisting circuit. Our model suggests that the switch from lateral excitation to lateral inhibition during the maturation of newborn DGCs plays a crucial role for their proper integration into the preexisting network of mature DGCs. In addition, it explains why newborn DGCs favor pattern separation of similar stimuli, but do not impact pattern separation of distinct stimuli.

3 Results

We ask how adult-born DGCs are integrated in the network of mature cells. To address this question, we study the hypothesis that functional integration of newborn cells requires the two-step maturation process caused by the switch of GABA from excitation to inhibition. Since excitatory GABAergic input potentially increases cooperativity within the dentate gyrus network, we predict that newborn DGCs respond to familiar stimuli during the early phase of maturation, but not during the late phase, when inhibitory GABAergic input leads to competition.

To test this prediction, we pretrain a model network of 100 mature DGCs with input from 144 EC cells to respond to patterns representing two different digits from MNIST, a standard data set in artificial intelligence (Methods). Below, we present results for a representative combination of three digits (digits 3, 4 and 5), but other combinations of digits have also been tested (Suppl. Table S1). Even though we do not expect EC neurons to show a 2-dimensional arrangement, the use of 2-dimensional patterns provides a simple way to visualize the activity of all 144 EC neurons in our model (Fig. 1d). Pretraining is based on a synaptic plasticity rule which combines LTP and LTD [24, 25] (Methods, Fig. 1c).

Different prototypes are represented by mature neurons

After pretraining with patterns from digits 3 and 4, we examine the receptive field of each DGC, consisting of the connections from all 144 EC neurons onto one DGC. We observe that out of the 100 DGCs, some have developed a receptive field that corresponds to digit 3, others a receptive field that corresponds to digit 4, and the remaining ones stay unselective (Fig. 2a). We classify the DGCs with unselective receptive fields as unresponsive units. The selective DGCs represent different prototypes of the two digits, reflected by different writing styles and inclinations (Fig. 2a). In fact, synaptic weights of each selective DGC converge to stable values, giving rise to the formation of a receptive field that represents the weighted average of all input patterns for which that DGC is highly responsive (Methods). At the end of pretraining, the classification error for patterns not used during pretraining (Methods) is low: 0.75% (classification error on digit 3: 1.29%; digit 4: 0.20%), indicating that nearly all input patterns of the two digits are well represented by the network of mature DGCs. The median classification error for ten random combinations is 1.46%, the 25th-percentile 0.50%, and the 75th-percentile 2.74% (Suppl. Table S1).

Newborn neurons can learn novel patterns

After convergence of synaptic weights during pretraining, unresponsive model neurons have very weak synaptic connections because of LTD (Fig. 1c), die, and are replaced by newborn DGCs. To mimic exposure of an animal to a novel set of stimuli, we now add input patterns from digit 5 to the previous set of stimuli, which was limited to patterns of digits 3 and 4. Model newborn DGCs go through two maturation phases (Methods). The early phase of maturation is cooperative because, for each pattern presentation, activated mature DGCs indirectly excite the newborn DGCs via GABAergic interneurons. This lateral activation of newborn DGCs drives the growth of their receptive fields in a direction similar to those of the currently active mature DGCs. As a result, at the end of the early phase of maturation, newborn DGCs show a receptive field corresponding to a mixture of several patterns (Fig. 2b).

In the late phase of maturation, model newborn DGCs receive inhibitory GABAergic input from interneurons, similar to the input received by mature DGCs. The inhibitory input shifts the distribution of firing rates of newborn DGCs to lower firing rates in the late phase compared to the early phase of maturation (Fig. 1g). Given that at the end of the early phase, newborn DGCs have receptive fields similar to those of mature DGCs, lateral inhibition induces competition with mature DGCs for activation during presentation of patterns from the novel digit. Because model newborn DGCs start their late phase of maturation with a lower threshold (higher excitability) compared to mature DGCs, the activation of newborn DGCs is facilitated for those input patterns for which no mature DGC has preexisting selectivity. Therefore, in the late phase of maturation, competition drives the synaptic weights of newborn DGCs towards a receptive field corresponding to a subcategory of the ensemble of input patterns of the novel digit 5 (Fig. 2c).

To better characterize how DGCs represent various input patterns, we perform Principal Component Analysis (PCA) on the vector of final firing rates of all DGCs in response to MNIST test patterns representing digits 3, 4, and 5. We then project the firing rates of all DGCs on Principal Components (PCs) that are selected based on visual inspection (Fig. 3a2). A projection on the same PCs is also performed at the end of the early phase of maturation (Fig. 3a1). First, we observe that firing rate responses are distributed over a wide range, rather than concentrated, even for a single digit (Fig. 3a2), indicating that many neurons change their firing rate when the input switches from one example of

a "3" to another one with different writing style. Second, we can identify a few extremal examples of, say, fours that set loose boundaries of the cloud of all fours (Fig. 3c). Third, we observe that the representation of novel patterns lies close to the representation of pretrained patterns at the end of the early phase of maturation of newborn DGCs (Fig. 3a1), while during the late phase of maturation the representation expands into a previously empty subspace (Fig. 3a2), consistent with the experimentally observed promotion of pattern separation of newborn DGCs. Finally, by examining the trajectories of the firing rates of the DGCs for a few example patterns, we notice that some novel patterns are located far from each other at the end of the early phase (Fig. 3b1), but close to each other at the end of the late phase of maturation of newborn DGCs (Fig. 3b2), indicating a self-organization of the neuronal representation of novel patterns. The zoom in insets further shows curved, rather than straight, trajectories indicating the influence of lateral inhibition on firing rate dynamics of the DGCs.

To test the quality of the learned representations, we compute classification performance by a linear classifier for the three ensembles of digits (Methods). We obtain an overall classification error of 5.44% (classification error for digit 3: 9.50%; digit 4: 1.83%; digit 5: 4.82%). We compare this performance with that of a network where all three digit ensembles are simultaneously pretrained (Suppl. Fig. S1). In this case, the overall classification error is 7.91% (classification error for digit 3: 13.17%; digit 4: 1.22%; digit 5: 9.30%). Across ten simulation experiments classification performance is significantly better (lower error) when a novel ensemble of patterns is learned sequentially by newborn DGCs, than if all patterns are learned simultaneously (Wilcoxon signed rank test: p-val = 0.0020, Wilcoxon signed rank = 55; one-way t-test: p-val = 0.0269, t-stat = 2.6401, df = 9; Suppl. Table S1).

Furthermore, if two novel ensembles of digits (instead of a single one) are introduced during maturation of newborn DGCs, we observe that some newborn DGCs become selective for one of the novel digits, while others become selective for the other novel digit (Fig. 2d). Therefore, newborn DGCs can ultimately promote separation of several novel ensembles of patterns, no matter if they are learned simultaneously (Fig. 2d) or sequentially (Suppl. Fig. S1).

The switch from excitation to inhibition guides learning of novel representations

To assess whether maturation of newborn DGCs promotes learning of a novel ensemble of digit patterns, we compare our results with a control model without neurogenesis. Similar to the neurogenesis case, patterns from the novel digit 5 are introduced after pretraining with patterns from digits 3 and 4. In the control model, the thresholds and weights of all unresponsive neurons remain plastic after pretraining, similar to the neurogenesis case, while the feedforward weights and thresholds of DGCs that developed selectivity during pretraining are fixed. The only differences to the model with neurogenesis are that unresponsive neurons: (i) keep their feedforward weights (i.e., no reinitialization), and (ii) keep the same connections from and to inhibitory neurons.

We find that without neurogenesis, the previously unresponsive DGCs do not become selective for the novel digit 5, no matter during how many epochs patterns are presented (we went up to 100 epochs) (Fig. 2e). Therefore, if patterns from digit 5 are presented to the network, it fails to discriminate them from the previously learned digits 3 and 4: the overall classification error is 18.31% (classification error for digit 3: 14.06%; digit 4: 2.44%; digit 5: 40.58%). This result suggests that integration of newborn DGCs is beneficial for sequential learning of novel patterns.

As a further control, we compare with a model where all DGCs keep plastic feedforward weights. We observe that in the case where all neurons are plastic, learning of the novel digit occurs at the cost of loss of selectivity of mature neurons. Several DGCs switch their selectivity to become sensitive to the novel ensemble of patterns (Fig. 2f), while none of the previously unresponsive units becomes selective for the novel digit (compare with Fig. 2a). Compared to the model with neurogenesis, we observe a drop in classification performance to 9.08% error (classification error for digit 3: 14.55%; digit 4: 1.63%; digit 5: 11.10%). We find that the classification error for digit 3 is the one which increases the most. This is due to the fact that many DGCs previously selective for digit 3 updated their weights to become selective for digit 5.

Newborn dentate granule cells become selective for similar novel patterns

To investigate whether our theory for integration of newborn DGCs can explain why adult dentate gyrus neurogenesis promotes discrimination of similar stimuli, but does not affect discrimination of distinct patterns [18, 19], we use a simplified competitive winner-take-all network (Methods) and construct an artificial data set (Fig. 1e) that allows us to control the similarity s of pairs of clusters (Methods). The MNIST data set is not appropriate to tackle this question, because all digit clusters are similar and highly overlapping, reflected by a high within cluster dispersion (e.g. across the set of all "3") compared to the separation between clusters (e.g. typical "3" versus typical "5").

After pretraining the simplified network such that a first mature DGC responds to patterns of cluster 1 and a second mature DGC to those of cluster 2 (Fig. 4b1,b2), we introduce a newborn DGC in the network, and present patterns from three clusters (the two pretrained ones, as well as a novel one). We observe that the newborn DGC ultimately becomes selective for the novel cluster if it is similar ($s = 0.8$) to the two pretrained clusters (Fig. 4d1), but not if it is distinct ($s = 0.2$, Fig. 4d2). The selectivity develops in two phases. In the early phase of maturation of the newborn model cell, a pattern from the novel cluster that is similar to one of the pretrained clusters activates the mature DGC that has a receptive field closest to the novel pattern. This rapidly firing mature DGC activates the newborn DGC via lateral excitatory GABAergic connections, to a level where LTP is triggered at active synapses onto the newborn DGC. LTP also happens when a pattern from one of the pretrained clusters is presented. Thus, synaptic plasticity leads to a receptive field that reflects the average of all stimuli from all three clusters (Fig. 4c1).

To summarize our findings in a more mathematical language, we characterize the receptive field of the newborn cell by the vector of its feedforward weights. Analogous to the notion of a firing rate vector that represents the set of firing rates of an ensemble of neurons, the feedforward weight vector represents the set of weights of all synapses projecting onto a given neuron (Fig. 1a). In the early phase of maturation, for similar clusters, the feedforward weight vector onto the newborn DGC grows in the direction of the center of mass of all three clusters (the two pretrained ones and the novel one), because for each pattern presentation one of the mature DGCs becomes active (compare Fig. 4c1 and Fig. 4e1). However, if the novel cluster has a low similarity to pretrained clusters, patterns from the novel cluster do not activate any of the mature DGCs. Therefore the receptive field of the newborn cell reflects the average of stimuli from the two pretrained clusters only (compare Fig. 4c2 and Fig. 4e2).

As a result of the different orientation of the feedforward weight vector onto the newborn DGC at the end of the early phase of maturation, two different situations arise in the late phase of maturation, when lateral GABAergic connections are inhibitory. If the novel cluster is similar to the pretrained clusters,

the weight vector onto the newborn DGC at the end of the early phase of maturation lies at the center of mass of all the patterns across the three clusters. Thus it is closer to the novel cluster than the weight vector onto either of the mature DGCs (Fig. 4c1). So if a novel pattern is presented, the newborn DGC wins the competition between the three DGCs, and its feedforward weight vector moves towards the center of mass of the novel cluster (Fig. 4d1). By contrast, if the novel cluster is distinct, the weight vector onto the newborn DGC at the end of the early phase of maturation is located at the center of mass of the two pretrained clusters (Fig. 4c2). If a novel pattern is presented, no output unit is activated since their receptive fields are not similar enough to the input pattern. Therefore the newborn DGC always stays silent and never updates its feedforward weights (Fig. 4d2). These results are consistent with studies that have suggested that dentate gyrus is only involved in the discrimination of similar stimuli, but not distinct stimuli [26, 27]. For discrimination of distinct stimuli, another pathway might be used, such as the direct EC to CA3 connection [28].

In conclusion, our model suggests that adult dentate gyrus neurogenesis promotes discrimination of similar patterns because newborn DGCs can ultimately become selective for novel stimuli which are similar to already learned stimuli. On the other hand, newborn DGCs fail to represent novel distinct stimuli, precisely because they are too distinct from other stimuli already represented by the network. Presentation of novel distinct stimuli in the late phase of maturation therefore does not induce synaptic plasticity of the newborn DGC feedforward weight vector toward the novel stimuli. In the simplified network, the transition between similar and distinct can be determined analytically (Methods). This analysis clarifies the importance of the switch from cooperative dynamics (excitatory interactions) in the early phase to competitive dynamics (inhibitory interactions) in the late phase of maturation.

After integration, the receptive field of a newborn DGC represents the average of novel stimuli

To illustrate the analytical results and characterize the evolution of the receptive field of the newborn DGC, we examine the angle ϕ of the feedforward weight vector with the center of mass of the novel cluster (i.e. the average of the novel stimuli), as a function of maturation time (Fig. 5d,e and Suppl. Fig. S2).

In the early phase of maturation, the feedforward weight vector onto the newborn DGC grows, while its angle with the center of mass of the novel cluster stays constant (Suppl. Fig. S2). In the late phase of maturation, the angle ϕ between the center of mass of the novel cluster and the feedforward weight vector onto the newborn DGC decreases in the case of similar patterns (Fig. 5e), but not in the case of distinct patterns (Suppl. Fig. S2), indicating that the newborn DGC becomes selective for the novel cluster for similar but not for distinct patterns.

The analysis of the simplified model thus leads to a geometric picture that helps us to understand how the similarity of patterns influences the evolution of the receptive field of the newborn DGC before and after the switch from excitation to inhibition of the GABAergic input. For novel patterns that are similar to known patterns, the receptive field of a newborn DGC at the end of maturation represents the average of novel stimuli.

4 Discussion

While experimental results stemming from the manipulation of the ratio of NKCC1 to KCC2 suggest that the switch from excitation to inhibition of the GABAergic input onto adult-born DGCs is crucial for their integration into the preexisting circuit [8, 10, 14] and that adult dentate gyrus neurogenesis promotes better pattern separation [18–20], the link between channel properties and behavior has remained puzzling [21, 22]. Our modeling work has shown that the switch enables newborn DGCs to become selective for novel stimuli which are similar to familiar, already stored, representations, consistent with the experimentally-observed function of pattern separation [18–20]. Previous modeling studies already suggested that newborn DGCs integrate novel inputs into the representation in dentate gyrus [9, 29–36]. However, in these models newborn DGCs are born with large weights and weights are algorithmically renormalized in each time step, rather than grown from small values with a synaptic plasticity rule that intrinsically limits the weights. In particular, previous models neglect the two-phase integration issue. To our knowledge, we present the first synaptic plasticity model that can explain both: (i) how adult-born DGCs integrate into the preexisting network, and (ii) why they promote pattern separation of similar stimuli.

Our work emphasizes why a two-phase maturation of newborn DGCs is beneficial for proper integration in the preexisting network. From a computational perspective, the early phase of maturation, when GABAergic inputs onto newborn DGCs are excitatory, corresponds to cooperative unsupervised learning. Therefore, the synapses grow in the direction of patterns that indirectly activate the newborn DGCs via GABAergic interneurons (Fig. 5a). At the end of the early phase of maturation, the receptive field of a newborn DGC represents the center of mass of all input patterns that led to its (indirect) activation. In the late phase of maturation, GABAergic inputs onto newborn DGCs become inhibitory, so that lateral interactions change from cooperation to competition, causing a shift of the receptive fields of the newborn DGCs towards novel features (Fig. 5b). At the end of maturation, newborn DGCs are thus selective for novel inputs. This integration mechanism is in agreement with the experimental observation that newborn DGCs are broadly tuned early in maturation, yet highly selective at the end of maturation [37]. Loosely speaking, the cooperative phase of excitatory GABAergic input promotes the growth of the synaptic weights coarsely in the relevant direction, whereas the competitive phase of inhibitory GABAergic input helps to specialize on detailed, but potentially important differences between patterns.

In the context of theories of unsupervised learning, the switch of lateral GABAergic input to newborn DGCs from excitatory to inhibitory provides a biological solution to the "problem of unresponsive units" [38]. Unsupervised competitive learning has been used to perform clustering of input patterns into a few categories [38, 39]. Ideally, after learning of the feedforward weights between an input layer and a competitive network, input patterns that are distinct from each other activate different neuron assemblies of the competitive network. After convergence of competitive Hebbian learning, the vector of feedforward weights onto a given neuron points to the center of mass of the cluster of input patterns for which it is selective [38]. Yet, if the synaptic weights are randomly initialized, it is possible that the set of feedforward weights onto some neurons of the competitive network point in a direction "quasi-orthogonal" (Methods) to the subspace of the presented input patterns (Fig. 5c). Therefore those neurons, called "unresponsive units", will never get active during pattern presentation. Different learning strategies have been developed in the field of artificial neural networks to avoid this problem [38–41]. However, most of these algorithmic approaches lack a biological interpretation. In our model, the synapses onto newborn DGCs form spontaneously after neuronal birth. The excitatory GABAergic

input in the early phase of maturation drives the growth of the synaptic weights in the direction of the subspace of presented patterns that succeed in activating some of the mature DGCs. Hence the early cooperative phase of maturation can be seen as a smart initialization of the synaptic weights onto newborn DGCs, close enough to novel patterns so as to become selective for them in the late competitive phase of maturation. However, the cooperative phase is helpful only if the novel patterns are similar to the input statistics defined by the set of known patterns.

Our results are in line with the classic view that dentate gyrus is responsible for decorrelation of inputs [42, 43], a necessary step for storage of similar memories in CA3, and with the observation that dentate gyrus lesions impair discrimination of similar but not distinct stimuli [26, 27]. To discriminate distinct stimuli, another pathway might be involved, such as the direct EC to CA3 connection [28]. Our theory for integration of newborn DGCs explains why enhanced adult dentate gyrus neurogenesis promotes better discrimination of similar stimuli, but not distinct stimuli, as observed experimentally [18, 19]. In our model, the early cooperative phase of maturation can only drive the growth of synaptic weights onto newborn cells if mature DGCs are activated by presented stimuli. Hence the stimuli should be similar enough to familiar stimuli that are already represented by the network. If for a long time only distinct stimuli are presented, none of the mature DGCs becomes active, so the newborn DGCs are not indirectly activated and their synaptic weights do not grow. Consequently, in an experimental paradigm with distinct patterns only (i.e. prolonged absence from home cage and other familiar environments), our model predicts that newborn cells stay silent. As they are poorly integrated into the preexisting circuit, they will probably not survive [5].

Experimental observations support the importance of the switch from early excitation to late inhibition of the GABAergic input onto newborn DGCs. An absence of early excitation using NKCC1-knockout mice has been shown to strongly affect synapse formation and dendritic development in vivo [14]. Conversely, a reduction in inhibition in the dentate gyrus through decrease in KCC2 expression has been associated with epileptic activity [44]. An analogous switch of the GABAergic input has been observed during development, and its proper timing has been shown to be crucial for sensorimotor gating and cognition [12, 45]. In addition to early excitation and late inhibition, our theory also critically depends on the duration of the switch. Indeed, it supposes that a sufficient number of newborn DGCs are just about -within a few hours- to switch the effect of their GABAergic input when novel inputs are presented, to become selective for new features of the environment. Several experimental results have suggested that the switch is indeed sharp and occurs within a single day, both during development [46, 47] and adult dentate gyrus neurogenesis [11]. Furthermore, in hippocampal cell cultures, expression of KCC2 is upregulated by GABAergic activity but not affected by glutamatergic activity [48]. A similar process during adult dentate gyrus neurogenesis would increase the number of newborn DGCs available for representing novel features by advancing the timing of their switch. In this way, instead of a few thousands of newborn DGCs ready to switch (3 to 6% of the whole population [16, 17], divided by 30 days), a larger fraction of newborn DGCs would be made available for coding, if appropriate stimulation occurs.

To conclude, our theory for integration of adult-born DGCs suggests that newborn cells have a coding –rather than a modulatory– role during dentate gyrus pattern separation function. Our theory highlights the importance of GABAergic input in adult dentate gyrus neurogenesis, and links the switch from excitation to inhibition to the integration of newborn DGCs into the preexisting circuit. Finally, it illustrates how Hebbian plasticity of EC to DGC synapses makes newborn cells suitable to promote pattern separation of similar but not distinct stimuli, a long-standing question in the field of adult dentate gyrus neurogenesis [21, 22].

References

[1] Deng, W., Aimone, J. B. & Gage, F. H. New neurons and new memories: how does adult hippocampal neurogenesis affect learning and memory? *Nature reviews neuroscience* **11**, 339–350 (2010).

[2] Amaral, D. G., Scharfman, H. E. & Lavenex, P. The dentate gyrus: fundamental neuroanatomical organization (dentate gyrus for dummies). *Progress in brain research* **163**, 3–22 (2007).

[3] Johnston, S. T., Shtrahman, M., Parylak, S., Gonçalves, J. T. & Gage, F. H. Paradox of pattern separation and adult neurogenesis: A dual role for new neurons balancing memory resolution and robustness. *Neurobiology of learning and memory* **129**, 60–68 (2016).

[4] Dayer, A. G., Ford, A. A., Cleaver, K. M., Yassaee, M. & Cameron, H. A. Short-term and long-term survival of new neurons in the rat dentate gyrus. *Journal of Comparative Neurology* **460**, 563–572 (2003).

[5] Tashiro, A., Sandler, V. M., Toni, N., Zhao, C. & Gage, F. H. NMDA-receptor-mediated, cell-specific integration of new neurons in adult dentate gyrus. *Nature* **442**, 929–933 (2006).

[6] Schmidt-Hieber, C., Jonas, P. & Bischofberger, J. Enhanced synaptic plasticity in newly generated granule cells of the adult hippocampus. *Nature* **429**, 184–187 (2004).

[7] Li, L. *et al.* Silent synapses generate sparse and orthogonal action potential firing in adult-born hippocampal granule cells. *Elife* **6**, e23612 (2017).

[8] Ge, S., Yang, C.-h., Hsu, K.-s., Ming, G.-l. & Song, H. A critical period for enhanced synaptic plasticity in newly generated neurons of the adult brain. *Neuron* **54**, 559–566 (2007).

[9] Temprana, S. G. *et al.* Delayed coupling to feedback inhibition during a critical period for the integration of adult-born granule cells. *Neuron* **85**, 116–130 (2015).

[10] Alvarez, D. D. *et al.* A disynaptic feedback network activated by experience promotes the integration of new granule cells. *Science* **354**, 459–465 (2016).

[11] Heigele, S., Sultan, S., Toni, N. & Bischofberger, J. Bidirectional GABAergic control of action potential firing in newborn hippocampal granule cells. *Nature neuroscience* **19**, 263–270 (2016).

[12] Wang, D. D. & Kriegstein, A. R. Blocking early GABA depolarization with bumetanide results in permanent alterations in cortical circuits and sensorimotor gating deficits. *Cerebral cortex* **21**, 574–587 (2010).

[13] Ben-Ari, Y. Excitatory actions of GABA during development: the nature of the nurture. *Nature Reviews Neuroscience* **3**, 728–739 (2002).

[14] Ge, S. *et al.* GABA regulates synaptic integration of newly generated neurons in the adult brain. *Nature* **439**, 589–593 (2006).

[15] Tashiro, A., Makino, H. & Gage, F. H. Experience-specific functional modification of the dentate gyrus through adult neurogenesis: a critical period during an immature stage. *Journal of Neuroscience* **27**, 3252–3259 (2007).

[16] Van Praag, H., Kempermann, G. & Gage, F. H. Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus. *Nature neuroscience* **2**, 266–270 (1999).

[17] Cameron, H. A. & McKay, R. D. Adult neurogenesis produces a large pool of new granule cells in the dentate gyrus. *Journal of Comparative Neurology* **435**, 406–417 (2001).

[18] Clelland, C. *et al.* A functional role for adult hippocampal neurogenesis in spatial pattern separation. *Science* **325**, 210–213 (2009).

[19] Sahay, A. *et al.* Increasing adult hippocampal neurogenesis is sufficient to improve pattern separation. *Nature* **472**, 466–470 (2011).

[20] Jessberger, S. *et al.* Dentate gyrus-specific knockdown of adult neurogenesis impairs spatial and object recognition memory in adult rats. *Learning & memory* **16**, 147–154 (2009).

[21] Sahay, A., Wilson, D. A. & Hen, R. Pattern separation: a common function for new neurons in hippocampus and olfactory bulb. *Neuron* **70**, 582–588 (2011).

[22] Aimone, J. B., Deng, W. & Gage, F. H. Resolving new memories: a critical look at the dentate gyrus, adult neurogenesis, and pattern separation. *Neuron* **70**, 589–596 (2011).

[23] Kee, N., Teixeira, C. M., Wang, A. H. & Frankland, P. W. Preferential incorporation of adult-generated granule cells into spatial memory networks in the dentate gyrus. *Nature neuroscience* **10**, 355–362 (2007).

[24] Sjöström, P., Turrigiano, G. & Nelson, S. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* **32**, 1149–1164 (2001).

[25] Artola, A., Bröcher, S. & Singer, W. Different voltage dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature* **347**, 69–72 (1990).

[26] Gilbert, P. E., Kesner, R. P. & Lee, I. Dissociating hippocampal subregions: A double dissociation between dentate gyrus and ca1. *Hippocampus* **11**, 626–636 (2001).

[27] Hunsaker, M. R. & Kesner, R. P. Evaluating the differential roles of the dorsal dentate gyrus, dorsal ca3, and dorsal ca1 during a temporal ordering for spatial locations task. *Hippocampus* **18**, 955–964 (2008).

[28] Yeckel, M. F. & Berger, T. W. Feedforward excitation of the hippocampus by afferents from the entorhinal cortex: redefinition of the role of the trisynaptic pathway. *Proceedings of the National Academy of Sciences* **87**, 5832–5836 (1990).

[29] Chambers, R. A., Potenza, M. N., Hoffman, R. E. & Miranker, W. Simulated apoptosis/neurogenesis regulates learning and memory capabilities of adaptive neural networks. *Neuropsychopharmacology* **29**, 747–758 (2004).

[30] Becker, S. A computational principle for hippocampal learning and neurogenesis. *Hippocampus* **15**, 722–738 (2005).

[31] Crick, C. & Miranker, W. Apoptosis, neurogenesis, and information content in hebbian networks. *Biological cybernetics* **94**, 9–19 (2006).

[32] Wiskott, L., Rasch, M. J. & Kempermann, G. A functional hypothesis for adult hippocampal neurogenesis: avoidance of catastrophic interference in the dentate gyrus. *Hippocampus* **16**, 329–343 (2006).

[33] Chambers, R. A. & Conroy, S. K. Network modeling of adult neurogenesis: shifting rates of neuronal turnover optimally gears network learning according to novelty gradient. *Journal of cognitive neuroscience* **19**, 1–12 (2007).

[34] Appleby, P. A. & Wiskott, L. Additive neurogenesis as a strategy for avoiding interference in a sparsely-coding dentate gyrus. *Network: Computation in Neural Systems* **20**, 137–161 (2009).

[35] Aimone, J. B., Wiles, J. & Gage, F. H. Computational influence of adult neurogenesis on memory encoding. *Neuron* **61**, 187–202 (2009).

[36] Weisz, V. I. & Argibay, P. F. A putative role for neurogenesis in neurocomputational terms: Inferences from a hippocampal model. *Cognition* **112**, 229–240 (2009).

[37] Marín-Burgin, A., Mongiat, L. A., Pardi, M. B. & Schinder, A. F. Unique processing during a period of high excitation/inhibition balance in adult-born neurons. *Science* **335**, 1238–1242 (2012).

[38] Hertz, J., Krogh, A. & Palmer, R. G. *Introduction to the Theory of Neural Computation* (Addison-Wesley, 1991).

[39] Du, K.-L. Clustering: A neural network approach. *Neural networks* **23**, 89–107 (2010).

[40] Grossberg, S. Adaptive pattern classification and universal recoding ii: Feedback, expectation, olfaction, illusions. *Biol. Cybern.* **23**, 187–202 (1976).

[41] Bienenstock, E. L., Cooper, L. N. & Munro, P. W. Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience* **2**, 32–48 (1982).

[42] Marr, D. Simple memory: a theory for archicortex. *Philosophical Transactions of the Royal Society of London* **262**, 23–81 (1971).

[43] Rolls, E. T. & Treves, A. *Neural networks and brain function*, vol. 572 (Oxford university press Oxford, 1998).

[44] Barmashenko, G., Hefft, S., Aertsen, A., Kirschstein, T. & Köhling, R. Positive shifts of the $GABA_A$ receptor reversal potential due to altered chloride homeostasis is widespread after status epilepticus. *Epilepsia* **52**, 1570–1578 (2011).

[45] Furukawa, M. *et al.* Neonatal maternal separation delays the GABA excitatory-to-inhibitory functional switch by inhibiting kcc2 expression. *Biochemical and biophysical research communications* **493**, 1243–1249 (2017).

[46] Tyzio, R., Holmes, G. L., Ben-Ari, Y. & Khazipov, R. Timing of the developmental switch in $GABA_A$ mediated signaling from excitation to inhibition in CA3 rat hippocampus using gramicidin perforated patch and extracellular recordings. *Epilepsia* **48**, 96–105 (2007).

[47] Leonzino, M. *et al.* The timing of the excitatory-to-inhibitory GABA switch is regulated by the oxytocin receptor via KCC2. *Cell reports* **15**, 96–103 (2016).

[48] Ganguly, K., Schinder, A. F., Wong, S. T. & Poo, M.-m. GABA itself promotes the developmental switch of neuronal GABAergic responses from excitation to inhibition. *Cell* **105**, 521–532 (2001).

5 Methods

Network architecture and neuronal dynamics

DGCs are the principal cells of the dentate gyrus. They mainly receive excitatory projections from the entorhinal cortex through the perforant path and GABAergic inputs from local interneurons, as well as excitatory input from Mossy cells. They project to CA3 pyramidal cells and inhibitory neurons, as well as local Mossy cells [2, 9].

In our model, we omit Mossy cells and describe the dentate gyrus as a competitive circuit consisting of N_{DGC} dentate granule cells and N_I GABAergic interneurons (Fig. 1a). The activity of N_{EC} neurons in EC represents an input pattern $\vec{x} = (x_1, x_2, ..., x_{N_{EC}})$. Because the perforant path also induces strong feedforward inhibition in the dentate gyrus [49], we assume that the effective EC activity is normalized, such that $||\vec{x}|| = 1$ for any input pattern $\vec{x}$. We use P different input patterns $\vec{x}^\mu$, $1 \leqslant \mu \leqslant P$ in the simulations of the model.

The EC neurons have excitatory all-to-all connections to the DGCs. In rodent hippocampus, spiking mature DGCs activate interneurons in DG, which in turn inhibit other mature DGCs [9, 10]. In our model, the DGCs are thus recurrently connected with inhibitory neurons (Fig. 1a). Connections from DGCs to interneurons exist in our model with probability p_{IE} and have a weight w_{IE}. Similarly, connections from interneurons to DGCs occur with probability p_{EI} and have a weight w_{EI}. All parameters are reported in Table 1(Biologically-plausible network).

Before an input pattern is presented, all rates of model DGCs are initialized to zero. Upon stimulation with input pattern $\vec{x}$, the firing rate ν_i of DGC i evolves according to [50]:

$$\tau_m \frac{\mathrm{d}\nu_i}{\mathrm{d}t} = -\nu_i + \tanh\left(\frac{[I_i - b_i]_+}{L}\right) \tag{1}$$

where $[.]_+$ denotes rectification: $[a] = a$ for $a > 0$ and zero otherwise. Here, b_i is a firing threshold, $L = 0.5$ is the smoothness parameter of the frequency-current curve (L^{-1} is the slope of the frequency-current curve at the firing threshold), and I_i the total input to cell i:

$$I_i = \sum_{j=1}^{N_{EC}} w_{ij} x_j + \sum_{k=1}^{N_I} w_{ik}^{EI} \nu_k^I \tag{2}$$

with x_j the activity of EC input neuron j, $w_{ij} \geqslant 0$ the feedforward weight from EC input neuron j to DGC i, and w_{ik}^{EI} the weight from inhibitory neuron k to DGC i. The sum runs over all inhibitory neurons, but the weights are set to $w_{ik}^{EI} = 0$ if the connection is absent. The firing rate ν_i is unit-free and normalized to a maximum of 1, which we interpret as a firing rate of 10 Hz. We take the synaptic weights as unit-less parameters such that I_i is also unit-free.

The firing rate ν_k^I of inhibitory neuron k, is defined as:

$$\tau_{\mathrm{inh}} \frac{\mathrm{d}\nu_k^I}{\mathrm{d}t} = -\nu_k^I + [I_k^I - p^* N_{DGC}]_+ \tag{3}$$

with p^* a parameter which relates to the desired ensemble sparsity, and I_k^I the total input towards interneuron k, given as:

$$I_k^I = \sum_{i=1}^{N_{DGC}} w_{ki}^{IE} \nu_i \tag{4}$$

with w_{ki}^{IE} the weight from DGC i to inhibitory neuron k. (We set $w_{ki}^{IE} = 0$ if the connection is absent.) The feedback from inhibitory neurons ensures a sparse activity of model DGCs for each pattern. With $p^* = 0.1$ we find that more than 70 % of model DGCs are silent (firing rate < 1 Hz [51]) when an input pattern is presented, and less than 10% are highly active (firing rate > 9 Hz) (Fig. 1g), consistent with the experimentally observed sparse activity in dentate gyrus [52].

Plasticity rule

Projections from EC onto newborn DGCs exhibit Hebbian plasticity [6, 8] [53]. Therefore, in our model the connections from EC neurons to DGCs are plastic, following a Hebbian learning rule which exhibits long-term depression (LTD) or long-term potentiation (LTP) depending on the firing rate ν_i of the postsynaptic

cell [24, 25, 41] [54]. Input patterns $\vec{x}^\mu$, $1 \leqslant \mu \leqslant P$, are presented in random order. For each input pattern, we let the firing rates converge for a time T where T was chosen long enough to achieve convergence to a precision of 10^{-6}. After $n-1$ presentations (i.e. at time $(n-1)\cdot T$) the weight vector has value $w_{ij}^{(n-1)}$. We then present the next pattern and update at time $n \cdot T$ ($w_{ij}^{(n)} = w_{ij}^{(n-1)} + \eta \Delta w_{ij}$), according to the following plasticity rule:

$$\Delta w_{ij} = -\alpha x_j \nu_i [\theta - \nu_i]_+ + \gamma x_j \nu_i [\nu_i - \theta]_+ - \beta w_{ij} [\nu_i - \theta]_+ \nu_i^3 \tag{5}$$

where x_j is the firing rate of presynaptic EC input neuron j, ν_i the firing rate of postsynaptic DGC i, θ marks the transition from LTD to LTP, and the relative strength α, γ of LTP and LTD depend on θ via $\alpha = \frac{\alpha_0}{\theta^3} > 0$ and $\gamma = \gamma_0 - \theta > 0$. The values of the parameters α_0, γ_0, β, and θ are given in Table 1(Biologically-plausible network). The weights are hard-bounded from below at 0, i.e. if equation (5) leads to a new weight smaller than zero, w_{ij} is set to zero. The first two terms of expression (5) are a variation of the BCM rule [41]. The third term implements heterosynaptic plasticity [55, 56]. Because the first two terms of the plasticity rule are Hebbian and proportional to the presynaptic activity x_j, the active DGCs ($\nu_i > \theta$) update their feedforward weights in direction of the input pattern $\vec{x}$. Moreover, all weights onto neuron i are downregulated heterosynaptically by an amount that increases supra-linearly with the postsynaptic rate ν_i. Similar to learning in a competitive network [38], the vector of feedforward weights onto active DGCs will move towards the center of mass of the cluster of patterns they are selective for, as we will discuss now.

For a given input pattern $\vec{x}^\mu$, there are three fixed points for the postsynaptic firing rate: $\nu_i = 0$, $\nu_i = \theta$, and $\nu_i = \hat{\nu}_i$ (the negative root is omitted, because $\nu_i \geqslant 0$ due to equation (1)). For $\nu_i < \theta$, there is LTD, so the weights move toward zero: $w_{ij} \to 0$, while for $\nu_i > \theta$, there is LTP, so the weights move toward $w_{ij} \to \frac{\gamma x_j^\mu}{\beta \hat{\nu}_i^2}$ (Fig. 1c). If a pattern $\vec{x}^\mu$ is presented only for a short time these fixed points are not reached during a single pattern presentation.

Winners, losers, and quasi-orthogonal inputs

We define the winners as the DGCs which become strongly active ($\nu_i > \theta$) during presentation of an input pattern. Since the input patterns are normalized to have an L2-norm of 1 ($||\vec{x}^\mu|| = 1$ by construction), and the L2-norm of the feedforward weight vectors is bounded (see Section Direction and length of the weight vector), the winning units are the ones whose weight vectors $\vec{w}_i$ (row of the feedforward connectivity matrix) align best with the current input pattern $\vec{x}^\mu$. Furthermore, we say that an input pattern $\vec{x}^\mu$ is "quasi-orthogonal" to a weight vector $\vec{w}_i$ if $I_i = \sum_{j=1}^{N_{EC}} w_{ij} x_j + \sum_{k=1}^{N_I} w_{ik}^{EI} \nu_k^I < b_i$. If an input pattern $\vec{x}^\mu$ is quasi-orthogonal to a weight vector $\vec{w}_i$, then neuron i does not fire in response to $\vec{x}^\mu$. Note that for a case without inhibitory neurons and with $b_i \to 0$, we recover the standard orthogonality condition.

Direction and length of the weight vector

Let us denote the ensemble of patterns for which neuron i is a winner by C_i and call this the set of winning patterns ($C_i = \{\mu | \nu_i > \theta\}$). Suppose that neuron i is quasi-orthogonal to all other patterns, so that for all $\mu \notin C_i$ we have $\nu_i = 0$. Then the feedforward weight vector of neuron i converges in expectation to:

$$\vec{w}_i = \frac{\gamma}{\beta} \frac{\langle G_1(\nu_i)\vec{x} \rangle_{\mu \in C_i}}{\langle G_2(\nu_i) \rangle_{\mu \in C_i}} \tag{6}$$

where $G_1(\nu_i) = (\nu_i - \theta)\nu_i$ and $G_2(\nu_i) = (\nu_i - \theta)\nu_i^3$. Hence $\vec{w}_i$ is a weighted average over all winning patterns.

The squared length of the feedforward weight vector can be computed by multiplying equation (6) with $\vec{w}_i$:

$$||\vec{w}_i||^2 = \vec{w}_i \cdot \vec{w}_i = \frac{\gamma}{\beta} \frac{\langle G_1(\nu_i)\,(\vec{w}_i \cdot \vec{x}) \rangle_{\mu \in C_i}}{\langle G_2(\nu_i) \rangle_{\mu \in C_i}} \tag{7}$$

Since input patterns have length one, the scalar product on the right-hand side can be rewritten as $\vec{w}_i \cdot \vec{x} = ||\vec{w}_i|| \cos(\alpha)$ where α is the angle between the weight vector and pattern $\vec{x}$. Division by $||\vec{w}_i||$ yields the L2-norm of the feedforward weight vector:

$$||\vec{w}_i|| = \frac{\gamma}{\beta} \frac{\langle G_1(\nu_i) \cos(\alpha) \rangle_{\mu \in C_i}}{\langle G_2(\nu_i) \rangle_{\mu \in C_i}} \tag{8}$$

where the averages run, as before, over all winning patterns.

Let us now derive bounds for $||\vec{w}_i||$. First, since $\cos(\alpha) \leqslant 1$ we have $\langle G_1(\nu_i) \cos(\alpha) \rangle_{\mu \in C_i} \leqslant \langle G_1(\nu_i) \rangle_{\mu \in C_i}$. Second, since for all winning patterns $\nu_i > \theta$, where θ is the LTP threshold, we have $\langle G_2(\nu_i) \rangle_{\mu \in C_i} \geqslant \langle (\nu_i - \theta)\, \nu_i \rangle \theta^2$. Thus the length of the weight vector is finite and bounded by:

$$||\vec{w}_i|| \leqslant \frac{\gamma}{\beta} \frac{\langle G_1(\nu_i) \rangle_{\mu \in C_i}}{\langle G_2(\nu_i) \rangle_{\mu \in C_i}} \leqslant \frac{\gamma}{\beta} \frac{1}{\theta^2} \tag{9}$$

It is possible to make the second bound tighter if we find the winning pattern with the smallest firing rate $\nu_{\min}$ such that $\nu_i \geqslant \nu_{\min}\ \forall i \in C_i$:

$$||\vec{w}_i|| \leqslant \frac{\gamma}{\beta} \frac{1}{(\nu_{\min})^2} \tag{10}$$

The bound is reached if neuron i is winner for a single input pattern.

We can also derive a lower bound. For a pattern $\mu \in C_i$, let us write the firing rate of neuron i as $\nu_i(\mu) = \bar{\nu}_i + \Delta\nu_i(\mu)$ where $\bar{\nu}_i$ is the mean firing rate of neuron i averaged across all winning patterns and $\langle \Delta\nu_i \rangle_{\mu \in C_i} = 0$. We assume that the absolute size of $\Delta\nu_i$ is small, i.e., $\langle (\Delta\nu_i)^2 \rangle_{\mu \in C_i} \ll (\bar{\nu}_i)^2$. Linearization of equation (8) around $\bar{\nu}_i$ yields:

$$||\vec{w}_i|| = \frac{\gamma}{\beta} \frac{G_1(\bar{\nu}_i)}{G_2(\bar{\nu}_i)} \langle \cos(\alpha) \rangle_{\mu \in C_i} + \frac{\gamma}{\beta} \frac{G_1'(\bar{\nu}_i)}{G_2(\bar{\nu}_i)} \langle \cos(\alpha) \Delta\nu_i \rangle_{\mu \in C_i} \tag{11}$$

Elementary geometric arguments for a neuron model with monotonically increasing frequency-current curve yield that the value of $\langle \cos(\alpha) \Delta\nu_i \rangle_{\mu \in C_i}$ is positive (or zero), because an increase in the angle α lowers both the cosine and the firing rate, giving rise to a positive correlation. Since we are interested in a lower bound, we can therefore drop the term proportional to G_1' and evaluate the ratio G_1/G_2 to find:

$$||\vec{w}_i|| \geqslant \frac{\gamma}{\beta} \frac{1}{(\bar{\nu}_i)^2} \langle \cos(\alpha) \rangle_{\mu \in C_i} \geqslant \frac{\gamma}{\beta} \frac{1}{(\nu_{\max})^2} \cos(\hat{\alpha}) \tag{12}$$

where $\nu_{\max}$ is the maximal firing rate of a DGC and $\hat{\alpha} = \max_{\mu \in \mathrm{C_i}} \{\alpha\}$ is the angle of the winning pattern that has the largest angle with the weight vector. The first bound is tight and is reached if neuron i is winner for only two patterns.

To summarize we find that the length of the weight vector remains bounded in a narrow range. Hence, for a reasonable distribution of input patterns and weight vectors, the value of $||\vec{w}_i||$ is similar for different neurons i, so that the weight vector will have, after convergence, similar lengths for all DGCs that are winners for at least one pattern. In our simulations with the MNIST data set, we find that the length of feedforward weight vectors lies in the range between 9.3 and 11.1 across all responsive neurons.

Early maturation phase

During the early phase of maturation, the GABAergic input onto a newborn DGC with index l has an excitatory effect. In the model, it is implemented as follows: $w_{lk}^{EI} = -w_{EI} > 0$ with probability p_{EI} for any interneuron k and $w_{lk}^{EI} = 0$ otherwise (no connection). Since newborn cells do not project yet onto inhibitory neurons [9],

we have $w_{kl}^{IE} = 0\ \forall l$. Newborn DGCs are known to have enhanced excitability [6, 7], so their threshold is kept at $b_l = 0\ \forall l$. Presentation of all patterns of the data set once (1 epoch) is sufficient to reach convergence of the feedforward weights onto newborn DGCs.

Because the newborn DGCs receive lateral excitation via interneurons and their thresholds are zero during the early phase of maturation, the lateral excitatory GABAergic input is always sufficient to activate them. Hence, if the firing rate of a newborn DGC exceeds the LTP threshold θ, the feedforward weights grow towards the presented input pattern, cf. equation (5).

Late maturation phase

During the late phase of maturation (starting at about 3 weeks [14]), the GABAergic input onto newborn DGCs switches from excitatory to inhibitory. In terms of our model, it means that all existing w_{lk}^{EI} connections switch their sign to $w_{EI} < 0$. Furthermore, since newborn DGCs develop lateral connections to inhibitory neurons in the late maturation phase [9], we set $w_{kl}^{IE} = w_{IE}$ with probability p_{IE}, and $w_{kl}^{IE} = 0$ otherwise. The thresholds of newborn DGCs are updated after presentation of pattern μ at time $n \cdot T$ ($b_l^{(n)} = b_l^{(n-1)} + \eta_b \Delta b_l$) according to $\Delta b_l = \nu_l - \nu_0$, where ν_0 is a reference rate, to mimic the decrease of excitability as newborn DGCs mature. Therefore the distribution of firing rates of newborn DGCs is shifted to the left (towards lower firing rates) at the end of the late phase of maturation compared to the early phase of maturation (Fig. 1g). A sufficient condition for a newborn DGC to win the competition upon presentation of patterns of the novel cluster is that the scalar product between a pattern of the novel cluster and the feedforward weight vector onto the newborn DGC is larger than the scalar product between the pattern of the novel cluster and the feedforward weight vector onto any of the mature DGCs. Analogous to the early phase of maturation, presentation of all patterns of the data set once (1 epoch) is sufficient to reach convergence of the feedforward weights onto newborn DGCs.

Input patterns

Two different sets of input patterns are used. Both data sets have a number K of clusters and several thousands of patterns per cluster. As a first data set, we use the MNIST 12x12 patterns [57] ($N_{EC} = 144$), normalized such that the L2-norm of each pattern is equal to 1. The training set contains approximately 6000 patterns per digit, while the testing set contains about 1000 patterns per digit (Fig. 1d).

As a second data set, we use hand-made artificial patterns designed such that the distance between the centers of any two clusters, or in other words their pairwise similarity, is the same. All clusters lie on the positive quadrant of the surface of a hypersphere of dimension $N_{EC} - 1$. The cluster centers are Walsh patterns shifted along the diagonal (Fig. 1e):

$$\begin{aligned}
\vec{P}^1 &= \frac{1}{c_0}\,(1+\xi, 1-\xi, 1+\xi, 1-\xi, ..., 1+\xi, 1-\xi, 1+\xi, 1-\xi) \\
\vec{P}^2 &= \frac{1}{c_0}\,(1+\xi, 1+\xi, 1-\xi, 1-\xi, ..., 1+\xi, 1+\xi, 1-\xi, 1-\xi) \\
&... \\
\vec{P}^K &= \frac{1}{c_0}\,(1+\xi, 1+\xi, 1+\xi, 1+\xi, ..., 1-\xi, 1-\xi, 1-\xi, 1-\xi)
\end{aligned} \tag{13}$$

with $|\xi| < 1$ a parameter that determines the spacing between clusters. c_0 is a normalization factor to ensure that the center of mass of all clusters has an L2-norm of 1:

$$c_0 = \sqrt{N_{EC}\,(1+\xi^2)}. \tag{14}$$

The number of input neurons N_{EC} is $N_{EC} = 2^K$. The scalar product, and hence the angle Ω, between the center of mass of any pair of clusters k and l ($k \neq l$) is a function of ξ (Fig. 1f):

$$\vec{P}^k \cdot \vec{P}^l = \frac{1}{1+\xi^2} = \cos(\Omega) \tag{15}$$

We define the pairwise similarity s of two clusters as: $s = 1 - \xi$. Highly similar clusters have a large s due to the small distance between their centers (hence a small ξ).

To make the artificial data set comparable to the MNIST 12x12 data set, we choose $K = 7$, so $N_{EC} = 128$, and we generate 6000 noisy patterns per cluster for the training set and 1000 other noisy patterns per cluster for the testing set. Since our noisy high-dimensional input patterns have to be symmetrically distributed around the centers of mass $\vec{P}^k$, yet lie on the hypersphere, we have to use an appropriate sampling method. The patterns $\vec{x}^{\mu(k)}$ of a given cluster k with center of mass $\vec{P}^k$ are thus sampled from a Von Mises-Fisher distribution [58]:

$$\vec{x}^{\mu(k)} \sim \left(\sqrt{1-a^2}\right)\vec{\zeta} + a\vec{P}^k \tag{16}$$

with $\vec{\zeta}$ an L2-normalized vector taken in the space orthogonal to $\vec{P}^k$. The vector $\vec{\zeta}$ is obtained by performing the singular-value decomposition of $\vec{P}^k$ ($U\Sigma V^* = \vec{P}^k$), and multiplying the matrix U (after removing its first column), which corresponds to the left-singular vectors in the orthogonal space to $\vec{P}^k$, with a vector whose elements are drawn from the standard normal distribution. Then the L2-norm of the obtained pattern is set to 1, so that it lies on the surface of the hypersphere. A rejection sampling scheme is used to obtain a [58]. The sample a is kept if $\kappa a + (N_{EC} - 1)\ln(1 - \psi a) - c \geqslant \ln(u)$, with κ a concentration parameter, $\psi = \frac{1-b}{1+b}$, $c = \kappa\psi + (N_{EC} - 1)\ln(1 - \psi^2)$, u drawn from a uniform distribution $u \sim U[0, 1]$, $a = \frac{1-(1+b)z}{1-(1-b)z}$, $b = \frac{N_{EC}-1}{\sqrt{4\kappa^2+(N_{EC}-1)^2}+2\kappa}$, and z drawn from a beta distribution $z \sim \mathcal{B}e(\frac{N_{EC}-1}{2}, \frac{N_{EC}-1}{2})$.

The concentration parameter κ characterizes the spread of the distribution around the center $\vec{P}^k$. In the limit where $\kappa \to 0$, sampling from the Von Mises-Fisher distribution becomes equivalent to sampling uniformly on the surface of the hypersphere, so the clusters become highly overlapping. In dimension $N_{EC} = 128$, if $\kappa > 10^3$ the probability of overlap between clusters is negligible. We use a value $\kappa = 10^4$.

Classification performance

To evaluate whether the newborn DGCs contribute to the function of the dentate gyrus network, we study classification performance. Once the feedforward weights have been adjusted upon presentation of many input patterns from the training set (Section Plasticity rule), we keep them fixed and determine classification on the test set using artificial readout units (RO).

To do so, the readout weights (w_{ki}^{RO} from model DGC i to readout unit k) are initialized at random values drawn from a uniform distribution: $w_{ki}^{RO} \sim \sigma\mathcal{U}(0, 1)$, with $\sigma = 0.1$. The number of readout units, N_{RO}, corresponds to the number of learned classes. To adjust the readout weights, all patterns of the training data set that belong to the learned classes are presented one after the other. For each pattern $\vec{x}^\mu$, we let the firing rate of the DGCs converge (values at convergence: ν_i^μ). The activity of a readout unit k is given by:

$$\nu_k^{RO,\mu} = g\left(I_k^{RO,\mu}\right) = g\left(\sum_{i=1}^{N_{DGC}} w_{ki}^{RO}\nu_i^\mu\right) \tag{17}$$

As we aim to assess the performance of the network of DGCs, the readout weights are adjusted by an artificial supervised learning rule. The loss function, which corresponds to the difference between the activity of the

readout units and a one-hot representation of the corresponding pattern label [38],

$$L(W^{RO}) = \frac{1}{2}\sum_{k=1}^{N_{RO}}(L_k^\mu - \nu_k^{RO,\mu})^2 \tag{18}$$

with L_k^μ the element k of a one-hot representation of the correct label of pattern $\vec{x}^\mu$, is minimized by stochastic gradient descent:

$$\Delta w_{ki}^{RO,\mu} = \eta(L_k^\mu - \nu_k^{RO,\mu})g'\left(I_k^{RO,\mu}\right)\nu_i^\mu. \tag{19}$$

The readout units have a rectified hyperbolic tangent frequency-current curve: $g(x) = \tanh(2[x]_+)$, whose derivative is: $g'(x) = 2\left(1 - (\tanh(2[x]_+))^2\right)$. We learn the weights of the readout units over 100 epochs of presentations of all training patterns with $\eta = 0.01$, which is sufficient to reach convergence.

Thereafter, the readout weights are fixed. Each test set pattern belonging to one of the learned classes is presented once, and the firing rates of the DGCs are let to converge. Finally, the activity of the readout units $\nu_k^{RO,\mu}$ is computed and compared to the correct label L_k^μ of the presented pattern. If the readout unit with the highest activity value is the one that represents the class of the presented input pattern, the pattern is said to be correctly classified. Classification error is given by the number of misclassified patterns divided by the total number of test patterns of the learned classes.

Control cases

In our standard setting, patterns from a third digit are presented to a network that has previously only seen patterns from two digits. The question is whether neurogenesis helps when adding the third digit. We use several control cases to compare with the neurogenesis case. In two control cases, we either keep all feedforward connections towards the DGCs plastic (Fig. 2f), or fix the feedforward connections for all selective DGCs but keep unselective neurons plastic (as in the neurogenesis case) (Fig. 2e). However, in both instances, the DGCs do not mature in the two-step process of our model of neurogenesis. Finally, in the third control case, all three digits are learned in parallel (Suppl. Fig. S1).

Pretraining with two digits

As we are interested by neurogenesis at the adult stage, we pretrain the network with patterns from two digits, such that it already stores some memories before neurogenesis takes place. To do so, we randomly initialize the weights from EC neurons to DGCs: they are drawn from a uniform distribution ($w_{ij} \sim U[0,1]$). The L2-norm of the feedforward weight vector onto each DGC is then normalized to 1, to ensure fair competition between DGCs during learning. Then we present all patterns from digits 3 and 4 in random order, as many times as needed for convergence of the weights. During each pattern presentation the firing rates of the DGCs are computed (Section Network architecture and neuronal dynamics) and their feedforward weights are updated according to our plasticity rule (Section Plasticity rule). We find that we need approximately 40 epochs for convergence of the weights, and use 80 epochs to make sure that all weights are stable. At the end of pretraining, our network is considered to correspond to an adult stage, because some DGCs are selective for prototypes of the pretrained digits (Fig. 2a).

Statistics

In the main text, we present a representative example with three digits from the MNIST data set (3, 4 and 5). It is selected from a set of ten random combinations of three different digits. For each combination, one network is pretrained with two digits for 80 epochs. Then the third digit is added and neurogenesis takes place (one epoch of early phase of maturation, and one epoch of late phase of maturation). Furthermore another network is pretrained directly with the three digits for 80 epochs. Classification performance is reported for all combinations (Suppl. Table S1).

Simplified rate network

We use a toy network and the artificial data set to determine if our theory of integration of newborn DGCs can explain why adult dentate gyrus neurogenesis helps for the discrimination of similar, but not for distinct patterns.

The rate network described above is simplified as follows. We use K dentate granule cells for K clusters. Their firing rate ν_i is given by:

$$\tau_m \frac{\mathrm{d}\nu_i}{\mathrm{d}t} = -\nu_i + \mathcal{H}\left(I_i - b_i\right) \tag{20}$$

where $\mathcal{H}$ is the Heaviside step function. As before, b_i is the threshold, and I_i the total input towards neuron i:

$$I_i = \sum_{j=1}^{N_{EC}} w_{ij} x_j + \sum_{k \neq j}^{N_{DGC}} w_{rec} \nu_k \tag{21}$$

with x_j the input of presynaptic EC neuron j, w_{ij} the feedforward weight between EC neuron j and DGC i, and ν_k the firing rate of DGC k. Inhibitory neurons are modeled implicitly: each DGC directly connects to all other DGCs via inhibitory recurrent connections of value $w_{rec} < 0$. During presentation of pattern $\vec{x}^{\mu}$, the firing rates of the DGCs evolve according to equation (20). After convergence, the feedforward weights are updated: $w_{ij}^{(\mu)} = w_{ij}^{(\mu-1)} + \eta \Delta w_{ij}$. The synaptic plasticity rule is the same as before, see equation (5), but with the parameters reported in Table 1(Simple network). They are different from those of the biologically-plausible network because we now aim for a single winning neuron for each cluster. Note that for an LTP threshold $\theta < 1$ all active DGCs update their feedforward weights, because of the Heaviside function for the firing rate (equation (20)).

Assuming a single winner i^* for each pattern presentation, the input (equation (21)) to the winner is:

$$I_{i^*} = \vec{w}_{i^*} \cdot \vec{x}, \tag{22}$$

while the input to the losers is:

$$I_i = \vec{w}_i \cdot \vec{x} + w_{\mathrm{rec}}. \tag{23}$$

Therefore, two conditions need to be satisfied for a solution with a single winner:

$$\vec{w}_{i^*} \cdot \vec{x} > b_i \tag{24}$$

for the winner to actually be active, and:

$$\vec{w}_i \cdot \vec{x} + w_{\mathrm{rec}} < b_i \tag{25}$$

to prevent non-winners to become active. The value of b_i is lower in the early phase than in the late phase of maturation to mimic enhanced excitability [6, 7].

Similar versus distinct patterns with the artificial data set

Using the artificial data set with $|\xi| < 1$ (equation (13)), the scalar product between the centers of mass of two different clusters, given by equation (15), satisfies: $0.5 \leqslant \frac{1}{1+\xi^2} \leqslant 1$. This corresponds to $0° \leqslant \Omega \leqslant \Omega_{\max} = 60°$.

After stimulation with a pattern $\vec{x}$, it takes some time before the firing rates of the DGCs converge. We call two patterns "similar" if they activate, at least initially, the same output unit, while we consider two patterns as "distinct" if they do not activate the same output unit, not even initially. We now show that, with a large concentration parameter κ, patterns of different clusters are similar if $\xi < \sqrt{\frac{||\vec{w}_i||}{b_i} - 1}$ and distinct if $\xi > \sqrt{\frac{||\vec{w}_i||}{b_i} - 1}$.

We first consider a DGC i whose feedforward weight vector has converged towards the center of mass of cluster k. If an input pattern $\vec{x}^{\mu(k)}$ from cluster k is presented, it will receive the following initial input:

$$I_i = \vec{w}_i \cdot \vec{x}^{\mu(k)} = ||\vec{w}_i|| \cdot ||\vec{x}^{\mu(k)}|| \cdot \cos(\vartheta_{\mathrm{kk}}) = ||\vec{w}_i|| \cdot \cos(\vartheta_{\mathrm{kk}}) \tag{26}$$

where ϑ_{kk} is the angle between the pattern $\vec{x}^{\mu(k)}$ and the center of mass $\vec{P}^k$ of the cluster to which it belongs. The larger the concentration parameter κ for the generation of the artificial data set, the smaller the dispersion of the clusters, and thus the larger $\cos(\vartheta_{\mathrm{kk}})$. If instead, an input pattern from cluster l is presented, that same DGC will receive a lower initial input:

$$I_i = \vec{w}_i \cdot \vec{x}^{\mu(l)} = ||\vec{w}_i|| \cdot ||\vec{x}^{\mu(l)}|| \cdot \cos(\vartheta_{\mathrm{kl}}) \approx \frac{||\vec{w}_i||}{1+\xi^2} \tag{27}$$

The approximation holds for a small dispersion of the clusters (large concentration parameter κ). We note that there is no subtraction of the recurrent input yet, because output units are initialized with zero firing rate before each pattern presentation. By definition, similar patterns stimulate (initially) the same DGCs. A DGC can be active for two clusters only if its threshold is:

$$b_i < \frac{||\vec{w}_i||}{1+\xi^2} \tag{28}$$

Therefore, with a high concentration parameter κ, patterns of different clusters are similar if $\xi < \sqrt{\frac{||\vec{w}_i||}{b_i} - 1}$, while patterns of different clusters are distinct if $\xi > \sqrt{\frac{||\vec{w}_i||}{b_i} - 1}$.

Parameter choice

The upper bound of the expected L2-norm of the feedforward weight vector towards the DGCs at convergence can be computed, see equation (10). With the parameters in Table 1(Simple network), the value is $||\vec{w}_i|| \leqslant 1.5$. Moreover, the input patterns for each cluster are highly concentrated, hence their angle with the center of mass of the cluster they belong to is close to 0, so we have $||\vec{w}_i|| \approx 1.5$. Therefore, at convergence, a DGC selective for a given cluster k receives an input $I_{i^*} = \vec{w}_{i^*} \cdot \vec{x}^{\mu(k)} \approx 1.5$ upon presentation of input patterns $\vec{x}^{\mu(k)}$ belonging to cluster k. We thus set $b_i = 1.2$ to satisfy condition (24). The threshold value ξ_{thresh} for which two clusters are similar (and above which two clusters are distinct) can be determined by equation (28) : $\xi_{\mathrm{thresh}} = 0.5$. We created a handmade data set with $\xi = 0.2$ for the case of similar clusters (therefore with similarity $s = 0.8$), and a handmade data set with $\xi = 0.8$ for the distinct case (hence with similarity $s = 0.2$).

Let us suppose that the weights of DGC i have converged and made this cell respond to patterns from cluster i. If another DGC k of the network is selective for cluster k, it ultimately gets the input $I_k = \vec{w}_k \cdot \vec{x}^{\mu(k)} + w_{\mathrm{rec}} \approx \frac{1.5}{1+\xi^2} + w_{\mathrm{rec}}$ upon presentation of input patterns $\vec{x}^{\mu(k)}$ belonging to cluster k ($k \neq i$). Hence, to satisfy condition (25), we need $w_{\mathrm{rec}} < b_i - \max_\xi \left(\frac{1.5}{1+\xi^2}\right)$. Furthermore, a newborn DGC is born with a null feedforward

weight vector, hence at birth, its input consists only of the indirect excitatory input from mature DGCs: $I_i = -w_{\text{rec}} > 0$. For the feedforward weight vector to grow, the condition $-w_{\text{rec}} > b_{\text{birth}}$ is also necessary (with b_{birth} the neuronal threshold of a newborn DGC at birth). We set $w_{\text{rec}} = -1.2$ and $b_{\text{birth}} = 0.9$, which satisfy the two above conditions.

Neurogenesis with the artificial data set

To save computation time, we initialize the feedforward weight vectors of two mature DGCs at two training patterns randomly chosen from the first two clusters, normalized such that they have an L2-norm of 1.5. We then present patterns from clusters 1 and 2, and let the feedforward weights evolve according to equation (5) until they reach convergence.

We thereafter introduce a novel cluster of patterns as well as a newborn DGC in the network. The sequence of presentation of patterns from the three clusters (a novel one and two pretrained ones) is random. The newborn DGC is born with a null feedforward weight vector, and its maturation follows the same rules as before. In the early phase, GABAergic input has an excitatory effect [14] and the newborn DGC does not inhibit the mature DGCs [9]. This is modeled by setting $w_{\text{rec}}^{NM} = -w_{\text{rec}}$ for the connections from mature to newborn DGC, and $w_{\text{rec}}^{MN} = 0$ for the connections from newborn to mature DGCs. The threshold of the newborn DGC starts at 0.9 at birth, mimicking enhanced excitability [6, 7], and increases linearly up to 1.2 (same threshold as the mature DGCs) over 12000 pattern presentations, reflecting loss of excitability with maturation. The exact time window is not critical. In the late phase of maturation of the newborn DGC, GABAergic input switches to inhibitory [14], and the newborn DGC recruits feedback inhibition onto mature DGCs [9]. It is modeled by switching the sign of the connection from mature to newborn DGC: $w_{\text{rec}}^{NM} = w_{\text{rec}}$, and establishing connections from newborn to mature DGCs: $w_{\text{rec}}^{MN} = w_{\text{rec}}$. Each of the 6000 patterns is presented once during the early phase of maturation, and once during the late phase of maturation.

The above paradigm is run separately for each of the two handmade data sets: the one where clusters are similar ($s = 0.8$), and the one where clusters are distinct ($s = 0.2$).

References methods

[49] Li, Y. *et al.* Molecular layer perforant path-associated cells contribute to feed-forward inhibition in the adult dentate gyrus. *Proceedings of the National Academy of Sciences* **110**, 9106–9111 (2013).

[50] Miller, K. D. & Fumarola, F. Mathematical equivalence of two common forms of firing rate models of neural networks. *Neural computation* **24**, 25–31 (2012).

[51] Senzai, Y. & Buzsáki, G. Physiological properties and behavioral correlates of hippocampal granule cells and mossy cells. *Neuron* **93**, 691–704 (2017).

[52] Chawla, M. *et al.* Sparse, environmentally selective expression of Arc RNA in the upper blade of the rodent fascia dentata by brief spatial experience. *Hippocampus* **15**, 579–586 (2005).

[53] McHugh, T. J. *et al.* Dentate gyrus NMDA receptors mediate rapid pattern separation in the hippocampal network. *Science* **317**, 94–99 (2007).

[54] Pfister, J.-P. & Gerstner, W. Triplets of spikes in a model of spike timing-dependent plasticity. *J. Neuroscience* **26**, 9673–9682 (2006).

[55] Chistiakova, M., Bannon, N., Bazhenov, M. & Volgushev, M. Heterosynaptic plasticity: Multiple mechanisms and multiple roles. *The Neuroscientist* **20**, 483–498 (2014).

[56] Zenke, F. & Gerstner, W. Hebbian plasticity requires compensatory processes on multiple timescales. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**, 1–17 (2017).

[57] LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86**, 2278–2324 (1998).

[58] Mardia, K. V. & Jupp, P. E. *Directional statistics*, vol. 494 (John Wiley & Sons, 2009).

6 Acknowledgments

This research was supported by the Swiss National Science Foundation (no. 200020_184615) and by the European Union Horizon 2020 Framework Program under grant agreement no. 785907 (HumanBrain Project, SGA2).

7 Authors contributions

O.G. developed the model and carried out the simulations. W.G. participated in discussions and helped designing the project. O.G. and W.G. wrote and validated the manuscript.

8 Competing interests statement

The authors declare no competing financial interests.

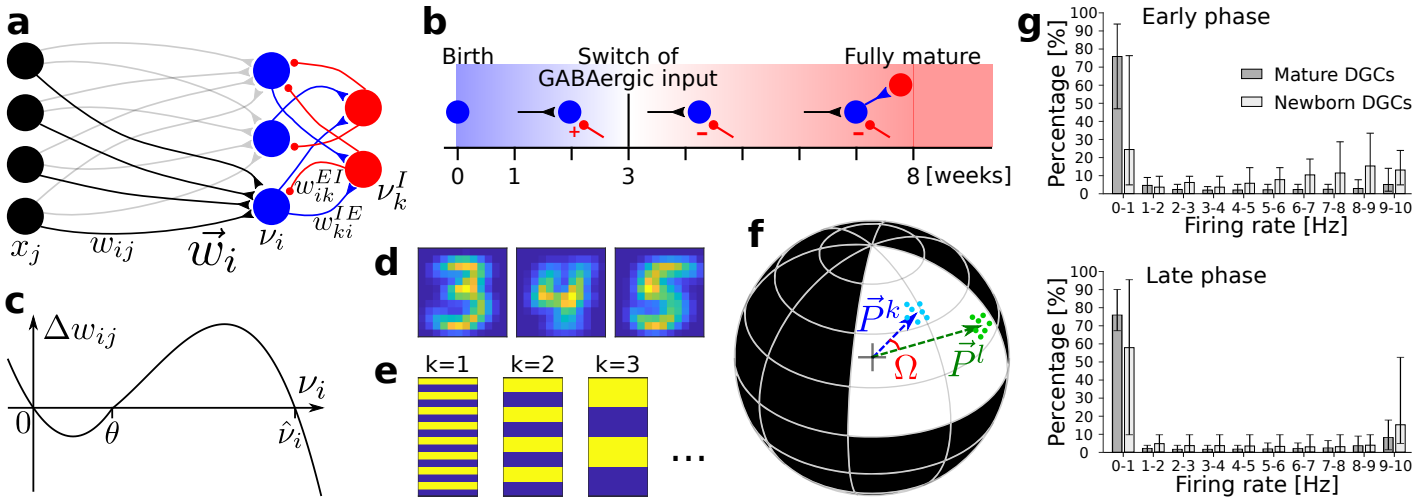

Figure 1: Model network and input patterns. (**a**) Network structure. EC neurons (black, rate x_j) are fully connected with weights w_{ij} to DGCs (blue, rate ν_i). The feedforward weight vector $\vec{w}_i$ onto neuron i is depicted in black. DGCs and interneurons (red, rate ν_k^I) are mutually connected with probability p_{IE} and p_{EI} and weights w_{ki}^{IE} and w_{ik}^{EI}, respectively. Connections with a triangular (round) end are glutamatergic (GABAergic). (**b**) Integration of an adult-born DGC (blue) as a function of time: GABAergic synaptic input (red) switches from excitatory (+) to inhibitory (-), glutamatergic synaptic input (black), interneuron (red). (**c**) Weight update Δw_{ij} as a function of the firing rate ν_i of the postsynaptic DGC showing LTD for $\nu_i < \theta$ and LTP for $\theta < \nu_i < \hat{\nu}_i$. (**d**) Center of mass for three ensembles of patterns from the MNIST data set, visualized as 12x12 pixel patterns. (**e**) Center of mass of three clusters of the artificial data set, visualized as 16x8 pixel patterns. 2-dimensional arrangement and color scale are for visualization only. (**f**) Center of mass of clusters k and l of the artificial data set ($\vec{P}_k$ and $\vec{P}_l$ respectively, separated by angle Ω) are represented by arrows that point to the surface of a hypersphere. Dots represent individual patterns. (**g**) Distribution of the percentage of model DGCs (mean with 10th and 90th percentiles) in each firing rate bin at the end of the early and late phase of maturation during stimulation with MNIST patterns. Percentages are per subpopulation.

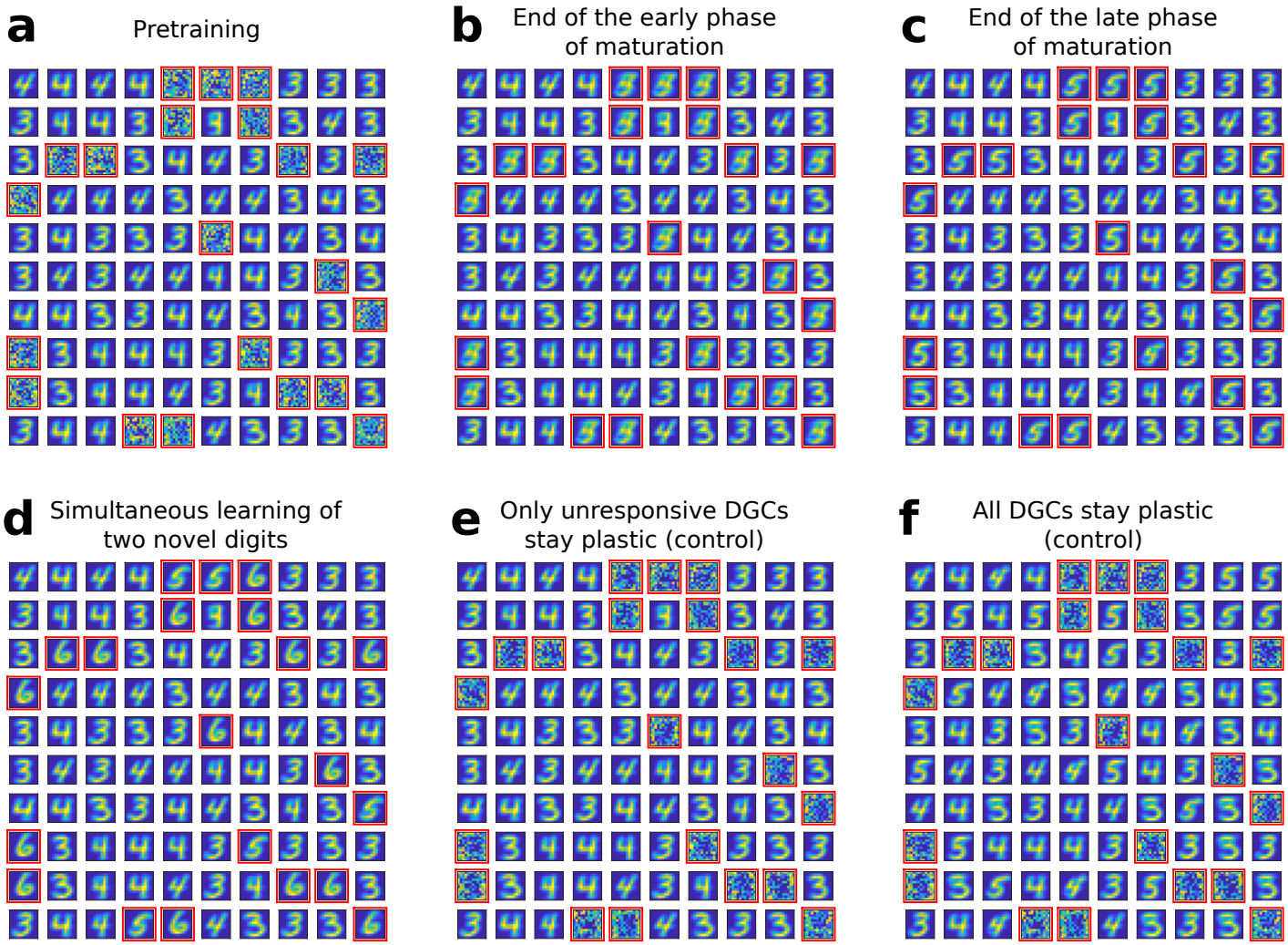

Figure 2: Newborn DGCs learn novel patterns. Receptive fields, defined as the set of feedforward weights, are represented in a 2-dimensional organization. (**a**) During pretraining, patterns from MNIST digits 3 and 4 are presented to the network. At the end of pretraining, 79 DGCs have receptive fields corresponding to threes and fours, while 21 remain unselective (highlighted by red frames). (**b**) Unselective neurons are replaced by newborn DGCs, which adapt their feedforward weights while patterns from digits 3, 4, and 5 are presented. At the end of the early phase of maturation, the receptive fields of all newborn DGCs (red frames) show mixed selectivity. (**c**) At the end of the late phase of maturation, newborn DGCs are selective for patterns from the novel digit 5, with different writing styles. (**d**) Several novel digits can be learned simultaneously. After pretraining as in (a), unresponsive neurons are replaced by newborn DGCs. When patterns from digits 3, 4, 5, and 6 are presented in random order, newborn DGCs exhibit after maturation receptive fields with selectivity for the novel digits 5 and 6. (**e**) Control without neurogenesis. After pretraining as in (a), the unresponsive neurons stay plastic, but they fail to become selective for digit 5 when patterns from digits 3, 4, and 5 are presented in random order. (**f**) If all DGCs stay plastic (no neurogenesis) when patterns from digit 5 are introduced, some of the DGCs previously responding to patterns from digits 3 or 4 become selective for digit 5.

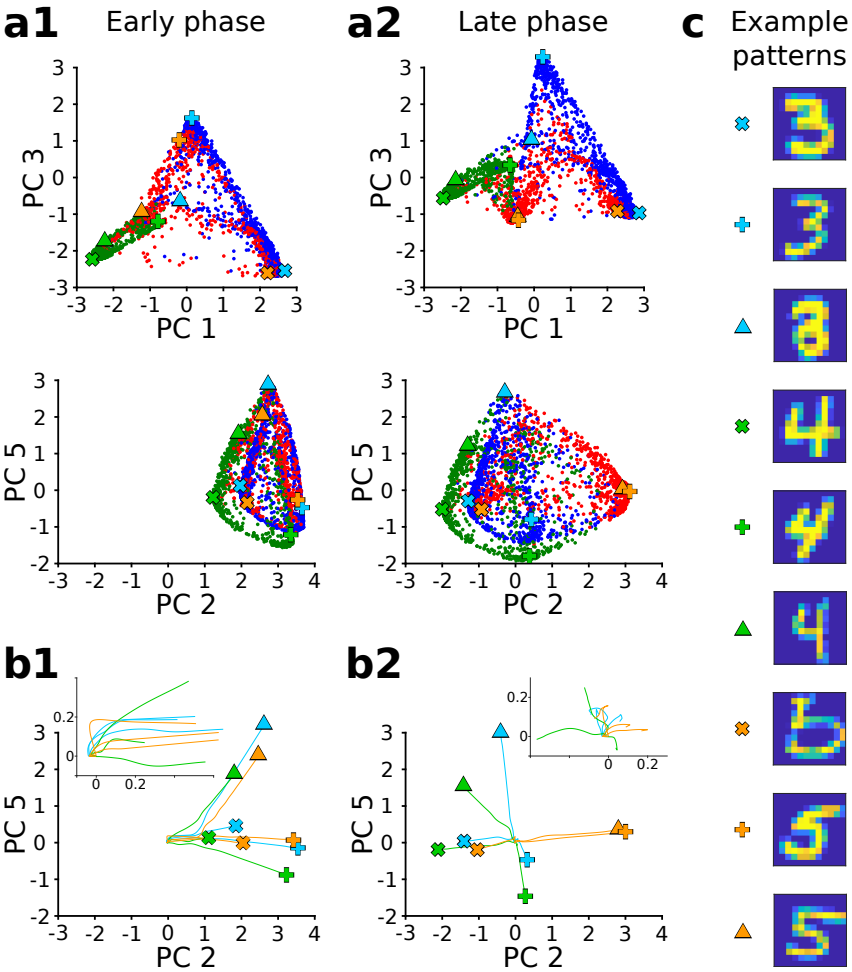

Figure 3: Novel patterns expand the representation into a previously empty subspace. (**a**) Projections of the final firing rates of all 100 DGCs on PC 1 and PC 3 (first row), and on PC 2 and PC 5 (second row), at the end of the early (a1) or late (a2) phase of maturation of the newborn DGCs. Each pattern of the MNIST test set corresponds to one point. Color indicates digit 3 (blue), 4 (green) or 5 (red). The PCs were determined at the end of the late phase of maturation. (**b**) Trajectories of firing rates in the PC-space for a few example patterns at the end of the early (b1) and late (b2) phase of maturation of newborn DGCs. It takes about 170 to 200 ms for the trajectories to converge to their final points, where symbols mark the corresponding patterns. Insets: zoom on trajectories during the first 10 ms. (**c**) Example patterns from the test set, with the symbols used in (a) and (b). Note that the fives corresponding to the orange swiss cross and the orange triangle are represented far from each other (and close to examples of resp. threes and fours) at the end of the early phase of maturation (a1,b1), but close to each other and far from any threes and fours at the end of maturation (a2,b2).

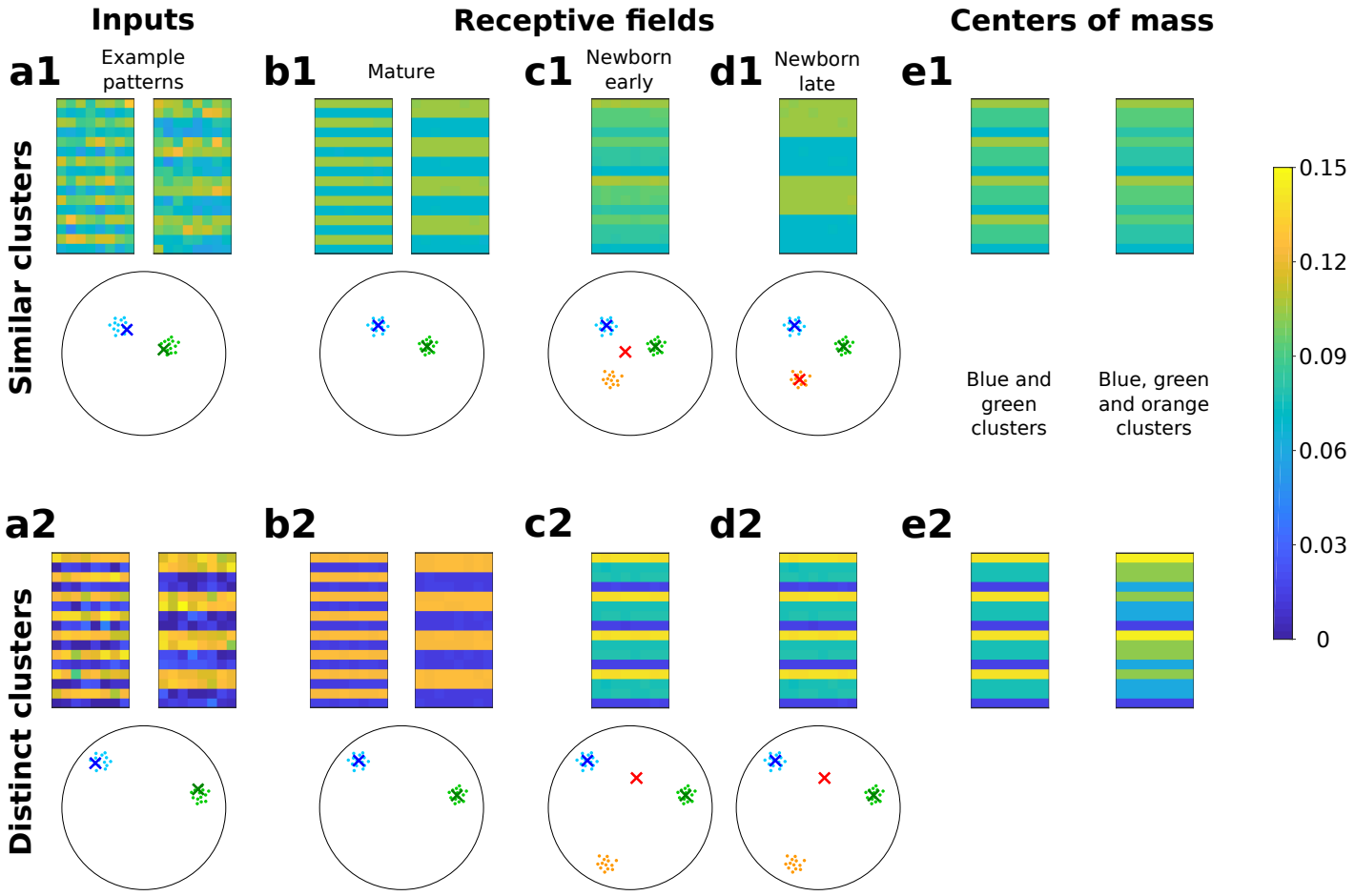

Figure 4: A newborn DGC becomes selective for similar but not distinct novel stimuli. (**a**) Example input patterns (activity of 16x8 input neurons) from clusters 1 and 2 for similar clusters (a1, $s = 0.8$), and distinct clusters (a2, $s = 0.2$). Below: dots correspond to patterns, crosses indicate the input patterns shown (schematic). (**b**) After pretraining with patterns from two clusters, the receptive fields (set of synaptic weights onto neurons 1 and 2) exhibit the center of mass of each cluster of input patterns (blue and green crosses). (**c**) Novel stimuli from cluster 3 (orange dots) are added. If the clusters are similar, the receptive field of the newborn DGC (red cross) moves towards the center of mass of the three clusters during its early phase of maturation (c1), and if the clusters are distinct towards the center of mass of the two pretrained clusters (c2). (**d**) Receptive field after the late phase of maturation for the case of similar (d1) or distinct (d2) clusters. (**e**) Total center of mass of all patterns of the blue and green clusters (left column) and of the blue, green and orange clusters (right column) for the case of similar (e1) or distinct (e2) clusters. Color scale: input firing rate x_j or weight w_{ij} normalized to $||\vec{w}_i|| = 1 = ||\vec{x}||$.

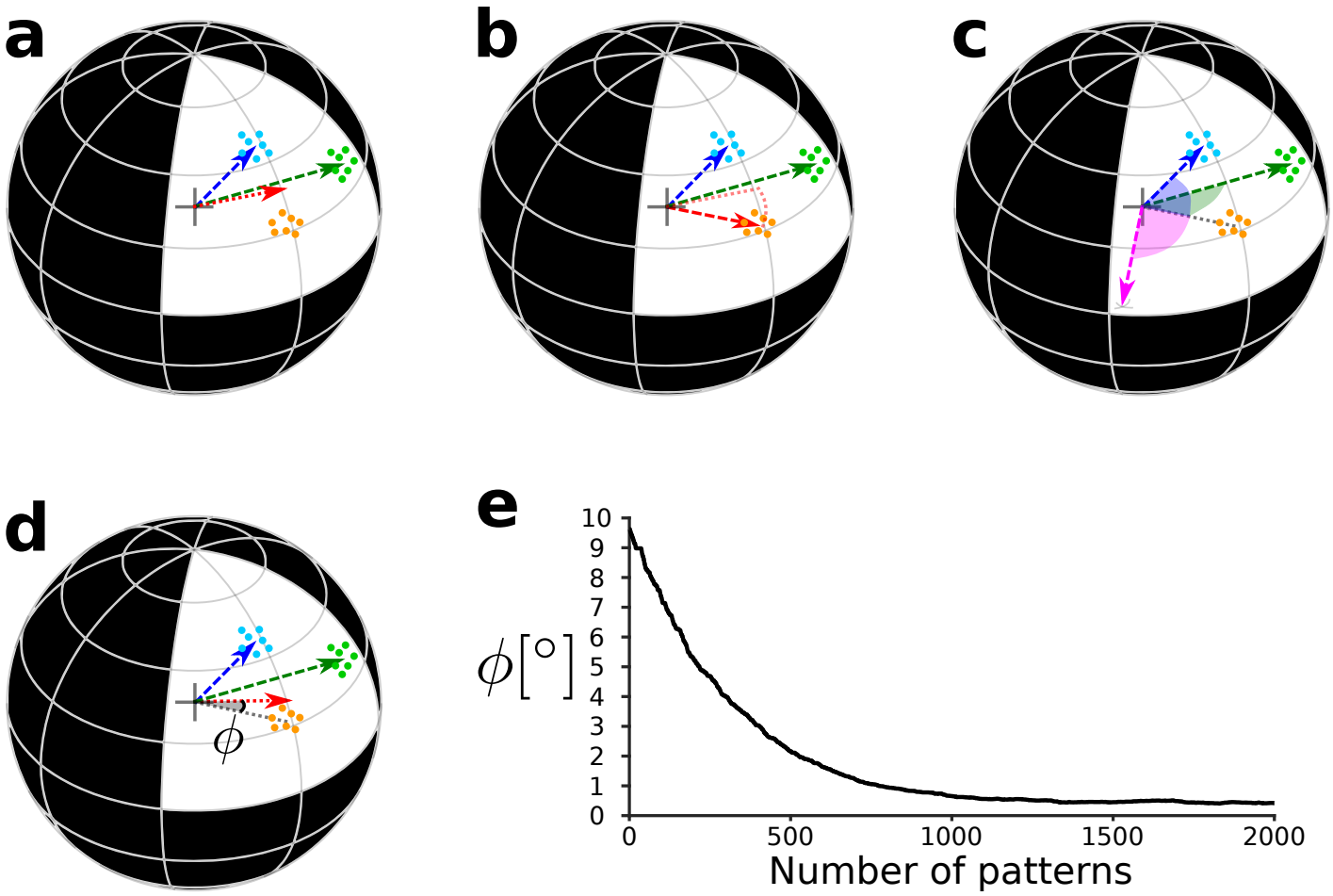

Figure 5: Maturation dynamics for similar patterns. Schematics of the unit hypersphere with three clusters of patterns (colored dots) and three normalized feedforward weight vectors (colored arrows). (**a**) After pretraining, the blue and green weight vectors point to the center of mass of the corresponding clusters. Patterns from the novel cluster (orange points) are presented only later to the network. During the early phase of maturation, the newborn DGC grows its vector of feedforward weights (red arrow) in the direction of the subspace of patterns which indirectly activate the newborn cell. (**b**) During the late phase of maturation, the red vector turns towards the novel cluster. (**c**) Control scenario: an unresponsive neuron with magenta feedforward weight vector is unable to learn the novel cluster. (**d**) Angle ϕ between the center of mass of the novel cluster and the feedforward weight vector onto the newborn cell. (**e**) The angle ϕ decreases in the late phase of maturation of the newborn DGC if the novel cluster is similar to the previously stored clusters. Its final average value of $\phi \approx 0.4$ is caused by the jitter of the weight vector around the center of mass of the novel cluster.

Table 1: Parameters for the simulations

	Biologically-plausible network		Simplified network	
Network	$N_{EC} = 144$ $N_I = 25$	$N_{DGC} = 100$	$N_{EC} = 128$	$N_{DGC} = 3$
Connectivity	$w_{IE} = 1$ $p_{IE} = 0.9$	$w_{EI} = -\frac{1}{p_{EI} * N_I}$ $p_{EI} = 0.9$	$w_{rec} = -1.2$	
Dynamics	$\tau_m = 20$ ms $L = 0.5$	$\tau_{\text{inh}} = 2$ ms $p^* = 0.1$	$\tau_m = 20$ ms	
Plasticity	$\alpha_0 = 0.05$ $\gamma_0 = 10$ $\nu_0 = 0.2$	$\beta = 1$ $\theta = 0.15$ $\gamma = 9.85$	$\alpha_0 = 0.03$ $\gamma_0 = 1.65$ $\gamma = 1.5$	$\beta = 1$ $\theta = 0.15$
Numerical simulations	$\Delta t = 0.1$ ms $\eta_b = 0.01$	$\eta = 0.01$	$\Delta t = 1$ ms	$\eta = 0.01$