Multiplicative Shot-Noise: A New Route to Stability of Plastic Networks

Bin Wang and Johnatan Aljadeff

Phys. Rev. Lett. **129**, 068101 — Published 2 August 2022

DOI: 10.1103/PhysRevLett.129.068101

Multiplicative shot-noise: a new route to stability of plastic networks

Bin Wang
Department of Physics, University of California San Diego

Johnatan Aljadeff
Department of Neurobiology, University of California San Diego
(Dated: June 29, 2022)

Fluctuations of synaptic-weights, among many other physical, biological and ecological quantities, are driven by coincident events of two 'parent' processes. We propose a multiplicative shot-noise model that can capture the behaviors of a broad range of such natural phenomena, and analytically derive an approximation that accurately predicts its statistics. We apply our results to study the effects of a multiplicative synaptic plasticity rule that was recently extracted from measurements in physiological conditions. Using mean-field theory analysis and network simulations we investigate how this rule shapes the connectivity and dynamics of recurrent spiking neural networks. The multiplicative plasticity rule is shown to support efficient learning of input stimuli, and gives a stable, unimodal synaptic-weight distribution with a large fraction of strong synapses. The strong synapses remain stable over long times but do not 'run away'. Our results suggest that the multiplicative shot-noise offers a new route to understand the tradeoff between flexibility and stability in neural circuits and other dynamic networks.

Usage:

Structure:

Introduction.—Many natural processes are triggered by coincidences of two 'parent' events. Examples include firefly flash synchronization [1]; effects of simultaneous environmental stressors [2]; applications of two-photon microscopy [3]; and stimulus-reward associations in reinforcement learning [4, 5]. In neuroscience, co-activation of pre- and postsynaptic neurons plays a crucial role in inducing synaptic plasticity [6, 7], a primary mechanism underlying learning and memory.

The parent processes are often described by event-based models [8], among which the Poisson process is an appealing starting point owing to its memory-less property. Experimental studies show that the aforementioned coincidence-based phenomena often cannot be accurately described as sums over the parent shot-noise (Poisson) processes [2, 7]. Specifically, induction of long-term plasticity was shown to depend strongly on the calcium flux into the postsynaptic neuron [9, 10]. This flux, in turn, depends on coincident spiking activity of pre- and postsynaptic neurons, and is well described by the product of two shot-noise processes [7, 11, 12]. In contrast, most network-level studies of spike-timing dependent plasticity (STDP) typically assume that the synaptic strength change is the sum over contributions of spike-pairs, ignoring cooperative effects between spikes [13–17]. These models often cannot reproduce realistic spiking activities observed *in vivo* [18].

Motivated by the converging theoretical and experimental evidence, we propose a stochastic process whose fluctuations are triggered by multiplicative interactions between two parent shot-noise processes $[c_\alpha(t),\, \alpha = 1, 2]$. The rates of parent events are λ_α, and their amplitudes $a_{\alpha,i}$ are exponentially-distributed with mean A_α. Events are referred to as "spikes", adopting the neuroscience terminology, but they may correspond to events in other domains. Spike-times are denoted $\{t_{\alpha,i}\}$ and may be temporally correlated. We then have,

$$\frac{\mathrm{d}c_\alpha(t)}{\mathrm{d}t} = -\frac{c_\alpha(t)}{\tau_\alpha} + \sum_i a_{\alpha,i}\delta(t - t_{\alpha,i}). \tag{1}$$

The decay timescales τ_α define a window during which coincidences can occur.

Our primary interest is a *multiplicative shot-noise* process (denoted $\mathcal{C}(t)$ and henceforth referred to as the *coincidence detector*), whose transient deviations from baseline are driven by the product $c_1 \times c_2$, with decay timescale $\tau_\mathcal{C}$ (Fig. 1A),

$$\frac{\mathrm{d}\mathcal{C}(t)}{\mathrm{d}t} = -\frac{\mathcal{C}(t)}{\tau_\mathcal{C}} + \eta c_1(t)c_2(t). \tag{2}$$

The stochastic calculus of Poisson processes makes it difficult to analyze their products [19]. In contrast, existing nonlinear shot-noise models [20–23] are equivalent to transformations of a single Poisson process, and are not suitable for studying statistics of coincidences.

We analyze the statistics of the coincidence detector and apply these results to gain insights to a longstanding problem in neuroscience: the stability of recurrent neuronal networks subject to STDP [24]. Here, c_1, c_2 are calcium transients induced by pre- and postsynaptic spikes; and $\mathcal{C}$ is the total calcium flux, which triggers plasticity [7]. Based on spiking network simulations and theoretical analysis of a reduced model, we show that when *individual synapses* in a recurrent network are subject to a nonlinear calcium-based plasticity rule, the empirical *macroscopic* network properties are reproduced (e.g.,

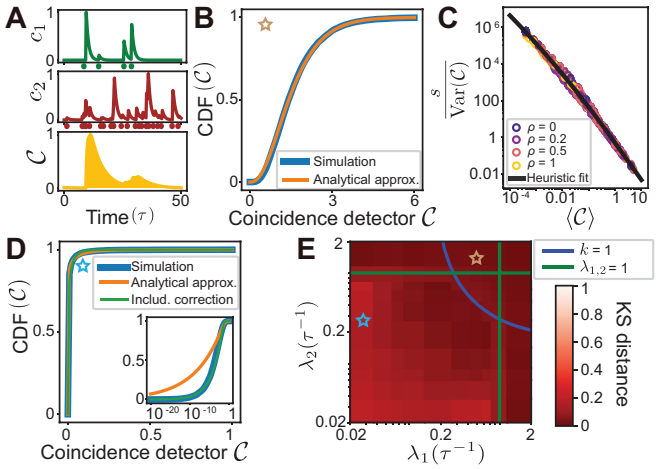

FIG. 1. Analytical approximation of multiplicative shot-noise statistics. (A) An illustration of the process $\mathcal{C}$, driven by the product $c_1 \times c_2$. (B) The cumulative distribution function (CDF) of $\mathcal{C}$ calculated through method of moments (MM) [Eq. (4)] matches simulations well at high firing-rates. (C) Relationship between the log-expectation, mean and variance of $\mathcal{C}$ [Eq. (5)]. Data collapses on a line for a range of firing-rates and spike-time correlations (ρ). (D) At low firing-rates, the CDF given by MM matches simulations well for $\mathcal{C} \gtrsim \eta\tau_{\mathcal{C}} A_1 A_2$ but poorly for $\mathcal{C} \to 0$. The heuristic MM matches the simulation over the entire range of $\mathcal{C}$. (E) Performance of heuristic MM for uncorrelated spike trains. The errors, measured by KS distance, are ≤ 0.2 for all firing-rates. See [27] for definition of ρ and KS distances for other parameters.

stable activity patterns, unimodal heavy-tailed synaptic-weight distributions [25]. Further, our results suggest that STDP in itself can support representations that remain stable over timescale of hours, making an important step towards understanding prolonged retention of spatial memories in the face of plasticity and noise [26].

Statistics of multiplicative shot-noise.—For simplicity, we assume that the decay timescales of the parent processes are identical, $\tau_1 = \tau_2 \equiv \tau$. This is consistent with calcium-induced plasticity [7, 28] and cases where the parent processes are generated by similar agents (e.g., firefly flashes [1]).

Solving Eq. (1) gives filtered spike-trains (Fig. 1A). Using these solutions, we evaluate $\mathcal{C}(t)$ [Eq. (2)] at steady-state. When the parent processes are uncorrelated,

$$\mathcal{C} \stackrel{d}{=} \eta\tau_{\mathcal{C}} \sum_{i,j} a_{1,i} a_{2,j} \times e^{-\frac{|t_{1,i}-t_{2,j}|}{\tau}} \times \frac{\mathcal{R}(\min(t_{1,i}, t_{2,j}))}{2\delta - 1}. \tag{3}$$

Here, $\delta = \frac{\tau_{\mathcal{C}}}{\tau}$ is the ratio between timescales [Eqs. (1, 2)], and $\stackrel{d}{=}$ means "equal in distribution". The natural interpretations of the three factors in Eq. (3) are the stochastic amplitudes of synaptic transmission; the temporal window for coincident spikes; and $\mathcal{R}(x) = e^{-\frac{x}{\tau_{\mathcal{C}}}} - e^{-\frac{2x}{\tau}}$ describes firing-rate dependent accumulation of multiple coincidences. Notably, $\mathcal{R}$ represents a departure from summation over spike-pairs [13–16]. Similar expressions for temporally correlated spike-trains appear in [27].

We begin by formulating an analytical approximation of $P_{\mathcal{C}}$, the distribution of $\mathcal{C}$. At steady-state, the shot-noise process [e.g., c_α in Eq. (1)] follows a Gamma distribution [29, 30] with shape and scale parameters $\lambda_\alpha\tau$, A_α. We think of the coincidences of the parent processes as events which drive fluctuations of $\mathcal{C}$. Therefore, we assume $P_{\mathcal{C}}$ can be approximated by a Gamma distribution whose shape (k) and scale (σ) parameters measure the effective rate and amplitude of the coincident spikes.

We used the method of moments (MM), i.e., matching the first and second moments of $P_{\mathcal{C}}$ to a Gamma distribution, to analytically estimate k, σ. We find for uncorrelated spike-trains,

$$k = \frac{\langle\mathcal{C}\rangle^2}{\mathrm{Var}(\mathcal{C})} = \frac{\lambda_1\lambda_2\tau^2}{\frac{\lambda_1+\lambda_2}{\delta+1}\tau + \frac{1}{2\delta+1}},$$
$$\sigma = \frac{\mathrm{Var}(\mathcal{C})}{\langle\mathcal{C}\rangle} = \eta\tau_{\mathcal{C}} A_1 A_2 \left(\frac{\lambda_1+\lambda_2}{\delta+1}\tau + \frac{1}{2\delta+1}\right). \tag{4}$$

See [27] for expressions including correlations.

We show that MM [Eq. (4)] is highly accurate in the high firing-rate regime ($\langle\mathcal{C}\rangle > \eta\tau_{\mathcal{C}} A_1 A_2$) by comparing it to numerical simulations (Fig. 1E, above blue line). However, for low firing-rates (i.e., $\lambda_1\tau, \lambda_2\tau < 1$ which implies $\langle\mathcal{C}\rangle < \eta\tau_{\mathcal{C}} A_1 A_2$) and particularly at $\mathcal{C} \to 0$, the MM is inaccurate (Fig. 1D). The reason is that for low firing-rates (and $k < 1$), the Gamma probability density diverges at 0. Such a singularity cannot be captured by the mean and variance of $P_{\mathcal{C}}$. Notably, either of λ_1, λ_2 being high suffices for the MM to be accurate, because the high-rate process provides a "background" on top of which the low-rate process can trigger coincidences.

In some applications of our theory, it may be important to accurately estimate $P_{\mathcal{C}}$ at $\mathcal{C} \to 0$, in the low firing-rate regime. We obtain such an estimate by first noticing that in this regime, the maximum-likelihood estimate of the Gamma distribution parameters (k, σ) yields a good approximation of $P_{\mathcal{C}}$. This estimate relies on the log-expectation variable $s = \ln\langle\mathcal{C}\rangle - \langle\ln\mathcal{C}\rangle$, which is indeed sensitive to the singularity at $\mathcal{C} \to 0$. Next we show that the log-expectation, the mean and variance of $\mathcal{C}$ obey a simple relationship, irrespective of the firing-rates and correlation (Fig. 1C),

$$\ln\left(\frac{s}{\mathrm{Var}(\mathcal{C})}\right) = -\ln 2 - 2\ln\langle\mathcal{C}\rangle + b(\delta)[-\ln\langle\mathcal{C}\rangle]_+^2, \tag{5}$$

where $[x]_+ = \max(x, 0)$ and $b(\delta)$ was fit directly to simulations. We found corrected shape and scale parameters $\tilde{k}$, $\tilde{\sigma}$ by computing s via Eq. (5) and using standard maximum-likelihood formulae [27]. We call this the "heuristic MM", and use it when coincidences are rare $k < 1$, and both firing-rates are small $\max(\lambda_1\tau, \lambda_2\tau) < 1$. This approximation is accurate for all values of $\mathcal{C}$, in the entire parameter space we explored (Fig. 1E).

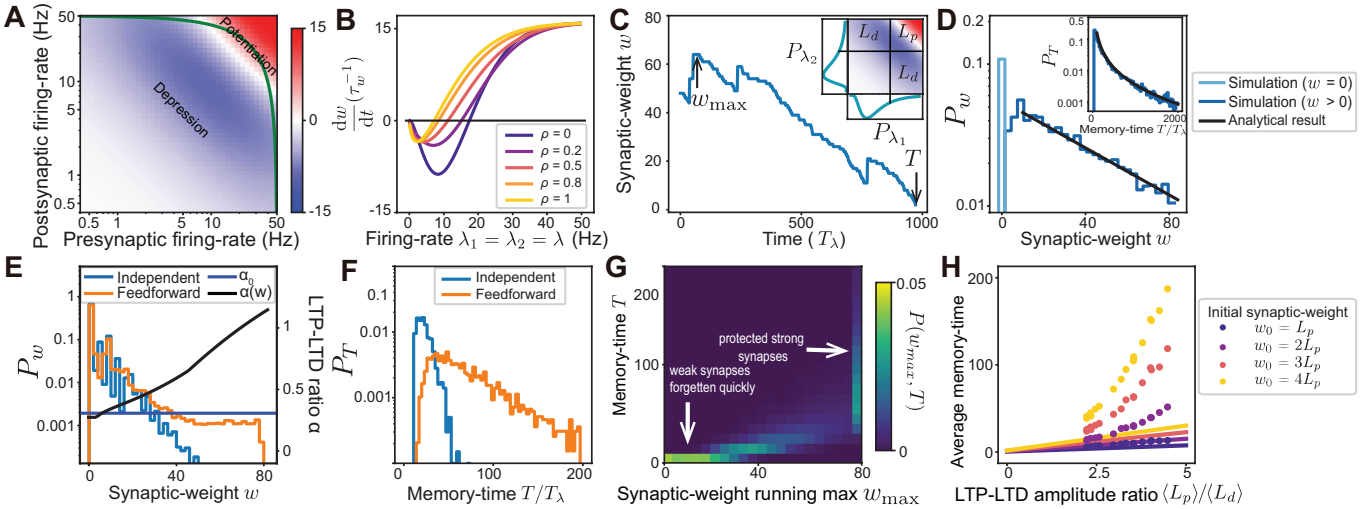

FIG. 2. Reduced model of synaptic-weight dynamics. Synaptic-weight change [$\frac{dw}{dt}$, Eq. (6)] arising from independent (A) and equal (B) pre- and postsynaptic firing-rates. In (A), the temporal correlation $\rho = 0$. The green line indicates $\lambda_1 + \lambda_2 = 50$Hz. (C) Example synaptic-weight trajectory, illustrating the memory-time T and the maximum weight $w_{\max}$. (D) Synaptic-weight distribution, including analytical result for the tail behavior in the weak-synapse scenario. (Inset) Memory-time distribution of a synapse. In the feedforward (strong-synapse) scenario, the synaptic-weight (E) and memory-time (F) distributions have heavier tails. Also shown in (E) is the potentiation-depression ratio $\alpha(w)$. (G) Joint distribution of ($w_{\max}$, T). Strong synapses are preferentially protected from forgetting. (H) Average memory-time increases linearly with $\langle L_p \rangle / \langle L_d \rangle$ in the independent case, and nonlinearly in the feedforward case. See [27] for parameter values.

Network stabilization by multiplicative synaptic plasticity.—We now leverage our results to study effects of a multiplicative plasticity on network structure. In this context, $\mathcal{C}$ represents the calcium influx into a neuron, triggered by coincident pre- and post-synaptic spikes. Large influx induces long-term potentiation (LTP; when $\mathcal{C} > \theta_p$), intermediate influx induces long-term depression (LTD; $\theta_p > \mathcal{C} > \theta_d$) [7, 9, 10]. Given the potentiation and depression rates $\gamma_{p,d}$ and thresholds $\theta_{p,d}$, the synaptic-weight dynamics are,

$$\tau_w \frac{\mathrm{d}w}{\mathrm{d}t} = \gamma_p \Theta(\mathcal{C}(t) - \theta_p) - \gamma_d \Theta(\mathcal{C}(t) - \theta_d) \equiv \Gamma_p - \Gamma_d. \tag{6}$$

Note that w has lower bound at 0, and typically $\tau_w \gg \tau_{\mathcal{C}}$. Based on our analytical approximation of $P_{\mathcal{C}}$, we computed the total potentiation/depression rates $\Gamma_{p,d} \simeq \frac{\gamma_{p,d}}{\Gamma(\tilde{k})} \int_{\theta_{p,d}}^{\infty} C^{\tilde{k}-1} e^{-C/\tilde{\sigma}} \mathrm{d}C$, which depend on the spike-train properties and the plasticity rule [Eqs. (4, 5)].

Neuronal activity *in vivo* undergoes substantial firing-rate fluctuations, generated by external input variability or intrinsic dynamics [31]. Below we formulate a mean-field approximation, reducing the joint dynamics of neurons and synapses and accounting for network structure, to the effective dynamics of a pair of pre- and postsynaptic neurons and the synapse connecting them. The statistical properties of the reduced system recapitulate the network behavior.

In the reduced model, we assume that the neurons' firing-rates (λ_1, λ_2) are sampled from P_λ, and have correlation time $T_\lambda \sim 0.1-1$s. P_λ may depend on the synaptic strength, and will be determined self-consistently, accounting for network interactions. During an interval T_λ, the rates (λ_1, λ_2) are approximately constant, so the weight change is $\Delta w = T_\lambda/\tau_w \times (\Gamma_p - \Gamma_d)$. Its distribution $P_{\text{step}}(\Delta w)$ is calculated from P_λ through Eq. (6) (Fig. 2A,B). Thus, the synaptic dynamics are reduced to a 1D random walk on $w \geq 0$ (Fig. 2C) with weight-dependent step-size distribution P_{step}. We identify its steady-state distribution P_w with the synaptic-weight distribution of the network. Importantly, for $T_\lambda \sim 1$s, Δw is not infinitesimal, so the small step-size approximation [13, 14, 32] is invalid. Next we use the mean-field approach to study representative network architectures.

Weak synapses. In this limit, the pre- and postsynaptic firing-rates (λ_1, λ_2) are sampled *independently* of the synaptic-weight w. P_{step} is further assumed to be discrete, such that a synapse can be unchanged/potentiated/depressed by fixed amounts $\Delta w = 0$, L_p, $-L_d$ with probabilities α_0, α_p, α_d. Equivalently, $P_{\text{step}}(\Delta w) = \alpha_0 \delta(\Delta w) + \alpha_d \delta(\Delta w + L_d) + \alpha_p \delta(\Delta w - L_p)$. Such a scenario would be expected in a network switching between "high" and "low" states. For illustration, we assume that L_p/L_d is an integer (see [27] for more general analysis). We define the potentiation-depression ratio $\alpha \equiv \alpha_p L_p / \alpha_d L_d$. When depression dominates ($\alpha < 1$), using results for random walks [33, 34], P_w is unimodal, and its tail follows $P_w \sim h(\beta) e^{-\frac{\beta w}{L_d}}$, $w \gg L_p$. The factors β, $h(\beta)$ were determined by analyzing the moment generating function of w [27]. We find that the tail becomes

heavier as overall potentiation and depression are more closely balanced ($\alpha_d L_d \gtrsim \alpha_p L_p$); or as L_p/L_d is larger with fixed α. Generally, when weight changes (Δw) are non-negligible relative to the mean, the distribution P_w is unimodal, and its tail behavior is sensitive to the high-order statistics of the step-size distribution P_{step}, in contrast to the case of infinitesimal Δw [14].

To study the process of forgetting in the reduced model, we envision a potentiated synapse with initial weight w_0 representing a certain memory. The memory-time $T(w_0)$ is that synapse's first-passage time to 0. Analysis of the random walk statistics [27] gives the average and tail behavior of $T(w_0)$,

$$\langle T(w_0)\rangle = \frac{w_0}{\alpha_d L_d - \alpha_p L_p} \quad \text{and} \quad P_T(t) \sim w_0 t^{-\frac{3}{2}} e^{-\frac{t}{\kappa}}. \tag{7}$$

See [27] for expressions of κ. Similarly to P_w, the average memory-time becomes longer and the tail becomes heavier as L_p/L_d increases, with fixed α (Fig. 2E).

Strong synapses. When synapses are strong [35, 36], P_λ becomes weight-dependent. The postsynaptic neuron receives *feedforward* weighted presynaptic input (firing-rate λ_1, weight w) and background input from the rest of the network. Both inputs switch between high and low firing-rates. Here, using the heuristic MM to compute $P_{\text{step}}(\Delta w)$ requires knowing how spike-time correlations depend on w and the background input. This relationship was determined by matching the postsynaptic neuron with a leaky-integrate-and-fire neuron driven by presynaptic shot-noise and background Gaussian noise [27]. We then numerically evaluated P_w, P_T at steady-state, showing a substantially heavier tail when compared to the independent case (while fixing the overall potentiation-depression ratio $\langle\alpha(w)\rangle_{P_w}$, Fig. 2E,F).

Inspection of the joint distribution of the synaptic-weight running maximum and the memory-time $(w_{\max}, T)$ suggests that strong synapses are specifically resistant to forgetting (Fig. 2G). Moreover, the average memory-time increases nonlinearly with the LTP-LTD amplitude ratio (Fig. 2H), compared to an approximately linear increase in the independent case [Eq. (7), again matching $\langle\alpha(w)\rangle_{P_w}$]. Similar results were observed in a reduced clustered recurrent network model [27].

Taken together, in the regime where a small number of inputs is sufficient to trigger a postsynaptic response, the multiplicative plasticity rule supports a unimodal synaptic-weight distribution in which strong synapses are preferentially protected from turnover.

Spiking network simulations.—We tested our results by simulating a network of leaky-integrate-and-fire neurons. The network consists of two excitatory (E) clusters which mutually inhibit each other indirectly via inhibitory clusters (Fig. 3A, see [27] for networks with > 2 clusters). Initially, intra-cluster $E \to E$ connections are strong while inter-cluster connections are weak. Crucially, the probability that an $E \to E$ connection exists is independent of the cluster assignment. The initial structure may represent two mutually exclusive memories stored in the network that spontaneously switch on timescale of $\sim$ 0.2s (Fig. 3A). Structured inhibition is consistent with experiments showing inhibitory stimulus-specific ensembles, and may arise from inhibitory plasticity [37–39]. We investigated memory retention when $E \to E$ synapses undergo multiplicative plasticity, by examining the steady-state statistics of $\mathcal{C}$, network structure, and dynamics.

In this network, potentiation is more likely in intra-cluster relative to inter-cluster synapses, so the tail of P_w for intra-cluster synapses is heavier (Fig. 3B). Yet, notably, P_w is unimodal with only minimal saturation to the upper bound. To examine stability of network structure, we plotted the steady-state $E \to E$ weight matrix spectrum (Fig. 3C). The spectral distribution's bulk follows the circular law for a network with independent, random weights, sampled from cluster-specific distributions [27, 40]. Additionally, there are two outlying eigenvalues. The fact that the larger eigenvalue (corresponding to the "DC" eigenvector) does not saturate to its maximum possible value, together with the stability of the switching dynamics, suggests that there is no runaway potentiation of either cluster. The smaller eigenvalue corresponds to an eigenvector that follows from the clustered connectivity. Angles between the plastic network's eigenvectors and those of a network with perfect cluster structure are stable and much smaller than angles computed for a network with shuffled connections, indicating that network structure is preserved despite ongoing plasticity. As predicted by the mean-field analysis, strong synapses are protected from rapid turnover. The dynamics of the plastic network also retain the cluster properties, exhibiting larger intra-cluster spike-time correlations and larger avalanches, than a shuffled network. Intriguingly, avalanche statistics are closely related to the synaptic-weight distribution [41].

To understand the implications of the multiplicative rule beyond stability, we extended the results in [42] and analytically computed the memory capacity of Hopfield-like network, defined such that the variance of the synaptic-weight distribution is independent of the memory load [27]. Fig. 3H shows that a heavier tail of the distribution, similarly to P_w in Fig. 2E, leads to a marked increase in capacity. Furthermore, we demonstrate in [27] that the multiplicative plasticity rule supports efficent learning of structured connectivity (akin to Fig. 3A), reflecting the structure of an external input.

In [27] we explored the sensitivity of the spiking-network stability results to changes of two key parameters, the potentiation/depression thresholds [$\theta_{p,\,d}$ in Eq. (6)] and the structure of inhibition.

We additionally showed that networks with intrinsically bistable synapses [10] are also stable, but do not exhibit realistic synaptic-weight distributions or activity-

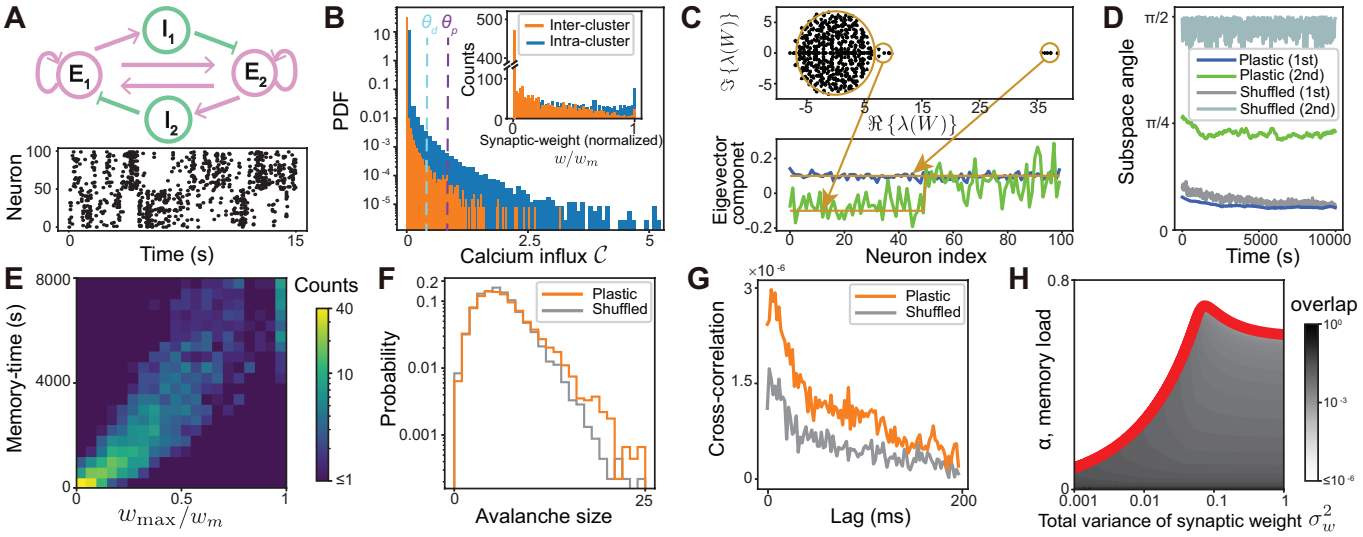

FIG. 3. Recurrent spiking neural network simulations with synapses subject to the multiplicative plasticity rule. (A) Initial network structure, and steady-state spiking activity. (B) Calcium influx and synaptic-weight (inset) distributions at steady-state, for inter- and intra-cluster connections. (C) Eigenvalues (top) and eigenvectors (bottom) of the $E \to E$ steady-state synaptic-weight matrix. The second outlier suggests that the network's cluster structure is preserved at steady-state. (D) The first two subspace angles between the plastic network and a network with perfect cluster structure as a function of time. The second angle in the plastic network remains far from $\pi/2$ compared to the shuffled network. (E) Joint distribution of $(w_{\max}, T)$ for intra-cluster $E \to E$ connections at steady-state. Similarly to the reduced model (Fig. 2C), strong synapses are preferentially protected. (F,G) The plastic network exhibits larger avalanches and stronger intra-cluster spike-timing correlations, compared to the shuffled network. (H) Memory capacity of a Hopfield-like network versus synaptic-weight variability. Synaptic-weights are normalized such that the variance is independent of the memory load α. See [27] for simulation and calculation details.

dependent protection of strong synapses. Highlighting the importance of the multiplicative rule's statistics, we found that an additive plasticity rule with $\mathcal{C} = c_1 + c_2$ [instead of Eq. (2)] rapidly leads to instability, and is unable to efficiently learn the structure of an external input [27].

Our analysis offers insights to the two-timescale problem, where synaptic interactions determine network dynamics on short timescales, and undergo neural activity-dependent modifications on longer timescales. Importantly, we analyze the network in a regime where strict separation of timescales does not hold. Previous studies utilizing plasticity rules where modifications depend (possibly nonlinearly) on sums over pre- and postsynaptic activity, typically resulted in unrealistic synaptic-weight or firing-rate distributions, or required fast homeostatic mechanisms for stability [13–16, 43–46]. The multiplicative structure of the plasticity rule analyzed here effectively eliminates modifications due to 'spurious' activity, while specific patterns of activity are responsible for potentiation and learning. The general structure of the multiplicative process introduced here suggests that our results could be applied to understand nonlinear and adaptive interacting systems in a broad range of scientific fields.

We thank Nicolas Brunel for discussions, and Guillaume Lajoie and Wilten Nicola for comments on the manuscript. This work was supported by DOE grant DE-SC0022042. B.W. thanks the UCSD Friends of the International Center for support.

[1] R. Sarfati, J. C. Hayes, É. Sarfati, and O. Peleg, Journal of the Royal Society Interface **17**, 20200179 (2020).
[2] C. M. Crain, K. Kroeker, and B. S. Halpern, Ecology Letters **11**, 1304 (2008).
[3] T. Winkler, U. Kettling, A. Koltermann, and M. Eigen, Proceedings of the National Academy of Sciences **96**, 1375 (1999).
[4] W. Schultz, Journal of Neurophysiology **80**, 1 (1998).
[5] M. Botvinick, J. X. Wang, W. Dabney, K. J. Miller, and Z. Kurth-Nelson, Neuron (2020).
[6] G.-q. Bi and M.-m. Poo, Journal of Neuroscience **18**, 10464 (1998).
[7] Y. Inglebert, J. Aljadeff, N. Brunel, and D. Debanne, Proceedings of the National Academy of Sciences **117**, 33639 (2020).
[8] D. L. Snyder and M. I. Miller, *Random point processes in time and space* (Springer Science & Business Media, 2012).
[9] H. Z. Shouval, M. F. Bear, and L. N. Cooper, Proceedings of the National Academy of Sciences **99**, 10831 (2002).
[10] M. Graupner and N. Brunel, Proceedings of the National Academy of Sciences **109**, 3991 (2012).
[11] S. S.-H. Wang, W. Denk, and M. Häusser, Nature Neu-

roscience **3**, 1266 (2000).
[12] T. Nevian and B. Sakmann, Journal of Neuroscience **26**, 11001 (2006).
[13] R. Kempter, W. Gerstner, and J. L. van Hemmen, Physical Review E **59**, 4498 (1999).
[14] J. Rubin, D. D. Lee, and H. Sompolinsky, Physical Review Letters **86**, 364 (2001).
[15] G. K. Ocker, A. Litwin-Kumar, and B. Doiron, PLoS Computational Biology **11**, e1004458 (2015).
[16] N. Ravid Tannenbaum and Y. Burak, PLoS Computational Biology **12**, e1005056 (2016).
[17] G. Lajoie, N. I. Krouchev, J. F. Kalaska, A. L. Fairhall, and E. E. Fetz, PLoS Computational Biology **13**, e1005343 (2017).
[18] F. Zenke, W. Gerstner, and S. Ganguli, Current Opinion in Neurobiology **43**, 166 (2017).
[19] G. Peccati and M. Reitzner, *Stochastic analysis for Poisson point processes: Malliavin calculus, Wiener-Itô chaos expansions and stochastic geometry*, Vol. 7 (Springer, 2016).
[20] H. Takayasu, A.-H. Sato, and M. Takayasu, Physical Review Letters **79**, 966 (1997).
[21] A. Manor and N. M. Shnerb, Physical Review Letters **103**, 030601 (2009).
[22] K. J. Rubin, G. Pruessner, and G. A. Pavliotis, Journal of Physics A: Mathematical and Theoretical **47**, 195001 (2014).
[23] I. Eliazar and J. Klafter, Proceedings of the National Academy of Sciences **102**, 13779 (2005).
[24] J. Aljadeff, M. Gillett, U. P. Obilinovic, and N. Brunel, Current Opinion in Neurobiology **70**, 24 (2021).
[25] S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, PLoS Biology **3**, e68 (2005).
[26] Y. Ziv, L. D. Burns, E. D. Cocker, E. O. Hamel, K. K. Ghosh, L. J. Kitch, A. El Gamal, and M. J. Schnitzer, Nature Neuroscience **16**, 264 (2013).
[27] "The Supplemental Material contains the derivation of the steady-state distribution of the coincidence detector; mean-field analysis of the reduced model for synaptic-weight dynamics; details of spiking network simulations; analysis of Hopfield-type model with constrained synaptic-weight variability; supplementary discussion; and citations of refs. [47-62],".
[28] M. Nishiyama, K. Hong, K. Mikoshiba, M.-M. Poo, and K. Kato, Nature **408**, 584 (2000).
[29] J. F. C. Kingman, *Poisson processes*, Vol. 3 (Clarendon Press, 1992).
[30] M. J. E. Richardson and R. Swarbrick, Physical Review Letters **105**, 178102 (2010).
[31] M. M. Churchland, M. Y. Byron, J. P. Cunningham, L. P. Sugrue, M. R. Cohen, G. S. Corrado, W. T. Newsome, A. M. Clark, P. Hosseini, B. B. Scott, *et al.*, Nature Neuroscience **13**, 369 (2010).
[32] M. O. Magnasco, O. Piro, and G. A. Cecchi, Physical Review Letters **102**, 258102 (2009).
[33] W. Feller, *An introduction to probability theory and its applications, vol 2* (John Wiley & Sons, 2008).
[34] P. Flajolet and R. Sedgewick, *Analytic combinatorics* (Cambridge University Press, 2009).
[35] J. Barral and A. D. Reyes, Nature Neuroscience **19**, 1690 (2016).
[36] Y. Ahmadian and K. D. Miller, Neuron **109**, 3373 (2021).
[37] A. G. Khan, J. Poort, A. Chadwick, A. Blot, M. Sahani, T. D. Mrsic-Flogel, and S. B. Hofer, Nature Neuroscience **21**, 851 (2018).
[38] T. Geiller, S. Sadeh, S. V. Rolotti, H. Blockus, B. Vancura, A. Negrean, A. J. Murray, B. Rózsa, F. Polleux, C. Clopath, *et al.*, Nature , 1 (2021).
[39] R. E. Field, J. A. Damour, R. Tremblay, C. Miehl, B. Rudy, J. Gjorgjieva, and R. C. Froemke, Neuron **106**, 842 (2020).
[40] J. Aljadeff, M. Stern, and T. Sharpee, Physical Review Letters **114**, 088101 (2015).
[41] L. Kuśmierz, S. Ogawa, and T. Toyoizumi, Physical Review Letters **125**, 028101 (2020).
[42] U. Pereira and N. Brunel, Neuron **99**, 227 (2018).
[43] A. Litwin-Kumar and B. Doiron, Nature Communications **5**, 1 (2014).
[44] D. Higgins, M. Graupner, and N. Brunel, PLoS Computational Biology **10**, e1003834 (2014).
[45] A. E. Akil, R. Rosenbaum, and K. Josić, PLoS Computational Biology **17**, e1008958 (2021).
[46] Y. F. K. Kossio, S. Goedeke, C. Klos, and R.-M. Memmesheimer, Proceedings of the National Academy of Sciences **118** (2021).
[47] M. Graupner, P. Wallisch, and S. Ostojic, Journal of Neuroscience **36**, 11238 (2016).
[48] B. Doiron, A. Litwin-Kumar, R. Rosenbaum, G. K. Ocker, and K. Josić, Nature Neuroscience **19**, 383 (2016).
[49] G. Last and M. Penrose, *Lectures on the Poisson process*, Vol. 7 (Cambridge University Press, 2017).
[50] T. P. Minka, Microsoft Research, Cambridge, UK, Tech. Rep (2002).
[51] M. Abramowitz, I. A. Stegun, and R. H. Romer, *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (American Association of Physics Teachers, 1988).
[52] I. Adan and J. Resing, *Queueing theory* (Eindhoven University of Technology, 2002).
[53] J. Wendel, The American Mathematical Monthly **82**, 494 (1975).
[54] S. G. Kou and H. Wang, Advances in Applied Probability **35**, 504 (2003).
[55] A. van Meegen and S. J. van Albada, Physical Review Research **3**, 043077 (2021).
[56] K. A. Wilmes and C. Clopath, Nature Communications **10**, 1 (2019).
[57] R. L. Redondo and R. G. Morris, Nature Reviews Neuroscience **12**, 17 (2011).
[58] G. G. Turrigiano, K. R. Leslie, N. S. Desai, L. C. Rutherford, and S. B. Nelson, Nature **391**, 892 (1998).
[59] G. G. Turrigiano, Cell **135**, 422 (2008).
[60] R. E. Mirollo and S. H. Strogatz, SIAM Journal on Applied Mathematics **50**, 1645 (1990).
[61] R. Sarfati, J. C. Hayes, and O. Peleg, Science Advances **7**, eabg9259 (2021).
[62] G. Hennequin, E. J. Agnes, and T. P. Vogels, Annual Review of Neuroscience **40**, 557 (2017).