Intrinsic Spine Dynamics Are Critical for Recurrent Network Learning in Models With and Without Autism Spectrum Disorder

James Humble 1, Kazuhiro Hiratsuka 1, Haruo Kasai 2 and Taro Toyoizumi 1*

1 Laboratory for Neural Computation and Adaptation, RIKEN Center for Brain Science, Saitama, Japan, 2 Laboratory of Structural Physiology, Faculty of Medicine, Center for Disease Biology and Integrative Medicine, University of Tokyo, Tokyo, Japan

It is often assumed that Hebbian synaptic plasticity forms a cell assembly, a mutually interacting group of neurons that encodes memory. However, in recurrently connected networks with pure Hebbian plasticity, cell assemblies typically diverge or fade under ongoing changes of synaptic strength. Previously assumed mechanisms that stabilize cell assemblies do not robustly reproduce the experimentally reported unimodal and long-tailed distribution of synaptic strengths. Here, we show that augmenting Hebbian plasticity with experimentally observed intrinsic spine dynamics can stabilize cell assemblies and reproduce the distribution of synaptic strengths. Moreover, we posit that strong intrinsic spine dynamics impair learning performance. Our theory explains how excessively strong spine dynamics, experimentally observed in several animal models of autism spectrum disorder, impair learning associations in the brain.

Keywords: spine dynamics, synaptic weight distribution, cell assembly, autism spectrum disorder, recurrent network, learning

INTRODUCTION

The operation of a neural circuit is shaped by the strength of synapses that mediate signal transduction between neurons. Activity-dependent modification of synaptic strength, termed synaptic plasticity, is considered to be an underlying mechanism of learning and memory (Malenka and Bear, 2004; Mongillo et al., 2017). A major form of synaptic plasticity is Hebbian plasticity (Hebb, 1949). While there are multiple molecular mechanisms (Malinow and Malenka, 2002; Matsuzaki et al., 2004; Nicoll et al., 2006) underlying Hebbian plasticity and experimental protocols (Neves et al., 2008), it is commonly induced by coactivation of pre- and postsynaptic neurons within a particular time window. One prominent biological mechanism for Hebbian plasticity is activity-dependent spine volume change. Spine volume is known to be tightly correlated with synaptic strength (Matsuzaki et al., 2001; Smith et al., 2003; Noguchi et al., 2005; Béïque et al., 2006; Asrican et al., 2007; Holbro et al., 2009; Zito et al., 2009), and both long-term potentiation (LTP) and long-term depression (LTD) involve spine change (Lang et al., 2004; Matsuzaki et al., 2004; Otmakhov et al., 2004; Zhou et al., 2004; Kopec et al., 2006; Hayama et al., 2013).

It was previously proposed (Hebb, 1949; Amari S., 1977; Hopfield, 1982) that a memory can be represented by coherent activity in a cell assembly, i.e., a group of cells mutually exciting each other,
and the memory can be stored in synaptic strengths between these neurons by Hebbian plasticity. Consistently, recent experiments have shown that the activation of a coherently active group of cells is necessary and sufficient for the expression of learned behavior (Liu et al., 2012; Nabavi et al., 2014). However, how a neural circuit maintains cell assemblies stably is not well understood. In some models, cell assemblies are stable because synaptic strength is modified only during learning, and then fixed (Amari S. I., 1977; Hopfield, 1982; Vogels et al., 2011). However, these models neglect changes in synaptic strength after learning and thus do not address the maintenance of an acquired memory.

Several studies modeled ongoing Hebbian plasticity during spontaneous activity and found that Hebbian plasticity alone is likely not sufficient to maintain a cell assembly. In additive Hebbian plasticity models (Gerstner et al., 1996; Song et al., 2000; Gütig et al., 2003), in which the dependencies of the LTP and LTD amplitudes on synaptic strength are the same, memory tends to become unstable due to a positive feedback during spontaneous activity (Fiete et al., 2010; Litwin-Kumar and Doiron, 2014; Zenke et al., 2015), namely, neurons that fire together are wired together, and then fire together more often. This kind of a positive feedback process typically fuses assemblies and expands the largest existing cell assembly. Some forms of stabilizing mechanisms, such as inhibitory plasticity (Litwin-Kumar and Doiron, 2014), homeostatic plasticity (Zenke et al., 2013), or heterosynaptic plasticity (Zenke et al., 2015) have been suggested to stabilize memory (Keck et al., 2017). However, even with these stabilizing mechanisms, the resulting distribution of synaptic strengths often becomes dissimilar to what has been experimentally observed (Toyoizumi et al., 2007). For example, while the models with positive feedback often produce a synaptic strength distribution that is bimodal, experiments have reported a unimodal and long-tailed distribution of synaptic strengths (Song et al., 2005; Cossell et al., 2015) and corresponding spine volumes (Yasumatsu et al., 2008; Loewenstein et al., 2011).

An alternative proposal is multiplicative Hebbian plasticity (van Rossum et al., 2000; Gütig et al., 2003; Morrison et al., 2007), in which the LTP amplitude is less prominent than the LTD amplitude for large synapses, in agreement with experimental observations (Bi and Poo, 1998; Tanaka et al., 2008; Hayama et al., 2013). This multiplicative form of Hebbian plasticity can avoid the above instability problem, and under spontaneous activity of neurons, synaptic strengths converge near a prefixed set point where LTP and LTD effects balance each other, regardless of the initial synaptic strengths (Morrison et al., 2007). This means that memories degrade in the presence of spontaneous neural activity.

Hence, in all the models described above, it is nontrivial to stably maintain cell assemblies and reproduce the experimentally observed distribution of synaptic strengths (or of spine volumes), which has a thick tail and a peak at a rather weak strength (Song et al., 2005; Yasumatsu et al., 2008; Loewenstein et al., 2011; Cossell et al., 2015). Interestingly, a similar distribution of spine volumes is robustly observed even in animal models of mental disorders (Pathania et al., 2014) and with an LTD deficiency in calcineurin KO animals (Okazaki et al., 2018). In contrast, in the above mathematical models, the distribution of synaptic strengths is fragile and strongly depends on the balance of LTP and LTD that is set by model parameters and input to neurons. Thus, conventional models have no mechanism to restore the distribution of synaptic strengths.

Despite the common assumption that only synaptic plasticity changes spines, they also dynamically change in the absence of neural activity. Recent studies showed that spine turnover, i.e., generation and elimination of spines, continues even under the blockade of neural activity and calcium signaling in vivo (Kim and Nabekura, 2011; Nagaoka et al., 2016; Ziv and Brenner, 2018). Further, spine volumes constitutively fluctuate in the absence of neural activity, calcium signaling, and activity-dependent plasticity in vitro (Yasumatsu et al., 2008). These intrinsic spine dynamics are characterized by a zero drift coefficient and a diffusion coefficient proportional to the square of spine volume, \( \nu^2 \). Interestingly, this volume-dependent diffusion reproduces the experimentally observed equilibrium distribution of spine volumes with a power-law tail of exponent \( \nu^{-2} \) (Yasumatsu et al., 2008; Ishii et al., 2018). This observation poses an important question: How do intrinsic spine dynamics affect the maintenance of cell assemblies?

We address this question by simulating a mathematical model of a recurrently connected neural network that implements both multiplicative spike-timing-dependent plasticity (STDP) (van Rossum et al., 2000; Morrison et al., 2007) and experimentally observed intrinsic spine dynamics (Yasumatsu et al., 2008). We also study how spine turnover and the distribution of spine volumes are affected by these two processes. Despite a possible perception of intrinsic spine dynamics as noise, we show that they can help to maintain cell assemblies by preventing unnecessary spines from growing and sustaining the physiological spine volume distribution.

Based on the model analysis, we hypothesize that intrinsic spine dynamics that are stronger than in wild type (WT) conditions can explain the abnormally high spine turnover rate observed in animal models of autism spectrum disorder (ASD) (Pan et al., 2010; Ishihki et al., 2014). By fitting model parameters to one ASD mouse model, fmr1KO (a model of fragile X syndrome) (Pfeiffer and Huber, 2009), we show how excessively strong intrinsic spine dynamics may cause learning deficits in ASD animals (Silverman et al., 2010; Padmashri et al., 2013).

RESULTS

We study the dynamics of spine volumes using a model of a cortical circuit. We stimulate recurrently connected spiking neurons (Figure 1; also see Methods) to explore if the network stores memories as cell assemblies. The cortical network consists of 1,000 excitatory and 200 inhibitory leaky-integrate-and-fire spiking neurons (Tuckwell, 1988), where the excitatory and inhibitory neurons are randomly connected by a 10% connection probability (Figure 1A). For simplicity, the excitatory neurons are embedded in a one-dimensional feature space that describes, for example, orientation selectivity in V1. We assume that synapses can be formed between a pair of excitatory neurons that have potential connectivity (Markram et al., 1997). Potential connectivity (either 0 or 1) from a neuron to another is randomly
A model of cortical circuitry. Excitatory and inhibitory neurons are modeled as leaky-integrate-and-fire units, which are sparsely connected. We assume that only the recurrent excitatory synapses are plastic. Excitatory neurons are aligned in a one-dimensional feature space, which is divided into four neighborhood quarters. Forty percent of randomly chosen excitatory neurons in each quarter comprise a stimulation group. The four stimulation groups are randomly stimulated one at a time during the learning period. (B1) Potential connectivity peaks at around 10% and decays with the tuning-distance between two excitatory neurons in the feature space. Synapses can grow if two excitatory neurons are potentially connected. (B2) If two excitatory neurons have potential connectivity, the number of contact points is randomly drawn from a truncated Poisson distribution in the range of 1 to 10. Each contact point can accommodate one spine. (B3) Spiking activity and membrane potential dynamics of a sample set of neurons at baseline. (B4) Excitatory neurons have an adaptation current, which builds up with firing activity and suppresses firing rate. During a learning period, one of the stimulated groups is randomly chosen with probability $\frac{1}{4}$ and receives elevated external input for 3 s. All inhibitory neurons receive external stimulation throughout the entire learning period. (C) Spine volumes are changed by the combination of STDP and intrinsic spine dynamics (except in Figure 2, where only STDP is considered). Arrows indicate possible changes in spine volume and the size of the arrow represents the possible maximum change in spine volume. A threshold at 0.02 $\mu$m$^3$ separates spines and non-spines, and STDP only affects spines. (D) The multiplicative STDP rule used for changing spine volume. The LTD amplitude is proportional to spine volume. (E) The diffusion coefficient characterizing intrinsic spine dynamics, which is proportional to the square of spine volume.

The multiplicative STDP rule used for changing spine volume. The LTD amplitude is proportional to spine volume. (F) The diffusion coefficient characterizing intrinsic spine dynamics, which is proportional to the square of spine volume.
we define a tiny protrusion of volume \( <0.02 \, \mu m^3 \) as a “non-spine”) because the correlation between the synaptic strength and spine volume has been experimentally demonstrated (Matsuzaki et al., 2004; Harvey and Svoboda, 2007; Bosch et al., 2014). The spine volume of an excitatory to excitatory synapse is modeled by multiplicative STDP (Figure 1E) (van Rossum et al., 2000), and thus the LTD amplitude is independent of synaptic strength, while the LTD amplitude is proportional to synaptic strength. Therefore, with an increase in synaptic strength, the LTD amplitude increases at a steeper rate than the LTP amplitude, and this is consistent with experimental observations (Bi and Poo, 1998; Tanaka et al., 2008; Hayama et al., 2013). We let the network acquire cell assemblies by providing additional external input to subsets of neurons. We divide the feature space of the excitatory network into four equally sized neighboring quarters (Figure 1A), and randomly select 40% of the neurons in each quarter as a stimulated group. We randomly select one of these four groups at a time during the learning period and stimulate it with an elevated rate of Poisson spikes for 3s (Figure 1C; see Methods). For simplicity, all inhibitory neurons are stimulated throughout the learning period. The spine volume, \( v \), of each spine is initially drawn randomly from a fixed distribution, proportional to \((v + 0.05 \, \mu m^3)^{-2}\). This initial distribution approximates an experimentally observed spine volume distribution (see Methods). As we will see below, changing synaptic strengths by multiplicative STDP alone fails to sustain cell assemblies in the presence of spontaneous activity.

Learning Without Intrinsic Spine Dynamics

Firstly, we consider the case where intrinsic spine dynamics are absent and spine volumes are only changed by multiplicative STDP. Figure 2 depicts the behavior of our network during and after the learning period. The learning period (Figure 2A) is terminated when the mean intra-group spine volume of at least one stimulated group reaches \( \geq 0.49 \, \mu m^3 \). During the learning period, four groups of neurons were randomly stimulated one at a time (Figure 2B, Left) with increased input, and after the learning period, only group 3’s neurons stayed active (Figure 2B, Right). Figure 2C plots the firing rates of all neurons in the 3rd-quarter during the entire simulation period. This shows that the number of active neurons monotonically increased both during the learning period and during the maintenance period, indicating an unstable learning outcome. Specifically, the cell assembly initially formed among the externally stimulated group 3 neurons and spread to neighboring neurons. This spreading of the cell assembly provided extra recurrent input from newly recruited neurons to this cell assembly. Figure 2D summarizes the population averaged firing rates of the four stimulated groups. All groups show elevated firing rates during the learning period due to external stimulations. After the learning period, physiological neural activity was maintained only for a day. After that, the mean firing rate of one group (i.e., group 3) exploded, and that of the other groups declined down to near zero values. The spine volume distribution at the end of the simulation included a non-physiological secondary peak at 0.5 \( \mu m^3 \) (Figure 2E). A small but non-negligible number of non-stimulated synapses also formed a peak at 0.5 \( \mu m^3 \). These synapses contribute to the extremely high firing rate of the group 3 neurons and further recruitment of non-stimulated neurons toward the end of the simulation. During the learning period, the mean volume of the spines connecting neurons within each stimulated group increases due to LTD (Figure 2F). Note that LTD is dominant over LTP for small spines because we assume that the LTD amplitude is fixed but the LTP amplitude is proportional to spine volume. After a few days of learning, mean spine volumes plateaued at around 0.5 \( \mu m^3 \), where LTD and LTP effects roughly balance. The mean spine volume of other spines (non-stimulated) exhibited a slow but steady increase (Figure 2F), reflecting the formation of the secondary peak of non-stimulated spines seen in the spine volume distribution (Figure 2E). The volumes of individual spines (Figure 2G) are homogeneous within each group and their development mirrors the corresponding mean spine volume (Figure 2F).

The spread of activity is slow in this simulation because of the lateral inhibition that tend to shut down spikes in surrounding neurons (Billings and van Rossum, 2009). However, the activity will eventually spread as long as neurons fire occasionally and spine volumes between active neurons converge near the prefixed LTP/LTD balancing point at around 0.5 \( \mu m^3 \). This convergence is innate to multiplicative STDP and spontaneously happens even in the absence of the learning period, albeit slowly (Figure S1). Hence, as expected from previous studies (Morrison et al., 2007), a recurrently connected network with multiplicative STDP has difficulty sustaining cell assemblies in spine volumes in the presence of spontaneous activity. Notably, the resulting spine volume distribution is dissimilar to an experimentally observed unimodal distribution with many small spines and a fat tail.

Learning With Normal Intrinsic Spine Dynamics

Next, we add intrinsic spine dynamics previously observed in hippocampal slices (Yasumatsu et al., 2008). Spine volumes fluctuate every day, and the amplitude of these fluctuations is spine volume dependent (Figure 1F). Importantly, these intrinsic spine dynamics were largely intact in the absence of neural activity and synaptic plasticity, i.e., under the pharmacological blockade of sodium channels, NMDA, and voltage-dependent calcium channels in cells. In the absence of neural activity and plasticity, the amplitude of spine volume fluctuations is roughly proportional to spine volume, \( \sim \alpha v + \beta \), where \( v \) is spine volume with parameters \( \alpha = 0.2 \, \text{day}^{-1/2} \) and \( \beta = 0.01 \, \mu m^3 \, \text{day}^{-1/2} \) (Yasumatsu et al., 2008). In other words, the effect of intrinsic spine dynamics is summarized by the volume-dependent diffusion coefficient \( D(v) = (\alpha v + \beta)^2 /2 \) with zero drift coefficient (Yasumatsu et al., 2008) (see also Methods). Therefore, the Fokker-Planck equation (Riskin, 1989) for describing the evolution of the spine volume distribution \( P(v,t) \) is \( \partial P(v,t)/\partial t = (\partial^2 /\partial v^2)\{D(v) P(v,t)\} \). It indicates that the equilibrium is reached when the diffusion intensity \( D(v) P(v,t) \) becomes volume-independent. This gives the equilibrium spine
FIGURE 2 | Network behavior in the absence of intrinsic spine dynamics. (A) Experimental protocol. External stimulus is provided during the learning period (blue bar) and the memory retention is studied in the maintenance period. The learning period finishes when one cell assembly becomes strong enough to sustain its activity. (B) Typical neural activity during (Left) and after (Right) the learning period. The panels show the firing rates of 1,000 excitatory neurons in a 50 s time window. During the learning, neural activity is driven by external stimulus, which randomly activates one group at a time. After learning, one group of neurons is strongly active spontaneously. (C) Firing rates of the 3rd-quarter neurons in the entire simulation period. Neurons are sorted by the first time their firing rates exceed 15 Hz. The number of active neurons expands both during the learning and retention periods beyond the initial 40% that are stimulated. (D) Mean firing rates of the four stimulated groups. (E) Spine volume distribution at the end of the simulation for intra-stimulated-group spines (Stimulated: pink) and other spines (Other: green). (F) Mean spine volume of intra-stimulated-group spines (pink) and other spines (green). (G) Individual volumes of a single intra-stimulated-group spine from each stimulated group and a single non-stimulated spine. The learning period is represented by a blue bar.
volume distribution $P_{eq}(v) \propto 1/D(v) \sim v^{-2}$, which has a power-law tail. Note that there are two mathematical conventions for interpreting the above equation (Gardiner, 1985), which lead to different semantic meanings of fluctuation. Here, we take the Itô interpretation, in which the intrinsic spine dynamics are interpreted as spine volume fluctuations (but see Methods for an alternative interpretation). We regard that spines smaller than 0.02 $\mu$m$^3$ are non-spines and do not exhibit multiplicative STDP, whereas intrinsic spine dynamics are still present even for these small protrusions (cf. Figure 1D). This assumption is consistent with the experimental observations that the baseline spine turnover is largely activity-independent (Kim and Nabekura, 2011; Nagaoka et al., 2016).

We used the same stimulation protocol as in Figure 2 to study cell assembly learning when both multiplicative STDP and intrinsic spine dynamics are involved (Figures 3A–H). Cell assembly learning was similar at the beginning of the simulation to the case without intrinsic spine dynamics: Firing rates increased (Figure 3D) and the intra-group spines enlarged (Figure 3F). In contrast to the previous case, a physiological neural activity level was maintained throughout (Figure 3D), and none of the cell assemblies aggressively spread to neighboring non-stimulated neurons (Figure 3C) during the maintenance period. The activity-dependent formation of cell assemblies is evident from the coherent reactivation of stimulated groups (Figure 3B) during the maintenance period at much lower spontaneous firing rates than during the learning period. This spontaneous reactivation of cell assemblies is essential for the maintenance of memory patterns (similarly to Mi et al., 2017). Nonetheless, these memory patterns can be successfully maintained even if neural activity is blocked for a whole day (e.g., by tetrodotoxin; Figure S2). Notably, the mean volume of the intra-assembly spines was stably maintained (Figure 3F), although individual spine volumes fluctuated throughout the simulation (Figure 3G). In contrast to the previous case without intrinsic spine dynamics (cf. Figure 2E), the spine volume distribution remained unimodal after learning, with no secondary peak around 0.5 $\mu$m$^3$ (Figure 3E; cf. Figure 2E with a secondary peak). This is because a possible secondary peak at the tail was smeared by large fluctuations in large spines. Despite this smearing, the memory patterns were still stored by enlarged spines located around the tail of the distribution. Finally, spine generation was increased during the learning period, as often seen experimentally (Hofer et al., 2009; Yang et al., 2009) (Figure 3H). The contribution of each model mechanism assumption is further explored in Figure S3.

### A Cell Assembly Is Maintained by an Ensemble Property of Intra-Assembly Spines

To elucidate how intrinsic spine dynamics stabilize our network learning and enable the storing of memories, we examined the effects of STDP and intrinsic spine dynamics separately (Figure 4). We initialized spine volumes randomly as described above except for the intra-group spines of one group, whose volumes were all set identically and changed systematically. While we simulated the spontaneous activity of neurons, we monitored how multiplicative STDP and intrinsic spine dynamics changed the mean intra-group spine volume, respectively. Multiplicative STDP leads to an overall potentiation of the intra-group spines (Figure 4; orange line) if they are small (roughly $<0.57 \mu$m$^3$) and a depression of the spines if they are large (roughly $>0.57 \mu$m$^3$). This transition is observed for large spines because the LTD amplitude, which is proportional to spine volume, dominates over the amplitude of LTP, which is independent of spine volume. LTP dramatically increases for spines $>0.30 \mu$m$^3$ because strong intra-group connections serve as positive feedback on coincident firing between the neurons within the group, increasing the frequency of both LTP and LTD events. Therefore, STDP on its own leads to one fixed point of mean spine volume at a non-physiologically high value at around 0.57 $\mu$m$^3$. This fixed point is controlled by the parameter setting the relative amplitude of LTD (Morrison et al., 2007). Intrinsic spine dynamics, on the other hand, normalize the spine volume distribution, restoring the distribution toward the equilibrium distribution $P_{eq}(v) \propto 1/D(v)$ with a mean spine volume of roughly 0.15 $\mu$m$^3$ (Figure 4; brown line).

When the contributions from STDP and intrinsic spine dynamics are added together with an appropriate balance (Figure 4; black line), the combination permits a bi-stability in the mean spine volume of a population of spines. In this case, changes in the mean spine volume are dominated by the intrinsic spine dynamics when small (roughly $<0.30 \mu$m$^3$), so that the spine volumes fluctuate around 0.15 $\mu$m$^3$. Conversely, changes in the mean spine volume are significantly affected by STDP when large (roughly $>0.30 \mu$m$^3$), creating a larger-volume fixed point at approximately $<0.55 \mu$m$^3$. In between these stable fixed points, there is a separation point (unstable fixed point) at around 0.50 $\mu$m$^3$, which prevents the mean spine volume from moving in between these two stable fixed points. Therefore, while single spines are largely fluctuating and sporadically moving around the small and large mean spine-volume fixed points, a large cell assembly is stably maintained by the ensemble property of intra-assembly spines: here quantified as the mean intra-assembly spine volume.

These results suggest that intrinsic spine dynamics can normalize the synaptic strength distribution to a stereotypical shape in the presence of ongoing Hebbian plasticity, and at the same time, enable the circuit to stably retain memory patterns in the form of cell assemblies with a bi-stable mean intra-cell assembly spine volume. The relative amplitudes of STDP and intrinsic spine dynamics are key parameters for achieving this bi-stability. For example, if STDP is too strong, relative to intrinsic spine dynamics, the small mean spine volume fixed point at 0.15 $\mu$m$^3$ would disappear. We have already seen in our model that in the absence of intrinsic spine dynamics, the spine volume distribution sharply peaks at around the large mean spine volume fixed point and activity becomes too high (Figure 1). Instead, if STDP is too weak, relative to intrinsic spine dynamics, the large mean spine volume fixed point around 0.55 $\mu$m$^3$ could disappear. Below, we explore more generally what might go wrong with excessively strong intrinsic spine dynamics.
FIGURE 3 | Network behavior in the presence of intrinsic spine dynamics. (A–G) Conventions are as in Figure 2. (H) Spine generation and elimination.
Learning With Excess Intrinsic Spine Dynamics as a Model of Fragile X Syndrome

Interestingly, experimental results suggest that intrinsic spine dynamics are abnormally high in a mouse model of fragile X syndrome, fmr1KO (Pan et al., 2010; Nagaoka et al., 2016). Below, we constrain the parameters $\alpha$ and $\beta$ of the diffusion coefficient $D(v) = (\alpha v + \beta)^2/2$ to characterize intrinsic spine dynamics in fmr1KO mice based on reported observations. Spine turnover is about twice as high in fmr1KO mice as in WT mice (Pan et al., 2010). Remarkably, this elevated spine turnover largely remains even when calcium activity is pharmacologically blocked, suggesting that this is due to abnormal intrinsic spine dynamics (Nagaoka et al., 2016). Another constraint is the spine volume distribution. Experimental reports comparing the spine volume distribution in fmr1KO and WT mice have mixed observations–some studies detected more immature spines in fmr1KO but others detected no significant difference (He and Portera-Cailliau, 2013). For simplicity, we assume that any differences in the spine volume distribution between fmr1KO and WT mice are negligible. Based on the numerical fitting to the observed spine turnover, these two constraints specify the parameters $\alpha = 0.43 \text{ day}^{-1/2}$ and $\beta = 0.021 \mu\text{m}^3\text{ day}^{-1/2}$ for fmr1KO mice (Figure 5; see also Figure S4 for a systematic parameter search). Intuitively speaking, $\alpha$ and $\beta$ are twice as high as the corresponding WT values. This is because, despite the presence of STDP, the spine volume distribution is largely set by the equilibrium distribution of the intrinsic spine dynamics, $P_{eq}(v) \propto 1/(v + \beta/\alpha)^2$. It suggests that the ratio $\beta/\alpha$ should be similar between WT and fmr1KO to have similar spine volume distributions, and $\beta$ in fmr1KO is twice as large as WT to account for the doubled spine turnover rate.

Learning and memory deficits have been reported in fmr1KO mice (Padmashri et al., 2013). We investigated whether these learning and memory deficits could potentially be explained by the abnormal intrinsic spine dynamics modeled above. Consistent with the experimental observations, storing memory patterns was impaired in our fmr1KO model (Figures 6A–H), where coherent spontaneous activation of cell assemblies rapidly faded after learning (Figures 6B–D). This is because the mean spine volumes of the stimulated groups decreased during the maintenance period (Figure 6F). This effect can be intuitively understood from the result in Figure 4—memory cannot be stably stored in a cell assembly if intrinsic spine dynamics are too strong, relative to multiplicative STDP, because the stable fixed point of the mean intra-group spine volume around 0.55 $\mu\text{m}^3$ disappears. Individual spines fluctuated as in the WT model but with a greater amount per unit time (Figure 6G). As a result of excessively strong intrinsic spine dynamics and the decay of...
mean intra-group spine volume for the stimulated groups, the stimulated spines scattered around the spine volume distribution (Figure 6E). This result is in contrast to the WT result, where stimulated spines are localized nearer the tail of the spine volume distribution (cf. Figure 3E). Similarly to the WT model, the external stimulation at the onset of learning increased the spine generation rate by about 2.5% without significantly altering the elimination rate. Note that the baseline turnover in this fmr1KO model was about twice as high as the WT model, consistent with the experimental observation (Nagaoka et al., 2016).

The learning and memory deficits reproduced above are however potentially more severe than observed experimentally because fmr1KO mice have some learning capability, albeit limited. Therefore, we considered whether there may be some compensating mechanism, via which animals with excessively strong intrinsic spine dynamics rescue some learning and memory performance. We consider the regulation of potential connectivity as one candidate mechanism. Figure 7 explores how the stability of cell assemblies changes in our WT and fmr1KO models (i.e., the simulations of Figures 3, 6) if the peak potential connectivity (Figure 1B1) is systematically altered. A cell assembly in either the WT or fmr1KO model can exhibit one of the following three scenarios: its mean firing rate either (1) explodes (≥100 Hz) due to unstable learning, (2) fades (≤1 Hz), or (3) is stably maintained within a physiological range (>1 Hz and <100 Hz) during the maintenance period. The peak potential connectivity of 10.4% was already optimized for the WT model, such that nearly all cell assemblies are stable. The frequency of fading assemblies increased if potential connectivity was too low and that of exploding assemblies increased if potential connectivity was too high. However, the WT model stably maintained cell assemblies over a range of the peak potential connectivity from 9.4 to 10.8%. In contrast, the maintenance of cell assemblies in the fmr1KO model was much more sensitive to potential connectivity. As indicated in Figure 7, most of the assemblies fade with the baseline peak connectivity of 10.4%. The stability in the fmr1KO model first improved up to ~50% maintenance rate, but soon started to decline again due to exploding assemblies, as potential connectivity increases. Given this result, it is an interesting possibility that fmr1KO mice with excessively strong intrinsic spine dynamics may locally up-regulate potential connectivity, relative to WT mice, to avoid catastrophic forgetting (Figure 7 Right; marked as compensation). This hypothesis, which predicts an elevated frequency of exploding assemblies in fmr1KO mice, is consistent with the experimentally observed hyper neural activity and synchrony in fmr1KO mice (Musumeci et al., 2000; Gonçalves et al., 2013).

DISCUSSION

In previous models of synaptic plasticity, changes in synaptic strengths are assumed to be activity-dependent (Poo et al., 2016; Mongillo et al., 2017). Recent experimental observations suggest that this is not the case. In this work, we modeled how activity-independent intrinsic spine dynamics (Yasumatsu et al., 2008; Nagaoka et al., 2016) affect learning and memory in recurrent circuit models. For simplicity, we assumed that spine volume is proportional to synaptic strength (Matsuzaki et al., 2004; Harvey and Svoboda, 2007; Bosch et al., 2014). Contrary to a view that noisy synaptic changes are harmful to memory, intrinsic spine dynamics in our model play a positive role in preventing Hebbian-plasticity-driven non-specific growth of synapses. Specifically, as a result of the interaction between STDP and intrinsic spine dynamics in our model, the mean intra-assembly spine volume exhibits bistability, which is suitable to sustain memory. STDP keeps spines large within a cell assembly due to the coherent spontaneous activity of the membership neurons (e.g., Kenet et al., 2003; Diekelmann and Born, 2010; Wei and Koukalov, 2014), and intrinsic spine dynamics constitutively normalize all spines toward a physiological distribution.

In our model, memory is maintained by the total strength of synapses that connect neurons within an assembly. Individual synapses fluctuate and exhibit constant turnover, but this does not degrade the memory as long as the net strength is kept. One property of the current model is that, given the innate Hebbian instability of cell assemblies in spontaneously active recurrent networks, highly overlapping cell assemblies likely merge during maintenance, even in the presence of intrinsic spine dynamics. In this sense, either a preprocessing mechanism that decorrelates memory patterns (Perez-Orive et al., 2002; Leutgeb et al., 2007) or a more elaborate additional mechanism that stabilizes individual synapses (Frey and Morris, 1997; Ziegler et al., 2015; Benna and Fusi, 2016) may be helpful for improving the memory capacity (Hopfield, 1982).

There are several experimental observations that support the role of intrinsic spine dynamics in shaping the spine volume distribution. First, the spine volume distribution in the absence of neural activity and plasticity was well predicted by intrinsic spine dynamics, and the distribution is similar to that in the normal condition (Yasumatsu et al., 2008). Second, the spine volume distribution was robust to experimental manipulations to synaptic plasticity. Remarkably, calcineurin KO mice with little LTD (Okazaki et al., 2018) exhibited a spine volume distribution that was similar to WT mice. This raises an argument against the hypothesis that the spine volume distribution is set by the balance between LTP and LTD. We argue that previous computational models that do not include intrinsic spine dynamics miss an important component underlying synaptic organization.

We studied the interplay between activity-dependent synaptic plasticity and intrinsic spine dynamics in the formation and maintenance of cell assemblies in recurrent networks. Conventional studies of intrinsic spine dynamics often focus on independent synapses (Yasumatsu et al., 2008) or learning in a local feedforward network (Matsubara and Uehara, 2016) instead of learning in a recurrently connected network. Another study explored the consequence of volatile synaptic strengths in recurrently connected networks without studying how such volatility affects activity-dependent plasticity (Mongillo et al., 2018). Other studies model stochastic changes of synaptic strength (Loewenstein et al., 2011) or connectivity (Deger et al., 2012; Fauth et al., 2015) without distinguishing activity-dependent and -independent parts. Hence, these works do
FIGURE 6 | Network behavior in the presence of predicted fmr1KO intrinsic spine dynamics. (A–H) Conventions are as in Figure 3.
not distinguish their separate roles in memory and synaptic normalization. Another line of studies (Litwin-Kumar and Doiron, 2014; Zenke et al., 2015) model activity-dependent synaptic plasticity to account for the formation and maintenance of cell assemblies in a recurrent network. However, because these models do not include intrinsic spine dynamics, their synaptic strength distributions are typically sensitive to the fine balance between LTP and LTD and counter to the experimental observations described above. Thus, the proposed model postulates how synaptic normalization, by intrinsic spine dynamics, maintains most synapses that are not participating in a cell assembly weak. In addition, intrinsic spine dynamics in our model work as a homeostatic mechanism that stabilizes Hebbian plasticity, although they are activity independent and an explicit sensing-and-control (Davis, 2006; Shah and Crair, 2008) mechanism is absent. For this to work, the relative magnitude of Hebbian plasticity and intrinsic spine dynamics is important; For example, this stabilization fails if Hebbian plasticity is too fast (see Figure 4). While we conjecture that intrinsic spine dynamics can stabilize slow Hebbian plasticity in adults, other fast forms of homeostatic plasticity, such as inhibitory plasticity (Vogels et al., 2011; Litwin-Kumar and Doiron, 2014), might be helpful in the young, or once neural activity deviates beyond the level that intrinsic spine dynamics can compensate for.

Another main contribution of this study is the relationship between intrinsic spine dynamics and ASD. The observation that baseline spine turnover is abnormally high in various ASD mouse models (Ishiki et al., 2014), including a model animal for fragile X syndrome (fmr1KO) (Pan et al., 2010) suggests abnormal intrinsic spine dynamics could be a candidate substrate of ASD. Indeed, a recent experiment has confirmed that high baseline spine turnover in fmr1KO mice is activity-independent (Nagaoka et al., 2016), and intrinsic spine dynamics are stronger in fmr1KO (Ishii et al., 2018). Based on these experimental observations, we fitted parameters characterizing intrinsic spine dynamics in fmr1KO mice and found that they explain the learning deficit observed in fmr1KO mice (Padmashri et al., 2013). Interestingly, when we included a compensatory increase in recurrent excitatory connectivity to rescue some memory, the model reproduced epileptic-like neural activity that has been reported experimentally in fmr1KO mice (Musumeci et al., 2000; Gonçalves et al., 2013). More generally, the disrupted cortical connectivity theory of ASD (Courchesne and Pierce, 2005; Kana et al., 2011) argues that deficiency of cortical long-range connections and compensatory strengthening of local connections is a general feature of ASD. We contend that because there are typically fewer long-range connections, which therefore limits the positive feedback effect of Hebbian plasticity that is required for maintaining cell assemblies, they are especially susceptible to degradation due to excessively strong intrinsic spine dynamics, such as in our fmr1KO model. Furthermore, the learning deficiency in our proposed local cortical circuit model of fmr1KO, and its rescue by a compensatory increase in the local connectivity (Figure 7), is consistent with this theory. Hence, it is an intriguing possibility that pharmacological manipulations (Nagaoka et al., 2016), or a future neurofeedback technology (Ganguly and Poo, 2013; Yahata et al., 2016), could be used to rescue memory and learning, and long-range neural association, by reducing intrinsic spine dynamics or producing network motifs (Watanabe and Rees, 2015) efficiently connecting target brain areas.

A similar concept to intrinsic spine dynamics that has been used to describe ASD is intrinsic forgetting (Davis and Zhong, 2017). These two mechanisms are both related to chronic molecular signaling, which slowly degrades synapses and therefore memories. However, important differences between the two are how they are regulated in ASD animals and how they could possibly be involved in producing flexible behavior. A decrease in intrinsic forgetting has been argued to disable flexible behavior by maintaining conflicting memories (Reaume et al., 2011; Davis and Zhong, 2017). In contrast, the excessively strong intrinsic spine dynamics found in ASD animals (Pan et al., 2010; Ishiki et al., 2014; Nagaoka et al., 2016) would work in the opposite manner, by facilitating forgetting. Namely, stable cell assemblies in our fmr1KO model, typically require more neurons than the WT model to counter excess synaptic normalization. Hence, given a hypothesis that the number of total neurons is roughly the same between fmr1KO and WT mice, fmr1KO mice

![Figure 7](image-url)
may afford a smaller number of cell assemblies in the brain, possibly reducing the number of behavioral repertoires.

Overall, our spine dynamics model provides a novel path relating spine statistics, memory, and ASD. Notably, it is currently difficult to block intrinsic spine dynamics experimentally in vivo without affecting plasticity, because they are thought to be caused by the thermal agitation of molecules. This underscores the importance of a modeling study. Selective manipulations of intrinsic spine dynamics are an intriguing candidate direction to influence memory and learning, and the current model serves as a guide.

METHODS

Network and Neurons

A local network of \( N_E \) excitatory and \( N_I \) inhibitory leaky-integrate-and-fire neurons [Figure 1A; see e.g., (Tuckwell, 1988)] is considered. We simulate a network with \((N_E,N_I) = (1000,200)\). The dynamics of membrane potential \( V_i \) of neuron \( i \) is described by

\[
\tau_m \frac{dV_i}{dt}(t) = - (V_i(t) - V_0) - A_i(t) + R_i(t) \sum_{jk} w_{ij}^{(k)} \int_{-\infty}^{t} f(t - t') S_j(t' - \Delta_{ij}) dt' + R_i(t) I_i^{\text{ext}}(t)
\]

with membrane time constant \( \tau_m = 20 \text{ ms} \), resting potential \( V_0 = -70 \text{ mV} \), adaptation \( A_i \), refractory coefficient \( R_i \), kth \((k = 1,2,...)\), synaptic strength \( w_{ij}^{(k)} \) from neuron \( j \) to neuron \( i \), spike train \( S_j(t) = \sum_n \delta(t - t_{ij}^{(n)}) \) of neuron \( j \) composed of \( n \) \((n = 1,2,...)\) spike time \( t_{ij}^{(n)} \), random axonal delay \( \Delta_{ij} \) from neuron \( j \) to neuron \( i \) uniformly sampled from interval \([0.5, 5.0]\) ms, and external input \( I_i^{\text{ext}} \) to neuron \( i \) (see below). Note that \( \delta \) is the Dirac delta function. The time course of postsynaptic input is modeled using the alpha function (Gerstner et al., 2014),

\[
f(t) = 20 \text{ mV} \cdot \frac{\tau_r}{\tau_r - \tau_f} \left( e^{-t/\tau_f} - e^{-t/\tau_r} \right) \Theta(t)
\]

with rise time \( \tau_r = 0.5 \text{ ms} \), fall time \( \tau_f = 2.0 \text{ ms} \), and the Heaviside step function \( \Theta(t) \) that takes 1 for \( t \geq 0 \) and 0 for \( t < 0 \). The peak value of \( f(t) \) is about 0.39 mV. A spike is emitted when \( V_i \) reaches a spiking threshold at \(-50 \text{ mV} \). A spike is reset to \( V_0 \). Excitatory neurons receive adaptation input \( A_i \) that reflects a slow \( \text{Na}^+ \)-activated \( \text{K}^+ \) current (Wang et al., 2003) (Figure 1B4), with dynamics described by

\[
dA_i(t) = - \frac{A_i(t)}{\tau_A} + 0.0017[20 \text{ mV} - A_i(t)] S_i(t)
\]

with time constant \( \tau_A = 13 \text{ s} \). In contrast, we assume no adaptation (\( A_i = 0 \)) in inhibitory neurons. Refractoriness is imposed by \( R_i \). \( R_i \) is fixed at 0 after each spike of neuron \( i \) for 1 ms (absolute refractoriness), and then recovers toward 1 following

\[
\tau_R \frac{dR_i}{dt}(t) = 1 - R_i(t)
\]

with time constant \( \tau_R = 3.5 \text{ ms} \) (relative refractoriness). The differential equations are numerically solved using the Euler method with bin size \( \Delta t = 0.1 \text{ ms} \).

Network Topology

In the network model of Figure 1A, excitatory neurons, \( E \), and inhibitory neurons, \( I \), are sparsely connected. The connections from excitatory to inhibitory neurons (\( E \rightarrow I \)) and those from inhibitory to excitatory neurons (\( I \rightarrow E \)) are randomly generated with 10% connection probability. Each of these directed connections is mediated by a single synapse, whose strength is randomly drawn from a uniform distribution in the range \([0, 31] \text{ mV} \) for \( E \rightarrow I \) and \([-31, 0] \text{ mV} \) for \( I \rightarrow E \), respectively. According to our neuron model described above, a typical \( E \rightarrow I \) or \( I \rightarrow E \) synapse produces a postsynaptic potential of \( 6 \text{ mV} \). For simplicity, we assume no direct connections between inhibitory neurons. Excitatory neurons are equidistantly placed on a one-dimensional ring of circumference 1 a.u. that represents the feature space. Note that the tuning distance in feature space \( d_{ij} \) does not necessarily correspond to the physical location of a neuron, or the distance between any two neurons. The potential connections from excitatory to excitatory neurons (\( E \rightarrow E \)) are randomly generated according to the Gaussian probability profile (0.14% \( \exp[-0.5(d_{ij}/0.1)^2] \)).

Figure 1B1), which peaks at 10.4% for similarly tuned neurons \((d_{ij} \approx 0)\), and decays toward 0 with a length-constant of 0.1 as \( d_{ij} \) increases. Although this \( E \rightarrow E \) peak connection probability is optimized for memory retention, simulation outcomes are robust with this parameter in our model (see, Figure 7, WT model). Effectively, this allows excitatory neurons closer in the feature space to be more interconnected, and those farther apart to be less so. If there is a potential connection from one neuron to another, we randomly assign a fixed number \( K(K = 1, 2, \cdots, 10) \) of spines from a Poisson distribution \((\lambda K/K')/ \sum_{K=1}^{10} (\lambda K'/K') \) with parameter \( \lambda = 3 \) (Figure 1B2).

Stimulation

In addition to the recurrent input, neuron \( i \) also receives spike train \( S_{i^{\text{ext}}}(t') \) as external input (dashed connection in Figure 1A). At baseline, \( S_{i^{\text{ext}}}(t') \) is generated by a Poisson process with firing intensity 60 Hz. Input neurons are not modeled explicitly here. The external input to each neuron is given by \( I_{i^{\text{ext}}}(t) = \int_{-\infty}^{t} f(t - t') S_{i^{\text{ext}}}(t') dt' \). This sets the baseline membrane potential and firing rate of neurons to \(-58.6 \pm 2.4 \text{ mV} \) and 0.13 \pm 0.08 Hz, respectively.

We divide the feature space of excitatory neurons into four equally-sized consecutive parts and randomly select 40% of the neurons from each part as a stimulated neuron group (Figure 1A). During the learning period (indicated by blue horizontal bars below the time axes in Figures 2, 3, 6), one of the 4 stimulated neuron groups is randomly selected with probability \( \frac{1}{4} \) and receives additional Poisson spikes at 750 Hz for 3 s. During the learning period, all inhibitory neurons also receive additional Poisson spikes at 300 Hz. The learning period starts at \( t = 0 \) and ends when at least one group’s mean intra-group spine volume reaches \( \geq 0.49 \mu \text{m}^3 \).
Spine Dynamics

Unlike E → I and I → E synapses, which have fixed weights, E → E synapses (orange and brown dashed in Figure 1A) change in time via two independent mechanisms: STDP and intrinsic spine dynamics. We assume that synaptic strengths for E → E synapses are essentially proportional to their spine volumes and therefore model their spine volume dynamics as changes in synaptic strengths. Changes in the kth \( k = 1, 2, ..., K_{ij} \) spine volume \( v_{ij}^{(k)} \) on neuron \( i \), receiving signal from neuron \( j \), is modeled by

\[
\frac{dv_{ij}^{(k)}(t)}{dt} = T_{\text{STDP}} \left( S_i(t)\overline{S}_j(t) - \frac{v_{ij}^{(k)}(t)}{\tau_{\text{LTD}}} S_j(t)\overline{S}_j(t) \right) \Theta(v_{ij}^{(k)}(t) - v_0) + \sqrt{T}(a\overline{v}_{ij}^{(k)}(t) + \beta)\xi(t)
\]

where the first and second terms on the right-hand side describe changes by multiplicative STDP (van Rossum et al., 2000) and intrinsic spine dynamics (Yasumatsu et al., 2008), respectively. \( T_{\text{STDP}} \) is a speed-up factor that we describe below, \( a = 7.6 \times 10^{-9} \text{m}^3 \) is the amplitude of STDP, \( \overline{S}_i \) is the running average of past spiking activity of neuron \( i \), i.e.,

\[
\frac{d\overline{S}_i(t)}{dt} = -\frac{\overline{S}_i(t)}{\tau_{\text{STDP}}} + S_i(t)
\]

with an averaging time constant \( \tau_{\text{STDP}} = 20 \text{ ms} \) and \( v_{\text{LTD}} = 0.5 \text{ m}^3 \) is the scaling factor for volume-dependent LTD (van Rossum et al., 2000). STDP is assumed to be absent for small spines of \( v_{ij}^{(k)} < v_0 = 0.02 \text{ m}^3 \). Slope parameter \( a = 0.2 \text{ day}^{-\frac{1}{2}} \) and offset parameter \( \beta = 0.01 \text{ m}^3 \text{ day}^{-\frac{1}{2}} \) for intrinsic spine dynamics are set as previously experimentally observed (Yasumatsu et al., 2008). \( \xi(t) \) is white noise with the autocorrelation function \( \langle \xi(t)\xi(t') \rangle = \delta(t - t') \). The above Langevin equation is numerically solved by the Euler method with bin size \( \Delta t = 0.1 \text{ ms} \). In addition, we set reflecting boundaries for spine volume to enforce \( 0 \leq v_{ij}^{(k)} \leq 1.0 \text{ m}^3 \) for all spines. One problem is that it is too time consuming to directly simulate the 10-day learning period studied with the fine time resolution required to simulate STDP and intrinsic spine dynamics. We therefore run \( T = 3.3 \times 10^4 \text{ times shorter simulations by speeding up both STDP and intrinsic spine dynamics by factors } T \) and \( \sqrt{T} \), respectively. (Note that volume changes \( \nu(t + \Delta t) - \nu(t) \) by intrinsic spine dynamics are diffusive and scale with the square root of time duration \( \sqrt{T}\Delta t \).) This way, we can extrapolate spine volume changes happening during 10 days based on shorter simulations up to 3,000 s. We display the time before this conversion in panels describing neural activity in seconds, but display the time after this conversion in panels describing learning in days. We initially set \( E \rightarrow E \) spine volumes by randomly sampling from the equilibrium distribution \( P_{eq}(v) \propto (av + \beta)^{-2} \), which is set by the intrinsic spine dynamics. Synaptic strength \( w \) of a spine with volume \( \nu \) is then assumed to be

\[
w = (43 \text{ m}^3)\nu
\]

for \( \nu \geq v_0 \) (a functional spine) and 0 for \( \nu < v_0 \) (a non-functional spine, e.g., filopodia). The median spine volume of \( P_{eq}(\nu) \) is 0.047 μm³ and such a spine produces 0.8 mV of excitatory postsynaptic potential. We set \( v_0 \) to be a threshold volume typically used in experiments to detect spines (Yasumatsu et al., 2008). We define spine gain and loss by a fraction of spines passing this threshold from below and above per day, respectively. The exact value of \( v_0 \) does not matter for our results as long as it is sufficiently small.

Itô vs. Stratonovich Interpretation

The meaning of fluctuation is different under the Itô and Stratonovich interpretations of intrinsic spine dynamics (Gardiner, 1985). The intrinsic spine dynamics under the Itô interpretation that we study in the main text are described by

\[
\frac{\partial P(v, t)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial v^2} (av + \beta)^2 P(v, t)
\]

In this view, changes in spine volume distribution are purely produced by the volume-dependent fluctuation with amplitude \( av + \beta \). In contrast, the Stratonovich interpretation represents the same equation in a different way:

\[
\frac{\partial P(v, t)}{\partial t} = -\frac{\partial}{\partial v} \left[ \frac{a(\nu + \beta)}{2} P(v, t) \right] + \frac{1}{2} \frac{\partial}{\partial v} \left[ (v + \beta) \frac{\partial}{\partial \nu} [(av + \beta) P(v, t)] \right]
\]

In this view, changes in spine volume distribution are described by two terms: the first term is the drift term produced by apparent force \( a(\nu + \beta)/2 \) and the second term is produced by the volume-dependent fluctuation \( av + \beta \). Hence, while the above two equations are identical, there is a semantic difference regarding what fluctuation means. According to the Itô interpretation, only the fluctuation drives spine volume changes and this fluctuation preserves mean spine volume [i.e., martingale (Øksendal, 2000)] except for a boundary effect. According to the Stratonovich interpretation, the apparent force shrinks and the fluctuation enlarges mean spine volume, respectively, and the two effects are canceled. These two interpretations become the same in the special case of \( \alpha = 0 \), namely, when the fluctuation is volume independent.

Simulation Environment

Simulations were performed in custom written C code with the forward Euler-integration method and a step size of 0.1 ms. Code was written to efficiently run on multiple GPUs to produce the trillions of random numbers needed to model the intrinsic spine dynamics per simulation up to 3,000 s. Post-simulation analysis was undertaken with MATLAB. Source code is available upon request.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.
AUTHOR CONTRIBUTIONS

HK and TT conceived the project. JH and KH performed numerical simulations. JH, KH, HK, and TT analyzed the results and drafted and revised the manuscript.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number JP18H05432 (TT), Brain/MINDS from AMED under Grant Number JP19dm0207001 (TT), and RIKEN Center for Brain Science (JH and TT). This manuscript has been released as a Pre-Print (Humble et al., 2019) at bioRxiv 525980.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2019.00038/full#supplementary-material

REFERENCES

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. Biol. Cybern. 27, 77–87. doi: 10.1007/BF00337259

Asrican, B., Lisman, J., and Otmakhov, N. (2007). Synaptic strength of individual spines correlates with bound Ca2+-calmodulin-dependent kinase II. J. Neurosci. 27, 14007–14011. doi: 10.1523/JNEUROSCI.3587-07.2007

Béqué, J.-C., Lin, D.-T., Kang, M.-G., Aizawa, H., Takamnya, K., and Huganir, R. L. (2006). Synapse-specific regulation of AMPA receptor function by PSD-95. Proc. Natl. Acad. Sci. U.S.A. 103, 19535–19540. doi: 10.1073/pnas.0608492103

Bennia, M. K., and Fusi, S. (2016). Computational principles of synaptic memory consolidation. Nat. Neurosci. 19, 1697–1706. doi: 10.1038/nrn.4401

Bi, G. Q., and Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. J. Neurosci. 18, 10464–10472. doi: 10.1523/JNEUROSCI.18-24-10464.1998

Billings, G., and van Rossum, M. C. (2009). Memory retention and spike-timing-dependent plasticity. J. Neurophysiol. 101, 2775–2788. doi: 10.1152/jn.90107.2008

Bosch, M., Castro, J., Saneyoshi, T., Matsuno, H., Sur, M., and Hayashi, Y. (2014). Structural and molecular remodeling of dendritic spine substructures during long-term potentiation. Neuron 82, 444–459. doi: 10.1016/j.neuron.2014.03.021

Cossell, L., Iacaruso, M. F., Muir, D. R., Houlton, B., Sader, E. N., Ko, H., et al. (2011). Functional organization of excitatory synaptic strength in primary visual cortex. Nature 518, 399–403. doi: 10.1038/nature14182

Courchesne, E., and Pierce, K. (2005). Why the frontal cortex in autism might be talking only to itself: local over-connectivity but long-distance disconnection. Curr. Opin. Neurobiol. 15, 225–230. doi: 10.1016/j.conb.2005.03.001

Davis, W. G. (2006). Homeostatic control of neural activity: from phenomenology to molecular design. Annu. Rev. Neurosci. 29, 307–323. doi: 10.1146/annurev.neuro.28.061604.135751

Davis, R. L., and Zhong, Y. (2017). The biology of forgetting: A perspective. Neuron 95, 490–503. doi: 10.1016/j.neuron.2017.05.039

Deger, M., Helias, M., Rottier, S., and Diesmann, M. (2012). Spike-timing dependence of structural plasticity explains cooperative synapse formation in the neocortex. PLoS Comput. Biol. 8:e1002689. doi: 10.1371/journal.pcbi.1002689

Diedenhofen, B., and Born, J. (2010). The memory function of sleep. Nat. Rev. Neurosci. 11, 114–126. doi: 10.1038/nrn2762

Fauth, M., Wörgötter, F., and Tetzlaff, C. (2015). Formation and maintenance of robust long-term information storage in the presence of synaptic turnover. PLoS Comput. Biol. 11:e1004684. doi: 10.1371/journal.pcbi.1004684

Fiete, I. R., Senn, W., Wang, C. Z., and Hahnloser, R. H. (2010). Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. Neuron 65, 563–576. doi: 10.1016/j.neuron.2010.02.003

Frey, U., and Morris, R. G. (1997). Synaptic tagging and long-term potentiation. Nature 385, 533–536. doi: 10.1038/385533a0

Ganguly, K., and Poo, M.-M. (2013). Activity-dependent plasticity from bench to bedside. Nature 80, 729–741. doi: 10.1038/nn.3304

Gerstner, W., Kempter, R., van Hemmen, J. L., and Wagner, H. (1996). A neuronal learning rule for sub-millisecond temporal coding. Nature 383, 76–81. doi: 10.1038/38376d0

Gerstner, W., Kistler, W. M., Naud, R., and Paninski, L. (2014). Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781107476155

Gonçalves, J. T., Anstey, J. E., Golshani, P., and Portera-Cailliau, C. (2013). Circuit level defects in the developing neocortex of Fragile X mice. Nat. Neurosci. 16, 903–909. doi: 10.1038/nn.3415

Gütig, R., Aharony, R., Rottier, S., and Sompolinsky, H. (2003). Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. J. Neurosci. 23, 3697–3714. doi: 10.1523/JNEUROSCI.23-09-03697.2003

Hardingham, N. R., Read, J. C., Trevelyan, A. J., Nelson, I. C., Jack, J. J., and Bannister, N. J. (2010). Quantal analysis reveals a functional correlation between presynaptic and postsynaptic efficacy in excitatory connections from rat neocortex. J. Neurosci. 30, 1441–1451. doi: 10.1523/JNEUROSCI.3244-09.2010

Harvey, C. D., and Svoboda, K. (2007). Locally dynamic synaptic learning rules in pyramidal neuron dendrites. Nature 450, 1195–1200. doi: 10.1038/nature06416

Hayama, T., Noguchi, J., Watanabe, S., Takahashi, N., Hayashi-Takagi, A., Ellis-Davies, G. C., et al. (2013). GABA promotes the competitive selection of dendritic spines by controlling local Ca2+ signaling. Nat. Neurosci. 16, 1409–1416. doi: 10.1038/nn.3496

He, C. X., and Portera-Cailliau, C. (2013). The trouble with spikes in fragile X syndrome: density, maturity and plasticity. Neuroscience 251, 120–128. doi: 10.1016/j.neuroscience.2012.03.049

Hebb, D. O. (1949). The Organization of Behavior: A Neuropsychological Theory. Mahwah, NJ: L. Erlbaum Associates.

Hofer, S. B., Mirlis-Fogl, T. D., Bonhoeffer, T., and Hübener, M. (2009). Experience leaves a lasting structural trace in cortical circuits. Nature 457, 313–317. doi: 10.1038/nature07487

Holbro, N., Grundtiz, A., and Oertner, T. G. (2009). Differential distribution of endoplasmic reticulum controls metabotropic signaling and plasticity at hippocampal synapses. Proc. Natl. Acad. Sci. U.S.A. 106, 15055–15060. doi: 10.1073/pnas.0905110106

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. U.S.A. 79, 2554–2558. doi: 10.1073/pnas.79.8.2554

Humble, J., Hiratsuka, K., Kasai, H., and Toyoizumi, T. (2019). Intrinsic spine dynamics are critical for recurrent network learning in models with and without autism spectrum disorder. bioRxiv 525980. doi: 10.1101/525980

Isshiki, M., Tanaka, S., Kuriu, T., Tabuchi, K., Takumi, T., and Okabe, S. (2014). Enhanced synapse remodelling as a common phenotype in mouse models of autism. Nat. Commun. 5:4742. doi: 10.1038/ncomms5742

Kana, R. K., Libero, L. E., and Moore, M. S. (2011). Disrupted cortical connectivity theory as an explanatory model for autism spectrum disorders. Phys. Life Rev. 8, 410–437. doi: 10.1016/j.plrev.2011.10.001
Keck, T., Toyozumi, T., Chen, L., Doiron, B., Feldman, D. E., Fox, K., et al. (2017). Integrating Hebbian and homeostatic plasticity: the current state of the field and future research directions. Philos. Trans. R. Soc. Lond. Ser B Biol. Sci. 372. doi: 10.1098/rstb.2016.0158

Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., and Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. Nature 425, 954–956. doi: 10.1038/nature02078

Kim, S. K., and Nabekura, J. (2011). Rapid synaptic remodeling in the adult somatosensory cortex following peripheral nerve injury and its association with neuropathic pain. J. Neurosci. 31, 5477–5482. doi: 10.1523/JNEUROSCI.0308-11.2011

Kopec, C. D., Li, B., Wei, W., Boehm, J., and Malinow, R. (2006). Glutamate receptor exocytosis and spine enlargement during chemically induced long-term potentiation. J. Neurosci. 26, 2000–2009. doi: 10.1523/JNEUROSCI.3918-05.2006

Lang, C., Barco, A., Zablow, L., Kandel, E. R., Siegelbaum, S. A., and Zakharenko, S. S. (2004). Transient expansion of synaptic connections. Brain Res. 1015, 206–221. doi: 10.1016/j.brainres.2004.06.022

Leutgeb, J. K., Leutgeb, S., Moser, M.-B., and Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. Science 315, 961–966. doi: 10.1126/science.1135801

Litwin-Kumar, A., and Doiron, B. (2014). Formation and maintenance of neuronal assemblies through synaptic plasticity. Nat. Commun. 5:3199. doi: 10.1038/ncomms4319

Liu, X., Ramirez, S., Pang, P. T., Purveyar, C. B., Govindarajan, A., Deisseroth, K., et al. (2012). Optogenetic stimulation of a hippocampal amigra engages fear memory recall. Nature 484, 381–385. doi: 10.1038/nature11028

Malenka, R. C., and Bear, M. F. (2004). LTP and LTD: an embarrassment of riches. Ann. Rev. Neurosci. 27, 103–126. doi: 10.1146/annurev.neuro.25.110101.142513

Nagaoa, K., Takehara, H., Hayashi-Takagi, A., Noguchi, J., Ishii, K., Shirai, F., et al. (2016). Abnormal intrinsic dynamics of dendritic spines in a fragile X syndrome mouse model in vivo. Science 353, 626–631. doi: 10.1126/science.aaf7151

Neves, G., Cooke, S. F., and Bliss, T. V. P. (2008). Synaptic plasticity, memory and the hippocampus: a neural network approach to causality. Nat. Rev. Neurosci. 9, 65–75. doi: 10.1038/nrn2303

Pathania, M., Davenport, E. C., Muir, J., Sheehan, D. F., López-Doménech, G., and Kittle, J. T. (2014). The autism and schizophrenia associated gene CYFIP1 is critical for the maintenance of dendritic complexity and the stabilization of mature spines. Transl. Psychiatry 4:e374. doi: 10.1038/tp.2014.16

Pan, F., Aldridge, G. M., Greenough, W. T., and Gan, W.-B. (2010). Dendritic spine instability and insensitivity to modulation by experience in a mouse model of fragile X syndrome. J. Neurosci. 30, 19715–19723. doi: 10.1523/JNEUROSCI.2354-13.2013

Poo, M.-M., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T., Martin, K. R., et al. (2000). The state of synapses in fragile X syndrome. J. Physiol. 522, 1–21. doi: 10.1113/jphysiol.2000.017978

Zhang, X., Smith, A. M., and Kandel, E. R. (2001). Mechanisms of response homeostasis during retinocollicular map formation. J. Physiol. 535.1, 1–16. doi: 10.1111/j.1469-7793.2001.02211.x

Samuel, L. W., and Sams, M. (1998). Stochastic Differential Equations: An Introduction With Applications, 5th Edn, 4th print. Berlin: Springer.
memory and weight distribution. *Neural Comput.* 19, 639–671. doi: 10.1162/neco.2007.19.3.639

Tuckwell, H. C. (1988). *Introduction to Theoretical Neurobiology.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511623202

van Rossum, M. C., Bi, G. Q., and Turrigiano, G. G. (2000). Stable Hebbian learning from spike timing-dependent plasticity. *J. Neurosci.* 20, 8812–8821. doi: 10.1523/JNEUROSCI.20-23-08812.2000

Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C., and Gerstner, W. (2011). Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science* 334, 1569–1573. doi: 10.1126/science.1211095

Wang, X.-J., Liu, Y., Sanchez-Vives, M. V., and McCormick, D. A. (2003). Adaptation and temporal decorrelation by single neurons in the primary visual cortex. *J. Neurophysiol.* 89, 3279–3293. doi: 10.1152/jn.00242.2003

Watanabe, T., and Rees, G. (2015). Age-associated changes in rich-club organisation in autistic and neurotypical human brains. *Sci. Rep.* 5:16152. doi: 10.1038/srep16152

Wei, Y., and Koulakov, A. A. (2014). Long-term memory stabilized by noise-induced rehearsal. *J. Neurosci.* 34, 15804–15815. doi: 10.1523/JNEUROSCI.3929-12.2014

Yahata, N., Morimoto, J., Hashimoto, R., Lisi, G., Shibata, K., Kawakubo, Y., et al. (2016). A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nat. Commun.* 7:11254. doi: 10.1038/ncomms11254

Yang, G., Pan, F., and Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature* 462, 920–924. doi: 10.1038/nature08577

Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J., and Kasai, H. (2008). Principles of long-term dynamics of dendritic spines. *J. Neurosci.* 28, 13592–13608. doi: 10.1523/JNEUROSCI.0603-08.2008

Zenke, F., Agnes, E. J., and Gerstner, W. (2015). Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nat. Commun.* 6:6922. doi: 10.1038/ncomms7922

Zenke, F., Hennequin, G., and Gerstner, W. (2013).Synaptic plasticity in neural networks needs homeostasis with a fast rate detector. *PLoS Comput. Biol.* 9:e1003330. doi: 10.1371/journal.pcbi.1003330

Zhou, Q., Homma, K. J., and Poo, M. (2004). Shrinkage of dendritic spines associated with long-term depression of hippocampal synapses. *Neuron* 44, 749–757. doi: 10.1016/j.neuron.2004.11.011

Ziegler, L., Zenke, F., Kastner, D. B., and Gerstner, W. (2015). Synaptic consolidation: from synapses to behavioral modeling. *J. Neurosci.* 35, 1319–1334. doi: 10.1523/JNEUROSCI.3989-14.2015

Zito, K., Scheuss, V., Knott, G., Hill, T., and Svoboda, K. (2009). Rapid functional maturation of nascent dendritic spines. *Neuron* 61, 247–258. doi: 10.1016/j.neuron.2008.10.054

Ziv, N. E., and Brenner, N. (2018). Synaptic tenacity or lack thereof: spontaneous remodeling of synapses. *Trends Neurosci.* 41, 89–99. doi: 10.1016/j.tins.2017.12.003

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Humble, Hiratsuka, Kasai and Toyoizumi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.