LETTER **Communicated by Christian Maes**

Cooperativity, Information Gain, and Energy Cost During Early LTP in Dendritic Spines

Jan Karbowski
jkarbowski@mimuw.edu.pl
Institute of Applied Mathematics and Mechanics, University of Warsaw, Warsaw 02-097, Poland

Paulina Urban
p.urban@student.uw.edu.pl
College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences and Laboratory of Functional and Structural Genomics, Centre of New Technologies, University of Warsaw, Warsaw 02-097, Poland; and Laboratory of Databases and Business Analytics, National Information Processing Institute, National Research Institute, Warsaw 00-608, Poland

We investigate a mutual relationship between information and energy during the early phase of LTP induction and maintenance in a large-scale system of mutually coupled dendritic spines, with discrete internal states and probabilistic dynamics, within the framework of nonequilibrium stochastic thermodynamics. In order to analyze this computationally intractable stochastic multidimensional system, we introduce a pair approximation, which allows us to reduce the spine dynamics into a lower-dimensional manageable system of closed equations. We found that the rates of information gain and energy attain their maximal values during an initial period of LTP (i.e., during stimulation), and after that, they recover to their baseline low values, as opposed to a memory trace that lasts much longer. This suggests that the learning phase is much more energy demanding than the memory phase. We show that positive correlations between neighboring spines increase both a duration of memory trace and energy cost during LTP, but the memory time per invested energy increases dramatically for very strong, positive synaptic cooperativity, suggesting a beneficial role of synaptic clustering on memory duration. In contrast, information gain after LTP is the largest for negative correlations, and energy efficiency of that information generally declines with increasing synaptic cooperativity. We also find that dendritic spines can use sparse representations for encoding long-term information, as both energetic and structural efficiencies of retained information and its lifetime exhibit maxima for low fractions of stimulated synapses during LTP.

Jan Karbowski is the corresponding author.

Neural Computation **36, 271–311** (2024)
https://doi.org/10.1162/neco_a_01632

Moreover, we find that such efficiencies drop significantly with increasing the number of spines. In general, our stochastic thermodynamics approach provides a unifying framework for studying, from first principles, information encoding, and its energy cost during learning and memory in stochastic systems of interacting synapses.

1 Introduction

Parts of synapses known as dendritic spines play an important role in learning and memory in neural networks (Bonhoeffer & Yuste, 2002; Kasai et al., 2003; Bourne & Harris, 2008; Takeuchi et al., 2014; Kandel et al., 2014). Learning can be thought of as acquiring information in synapses through a plasticity mechanism such as LTP and LTD (long-term potentiation and depression, respectively), and memory can be regarded as storing that information. (For an experimental overview, see Yang et al., 2009; Bourne & Harris, 2008; Takeuchi et al., 2014; and Poo et al., 2016. For a theoretical overview, see Fusi et al., 2005; Benna & Fusi, 2016; and Chaudhuri & Fiete, 2016.) Thus, learning and memory are strictly related to processing and maintaining long-term information, which in principle could be quantified in terms of information theory and statistical mechanics, much as it can be done for neural spiking activity (Rieke et al., 1999). Indeed, recent results show that information (entropy) contained in the distributions of dendritic spine volumes and areas is nearly maximal for any of their average sizes across different brains and cerebral regions (Karbowski & Urban, 2022). This suggests that the concept of information can be useful in quantifying synaptic learning and memory and that actual synapses might "use" and optimize certain information-theoretic quantities.

Physics teaches us that there are close relationships between information and thermodynamics, and the smaller the system, the stronger the mutual link, since smallness contributes to fluctuations and thus unpredictability in the system (Bennett, 1982; Leff & Rex, 1990; Berut et al., 2012; Parrondo et al., 2015). This means that information processing always requires some energy, and it is reasonable to assume that biological evolution favors systems that save energy while handling information because of the limited resources and/or competition (Niven & Laughlin, 2008). This line of thought was explored in neuroscience for estimating energy-efficient coding capacity in (short-term) neural activities and synaptic transmissions (Rieke et al., 1999; Levy & Baxter, 1996, 2002; Levy & Calvert, 2021; Laughlin et al., 1998; Balasubramanian et al., 2001). All of these approaches and calculations for neural and synaptic activities, however valuable, missed one key ingredient of real biological systems or did not make it explicit: biological systems, including neurons and synapses, always operate far from thermodynamic equilibrium, where balance between incoming and outgoing energy and matter (or probability) fluxes is broken—the so-called broken detailed

balance. (For a general physical approach, see Maes et al., 2000; Seifert, 2012; and Gardiner, 2004. For a biophysical approach to synapses, see Karbowski, 2019.) As a consequence, equilibrium thermodynamics (with static or stationary variables and Gibbs distributions) does not seem to be the right approach, and to be more realistic, one has to use nonequilibrium statistical mechanics, where the concepts of stochasticity and entropy production play central roles (and where we do not know a priori the probability distributions). Nonequilibrium statistical mechanics (or nonequilibrium stochastic thermodynamics) provides a unifying description for stochastic dynamics because it treats microscopic information and energy on the same footing, which allows us to get the right estimates of information and energy rates from first principles. Such an approach was initiated in Karbowski (2019, 2021) for studying nonequilibrium thermodynamics of synaptic plasticity. Both of these studies suggested that synaptic plasticity can use energy rather economically, since (1) it consumes only about 4% to 11% of energy devoted for fast synaptic transmission (Karbowski, 2019), and (2) it can provide higher coding accuracy and longer memory time for a lower energy cost at certain regimes (Karbowski, 2019, 2021).

There exists large experimental evidence that local synaptic cooperativity on a dendrite takes place during learning and memory formation (Makino & Malinow, 2011, and Yadav et al., 2012; for a review see Winnubst et al., 2012), and it can also be useful for long-term memory stability (Govindarajan et al., 2006; Kastellakis & Poirazi, 2019). Thus, it seems that any realistic model of synaptic plasticity relevant for memory formation should include correlations between neighboring synapses. Moreover, it would be good to know how such correlations influence the lifetime of memory trace, as well as information gain and the energy cost associated with it.

This study explores nonequilibrium statistical mechanics for investigating the efficiency of learning and storing information in a system of interacting synapses with stochastic dynamics during early LTP induction and its maintenance (e-LTP phase, without consolidation). Not only fast synaptic transmission is noisy (Volgushev et al., 2004); the noise is also present in long-term synaptic dynamics associated with slow plasticity due to large thermal fluctuations in internal molecules and presynaptic input (Bonhoeffer & Yuste, 2002; Holtmaat et al., 2005; Choquet & Triller, 2013; Statman et al., 2014; Meyer et al., 2014; Kasai et al., 2003; Loewenstein et al., 2011). Thus, synaptic plasticity requires a probabilistic approach (Yasumatsu et al., 2008), and we assume that it can be described as dynamics involving transitions between discrete mesoscopic states (Montgomery & Madison, 2004; Fusi et al., 2005; Leibold & Kempter, 2008; Barrett et al., 2009; Benna & Fusi, 2016). Our letter extends the previous two studies (Karbowski, 2019, 2021) in three important methodological ways. First, it provides a general framework for approximating the dynamics of a multidimensional stochastic system of N interacting synapses each with four internal states (which in practice is computationally intractable for large N) by reducing it to a set of

coupled lower-dimensional stochastic subsystems (which are computationally tractable). This is done by applying the so-called pair approximation, which allows us to reduce the system with 4^N equations to the system with about $4(5N - 4)$ equations. Second, our letter provides explicit formulas for investigating information gain (rate of Kullback-Leibler divergence) and energy consumption (entropy production rate) for arbitrary time during learning and memory retention phases. Third, we use data-driven estimates for transition rates between different synaptic states, and hence our values of information and energy are realistic.

On a conceptual level, our study investigates how the efficiency of encoded information (both amount and duration) depends on a degree of correlation between neighboring synapses, a percentage of their activation by presynaptic neurons, and the magnitude and duration of synaptic stimulation during LTP. The first relates to cooperativity between synapses, the second to sparseness of synaptic coding, and the last to the strength of learning. All of these parameters can in principle be compared to empirical data once those are available, thus providing an important link between theory and experiment.

Throughout the letter, we use interchangeably the terms *dendritic spine* and *synapse*.

2 Model of Plastic Interacting Dendritic Spines

We consider a single postsynaptic neuron having one basal (main) dendrite with N dendritic spines located along its length (see Figure 1). Because of the small sizes of dendritic spines (about 1 μm), their dynamics are necessarily probabilistic due to thermodynamic fluctuations of local environment (with high temperature of about 300 K), as well as due to activity fluctuations in neurons (both an electric and chemical nature). Moreover and importantly, the spines are locally coupled by nearest-neighbor interactions, as the experimental data suggest (Makino & Malinow, 2011; Yadav et al., 2012; Winnubst et al., 2012).

2.1 Morphological Spine States and Stochastic Multidimensional Dynamics. Empirical data indicate that a single dendritic spine can be regarded as a four-state stochastic system with well-defined morphological (mesoscopic) states (Montgomery & Madison, 2004; Bokota et al., 2016; Basu et al., 2018; Urban et al., 2020; see Figure 1A). These states are denoted here as s_i at each location i, with values $s_i = 0, 1, 2, 3$, corresponding respectively to the following morphological states: nonexistent, stubby, filopodia/thin, and mushroom (the larger s_i, the larger the spine size; see appendix A). These mesoscopic states are quasi-stable, which means that there are slow transitions between them that are much slower than microscopic transitions between molecular, mostly unknown, processes comprising internal microscopic dynamics of a dendritic spine (Kennedy, 2000; Sheng & Hoogenraad,

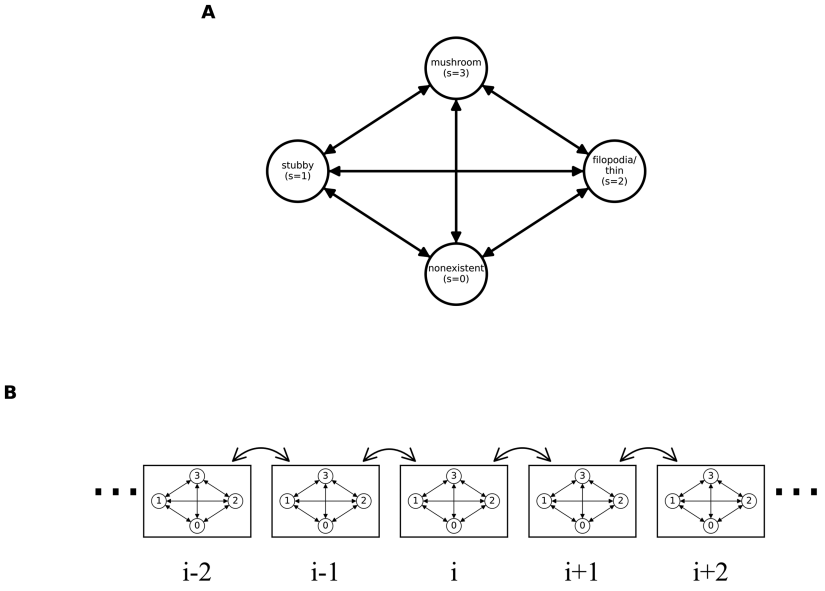

Figure 1: Mesoscopic model of dendritic spines and their interactions on a dendrite. (A) Four state stochastic models of a dendritic spine with transitions between the mesoscopic states. (B) Interactions between spines on a dendrite are determined only by nearest neighbors, in such a way that the spine's intrinsic transition rates depend on the states of neighboring spines. These interactions can be thought of as representing inflow and outflow of different molecules between spines.

2007; Miller et al., 2005; Kandel et al., 2014). This approach can be treated as a coarse-grained description of intrinsic spine dynamics in terms of stochastic Markov process on a mesoscopic scale.

We assume stochastic dynamics for N-coupled dendritic spines and denote by $P(s_1, s_2, \ldots, s_N; t)$ the probability that the spine system has the configuration of internal states $s_1, s_2, \ldots, s_N$. The dynamic of this global stochastic state is motivated by the Glauber model of time-dependent Ising model (Glauber, 1963), and it is represented by the following master equation (see appendix A)

$$\frac{dP(s_1, \ldots, s_N)}{dt} = \sum_{i=1}^{N} \sum_{s_i'} w_{s_i, s_i'}(s_{i-1}, s_{i+1}) P(s_1, \ldots, s_i', \ldots, s_N) - P(s_1, \ldots, s_N) \sum_{i=1}^{N} \sum_{s_i'} w_{s_i', s_i}(s_{i-1}, s_{i+1}), \tag{2.1}$$

Table 1: Values of Intrinsic Basic Transition Rates $v_{s,s'}$ in Each Synapse.

Matrix Element	Value
$v_{0,1}$	0.019
$v_{1,0}$	0.065
$v_{0,2}$	0.015
$v_{2,0}$	0.007
$v_{0,3}$	0.022
$v_{3,0}$	0.001
$v_{2,1}$	0.003
$v_{1,2}$	0.061
$v_{2,3}$	0.009
$v_{3,2}$	0.015
$v_{1,3}$	0.049
$v_{3,1}$	0.008

Notes: All diagonal elements (that is $v_{k,k}$) are zero. The units are in min^{-1}.

where $w_{s_i,s_i'}(s_{i-1}, s_{i+1})$ is the transition rate for the jumps inside spine i from state s_i' to state s_i. These jumps also depend on the states of neighboring spines s_{i-1} and s_{i+1}, because of the nearest-neighbor coupling between the spines (see Figure 1B). Generally we take the following form of the transition matrix $w_{s_i,s_i'}(s_{i-1}, s_{i+1})$,

$$w_{s_i,s_i'}(s_{i-1}, s_{i+1}) = v_{s_i,s_i'}\left[1 + \theta(d(s_i) - d(s_i'))\frac{\gamma}{2d(3)}[d(s_{i-1}) + d(s_{i+1})]\right] \times\big[1 + a_i(1 + \theta(d(s_i) - d(s_i')))f(t)\big], \tag{2.2}$$

where $v_{s_i,s_i'}$ is the intrinsic basic transition rate between s_i' and s_i at spine i, and it is independent of the neighboring synapses. These intrinsic rates are the same for each i, setting the temporal scale for basal synaptic plasticity, and they were estimated based on data in Urban et al. (2020) and are presented in Table 1. The symbol $d(s_i)$ is the spine size at state s_i of spine i, and it is proportional to the spine surface area (see appendix A), whereas $\theta(x)$ denotes the sign function of the argument x, where $\theta(x) = 1$ if $x \geq 0$ and $\theta(x) = -1$ if $x < 0$. Note that larger spine sizes of neighboring spines generally influence more the transition rates of the given spine, which relates to their cooperativity. The parameter γ corresponds to the magnitude of spine cooperativity between nearest neighbors, with $-1 \leq \gamma \leq 1$, and there is rescaling by the maximal spine size $d(3)$ (size in state $s = 3$ called mushroom), which ensures the positivity of all elements of the transition matrix. The positive values of γ indicate positive correlations (positive cooperativity), while its negative values mean negative correlations (negative

Table 2: Nominal Values of Global Parameters Used in Computations.

Description	Variable	Value	Units
Number of synapses	N	1000	Unitless
Cooperativity	γ	$(-1.0, 1.0)$	Unitless
Probability of stimulation	p_{act}	0.3	Unitless
Stimulation amplitude	A	200	Unitless
Decay time of stimulation	τ_1	15	min
Rising time of stimulation	τ_2	2	min
Energy scale for plasticity	ϵ	$4.6 \cdot 10^5$	kT

Notes: Value of stimulation amplitude A is motivated by experimental facts that during LTP, the rates of protein phosphorylation at PSD increase by two or three orders of magnitude with respect to baseline rates (Miller et al., 2005). Energy scale ϵ for synaptic plasticity was taken from the estimate in Karbowski (2021).

cooperativity). Note that for positive cooperativity, the local spine interactions amplify the transitions that lead to an increase of spine size and reduce transitions that decrease spine size. The opposite is true for negative cooperativity. It should be added that distant spines also can affect a given spine at location i, but that interaction has an indirect character and thus is weaker and mediated with some delay.

The last factor on the right in equation 2.2 indicates the effect of external (presynaptic) stimulation, leading to LTP, with a time-varying function known as alpha function $f(t)$ given by

$$f(t) = A(e^{-t/\tau_1} - e^{-t/\tau_2}), \tag{2.3}$$

where A is the stimulation amplitude, t is the time after stimulation onset, and τ_1, τ_2 are time constants related to falling and rising phases of the stimulation, respectively. The latter means that LTP-related stimulation lasts only about $\tau_1 + \tau_2$ (around 17 minutes; see Table 2), which we call the duration of the learning phase. After that time, all the transition rates essentially decay to their prestimulation basal values. Consequently, after the stimulation, the dynamics of synaptic plasticity is driven only by the interactions between neighboring spines and their internal states. These dynamics are slow, and we call this stage the memory phase. The prefactor a_i of $f(t)$ in equation 2.2 assumes two values: $a_i = 1$ when the spine i is stimulated with the probability p_{act}, and $a_i = 0$ when there is no stimulation with the probability $1 - p_{act}$. Note that LTP-related stimulation amplifies only the transitions increasing the spine size (the sign function θ is then 1). For the transitions decreasing the spine size, there is no amplification because then the prefactor of $f(t)$ is zero.

The important point is that during learning (the phase when the function $f(t)$ is activated), the information about the stimulation is encoded in the patterns of the probabilities $P(s_1, \ldots, s_N)$. Thus, knowing how these patterns change in time provides necessary "data" for computing physical characteristics of learning and memory.

2.2 Reduction of Multidimensional Spine Stochastic Dynamics into Low-Dimensional Dynamics: Pair Approximation. Equation 2.1 describes the dynamics of the multidimensional probability that involves a gigantic 4^N number of equations. For a typical number of synapses on a dendrite $N \sim 10^3$, the dynamics represented by equation 2.1 require $\sim 10^{600}$ equations, which are impossible to simulate and solve on any computer. This numerical impossibility forces us to find an approximation to the dynamics in equation 2.1. Consequently, we consider a lower-dimensional dynamics involving only singlets and pairs of locally interacting spines. This strategy is sufficient to describe the global dynamics of the synaptic system and to compute information and energy rates if we make a certain reasonable assumption (see below).

The probabilities for singlets and doublets of spine states $P(s_i)$ and $P(s_i, s_{i+1})$ can be obtained from equation 2.1 by summations over all other states in other synapses as

$$P(s_i) = \sum_{s_1} \cdots \sum_{s_{i-1}} \sum_{s_{i+1}} \cdots \sum_{s_N} P(s_1, \ldots, s_i, \ldots, s_N)$$

and, similarly,

$$P(s_i, s_{i+1}) = \sum_{s_1} \cdots \sum_{s_{i-1}} \sum_{s_{i+2}} \cdots \sum_{s_N} P(s_1, \ldots, s_i, \ldots, s_N).$$

As a result, we obtain two equations:

$$\frac{dP(s_i)}{dt} = \sum_{s_{i-1}} \sum_{s_{i+1}} \sum_{s_i'} \left[w_{s_i, s_i'}(s_{i-1}, s_{i+1}) P(s_{i-1}, s_i', s_{i+1}) - w_{s_i', s_i}(s_{i-1}, s_{i+1}) P(s_{i-1}, s_i, s_{i+1}) \right], \tag{2.4}$$

which is valid for $i = 2, \ldots, N-1$, and

$$\frac{dP(s_i, s_{i+1})}{dt} = \sum_{s_{i-1}} \sum_{s_i'} \left[w_{s_i, s_i'}(s_{i-1}, s_{i+1}) P(s_{i-1}, s_i', s_{i+1}) - w_{s_i', s_i}(s_{i-1}, s_{i+1}) P(s_{i-1}, s_i, s_{i+1}) \right]$$

$$+ \sum_{s_{i+2}} \sum_{s'_{i+1}} \left[w_{s_{i+1},s'_{i+1}}(s_i, s_{i+2}) P(s_i, s'_{i+1}, s_{i+2}) \right.$$

$$\left. - w_{s'_{i+1},s_{i+1}}(s_i, s_{i+2}) P(s_i, s_{i+1}, s_{i+2}) \right], \tag{2.5}$$

which is valid for $i = 2, \ldots, N-2$. For the boundary probabilities, that is, for the dynamics of $P(s_1), P(s_N)$ and $P(s_1, s_2), P(s_{N-1}, s_N)$, we have similar equations, except we drop the boundary terms s_0, s_{N+1} as they are nonexistent.

Equations 2.4 and 2.5 for the dynamics of $P(s_i)$ and $P(s_i, s_{i+1})$ involve additionally the probabilities of spine triplets $P(s_{i-1}, s_i, s_{i+1})$ and $P(s_i, s_{i+1}, s_{i+2})$, and thus they do not form a closed system of equations. To close these equations, we use the so-called pair approximation for probabilities. The main idea in this approximation is that the biggest influence on a given synapse is exerted only by the nearest-neighbor synapses, and the effects from remote neighbors can be neglected, as is implied by the form of the transition rate $w_{s_i,s'_i}(s_{i-1}, s_{i+1})$. Specifically, for three neighboring synapses indexed spatially as $i-1, i, i+1$, the dynamic of synapse $i-1$ depends directly only on the state of synapse i, and the influence of $i+1$ synapse can be neglected as coming from the remote site. In terms of probabilities, this can be written as

$$P(s_{i-1}, s_i, s_{i+1}) \approx P(s_{i-1}, s_i) P(s_i, s_{i+1}) / P(s_i), \tag{2.6}$$

where we used the approximation for the conditional probability $P(s_{i-1}|s_i, s_{i+1}) \approx P(s_{i-1}|s_i)$ and the fact that $P(s_{i-1}|s_i) = P(s_{i-1}, s_i)/P(s_i)$. Thus, the probabilities of the spine triplets can be effectively written as combinations of the probabilities for spine singlets and doublets, which forms the essence of the pair approximation. A similar expression can be obtained for synapses with other combinations of indexes.

The above pair approximation allows us to write the dynamics of probabilities $P(s_i)$ and $P(s_i, s_{i+1})$ as

$$\frac{dP(s_i)}{dt} = \sum_{s_{i-1}} \sum_{s_{i+1}} \sum_{s'_i} \left[w_{s_i,s'_i}(s_{i-1}, s_{i+1}) \frac{P(s_{i-1}, s'_i) P(s'_i, s_{i+1})}{P(s'_i)} \right.$$

$$\left. - w_{s'_i,s_i}(s_{i-1}, s_{i+1}) \frac{P(s_{i-1}, s_i) P(s_i, s_{i+1})}{P(s_i)} \right] \tag{2.7}$$

for $i = 2, \ldots, N-1$, and

$$\frac{dP(s_i, s_{i+1})}{dt} = \sum_{s_{i-1}} \sum_{s'_i} \left[w_{s_i,s'_i}(s_{i-1}, s_{i+1}) \frac{P(s_{i-1}, s'_i) P(s'_i, s_{i+1})}{P(s'_i)} \right.$$

$$\left. - w_{s'_i,s_i}(s_{i-1}, s_{i+1}) \frac{P(s_{i-1}, s_i) P(s_i, s_{i+1})}{P(s_i)} \right]$$

$$+\sum_{s_{i+2}}\sum_{s'_{i+1}}\left[w_{s_{i+1},s'_{i+1}}(s_i,s_{i+2})\frac{P(s_i,s'_{i+1})P(s'_{i+1},s_{i+2})}{P(s'_{i+1})}\right.$$

$$\left.-\,w_{s'_{i+1},s_{i+1}}(s_i,s_{i+2})\frac{P(s_i,s_{i+1})P(s_{i+1},s_{i+2})}{P(s_{i+1})}\right], \tag{2.8}$$

for $i = 2, \ldots, N-2$. Similar expressions can be written for the boundary probabilities with $i = 1$ and $i = N$.

It is clear that equations 2.7 and 2.8 for the dynamics of $P(s_i)$ and $P(s_i, s_{i+1})$ form the closed system of equations, since now they only depend on each other. Importantly, the number of equations in the reduced dynamics (equations 2.7 and 2.8) is only $20N - 16$, which is linear in N and thus much smaller than the original 4^N equations, and hence feasible for numerical analysis. These two types of probabilities are sufficient to compute quantities of interest, which are associated with LTP induction, such as memory trace and its duration, the average sizes of spines, and the rates of information gain (Kullback-Leibler divergence) and energy dissipated (entropy production rate). However, first we check the accuracy of the pair approximation.

2.3 Validity of the Pair Approximation. In this section, we check how accurate the pair approximation is by considering a small system of dendritic spines with $N = 4$, for which one can find an exact numerical solution for the dynamics in equation 2.1. Our goal is to compare this exact solution with its approximation given by equations 2.7 and 2.8.

Numerical calculations indicate that the pair approximation is accurate, as exact and approximate probabilities are practically indistinguishable (see Figure 2A), even for very strong couplings between spines ($\gamma = -0.9$ and $\gamma = 0.9$). Moreover, the pair approximation is well defined, since it preserves positivity of all probabilities and their normalization (see Figures 2A and 2B). Additionally, as an example of the main observable used in this study, we also compared entropy production rates computed for the exact dynamics in equation 2.1, denoted as EPR_{ex}, with that computed from the approximate dynamics in equations 2.7 and 2.8 and denoted as EPR_{pa} (the formulas for the exact and approximate EPR are given, respectively, in equations 3.1 to 3.4 and 3.7 to 3.10). Both entropy production rates are also essentially indistinguishable, with a small difference between them, at most 0.6% (see Figure 2C).

In order to give a measure of the pair approximation accuracy, we introduce the ratio R for $N = 4$ spines, defined as

$$R(s_1,s_2,s_3,s_4) = \frac{P_{pa}(s_1,s_2)P_{pa}(s_2,s_3)P_{pa}(s_3,s_4)}{P_{pa}(s_2)P_{pa}(s_3)P_{ex}(s_1,s_2,s_3,s_4)}, \tag{2.9}$$

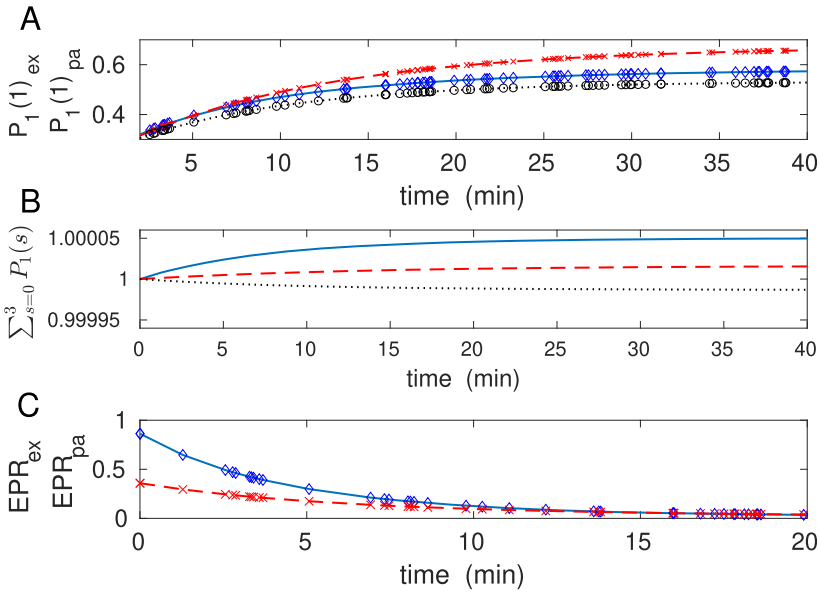

Figure 2: Comparison of an exact solution and pair approximation for $N = 4$ interacting spines. (A) Time dependence of exact $P_1(1)_{ex}$ and approximated $P_1(1)_{pa}$ probability $P(s_1 = 1)$ for three different couplings γ. Both probabilities are essentially indistinguishable. Exact solutions correspond to solid ($\gamma = -0.9$), dashed ($\gamma = 0.1$), and dotted ($\gamma = 0.9$) lines. Pair approximations correspond to diamonds ($\gamma = -0.9$), x ($\gamma = 0.1$), and circles ($\gamma = 0.9$). (B) Normalization condition for probabilities of spine 1, $\sum_{s_1=0}^{3} P(s_1)$. Solid line for $\gamma = -0.9$, dashed line for $\gamma = 0.1$, and dotted line for $\gamma = 0.9$. Note that the sum deviates from unity by a very small number—less than $5 \cdot 10^{-5}$. (C) Time dependence of exact (EPR_{ex}) and approximated (EPR_{pa}) entropy production rate. Exact EPR_{ex} correspond to solid ($\gamma = -0.9$) and dashed ($\gamma = 0.1$) lines. Pair approximations EPR_{pa} correspond to diamonds ($\gamma = -0.9$) and x ($\gamma = 0.1$). Note an excellent matching of EPR_{pa} to EPR_{ex}.

where the subscript *pa* refers to the pair approximation (see equations 2.7 and 2.8), while *ex* corresponds to the exact solution (see equation 2.1). When R approaches 1, the pair approximation matches the exact solution perfectly. The larger the deviation of R from unity, the less accurate is the approximation. This follows from the form of the pair approximation for four spines: $P(s_1, s_2, s_3, s_4) \approx P(s_1, s_2)P(s_2, s_3)P(s_3, s_4)/[P(s_2)P(s_3)]$. To have a global numerical accuracy, we have to average R over all states, which yields a mean ratio $\langle R \rangle = \frac{1}{4^4} \sum_{s_1,\ldots,s_4} R(s_1, \ldots, s_4)$, and its standard deviation $\mathrm{SD(R)} = \sqrt{\langle R^2 \rangle - \langle R \rangle^2}$, serving as a global error. In Figure 3, we show that $\langle R \rangle$ is very close to 1, and SD(R) is generally small, at most 0.1.

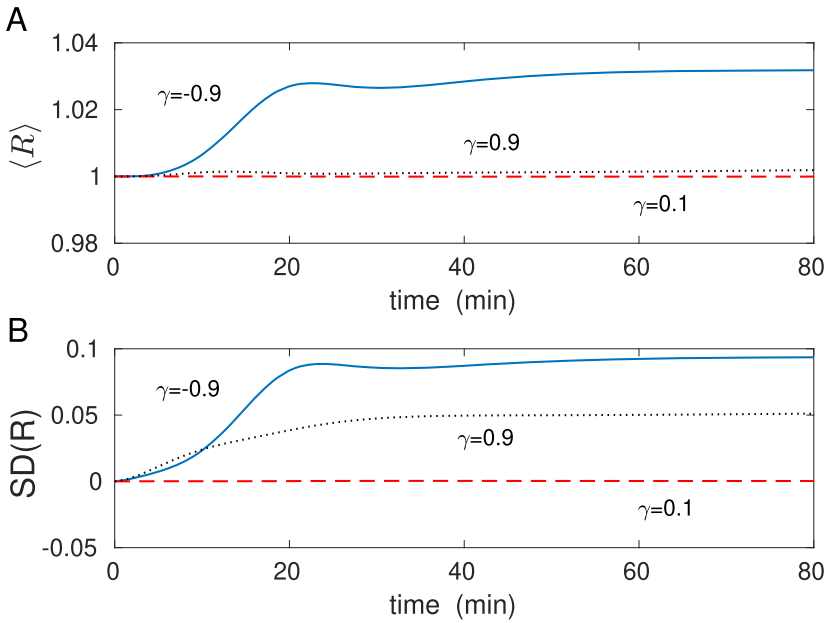

Figure 3: Accuracy measure for pair approximation with $N = 4$ interacting spines. (A) Average ratio $\langle R \rangle$ as a function of time for different magnitudes of coupling between the spines. (B) Similar to panel A but for standard deviation of the ratio R. Note that SD(R) is for moderate values of γ much less than 0.05, and it maximally achieves the value $\sim$0.1 in the extreme case $\gamma \mapsto -1$.

Taken together, the numerical results in Figures 2 and 3 indicate that the pair approximation derived in this study is quite accurate, and its accuracy is preserved in time. For an analytical example related to the pair approximation, see appendix B.

3 Derivation of Entropy Production Rate as an Energy Cost for Interacting Spines

Energy expenditure of synaptic plasticity is associated with transitions between different states of a dendritic spine. The faster the transitions, the more energy is used, and vice versa. Generally, a spine is in a thermodynamic nonequilibrium with its environment, and thus the energy cost is strictly related to the entropy production rate of the spine (for general ideas of nonequilibrium thermodynamics, see Nicolis & Prigogine, 1977, and Peliti & Pigolotti, 2021). Specifically, we assume that the energy rate associated with plasticity processes in dendritic spines is equal to the entropy production associated with stochastic transitions between spine mesoscopic

states, similar as in Karbowski (2019). In our case of N dendritic spines, the entropy production rate of the whole system $\mathrm{EPR}(s_1, \ldots, s_N)$ is given by a general formula for the entropy production (Schnakenberg, 1976; Maes et al., 2000; Seifert, 2012; Van den Broeck & Esposito, 2015):

$$\mathrm{EPR}(s_1, \ldots, s_N) = \sum_{i=1}^{N} \mathrm{EPR}_i \tag{3.1}$$

where EPR_i is the individual entropy productions of each interacting spine,

$$\mathrm{EPR}_i = \frac{\epsilon}{2} \sum_{s_1,\ldots,s_N} \sum_{s_i'} \Big[w_{s_i,s_i'}(s_{i-1}, s_{i+1}) P(s_1, \ldots, s_i', \ldots, s_N) - w_{s_i',s_i}(s_{i-1}, s_{i+1}) \times P(s_1, \ldots, s_i, \ldots, s_N) \Big] \ln \frac{w_{s_i,s_i'}(s_{i-1}, s_{i+1}) P(s_1, \ldots, s_i', \ldots, s_N)}{w_{s_i',s_i}(s_{i-1}, s_{i+1}) P(s_1, \ldots, s_i, \ldots, s_N)}, \tag{3.2}$$

for $i = 2, \ldots, N-1$, and for the boundary terms

$$\mathrm{EPR}_1 = \frac{\epsilon}{2} \sum_{s_1,\ldots,s_N} \sum_{s_1'} \big[w_{s_1,s_1'}(s_2) P(s_1', s_2, \ldots, s_N) - w_{s_1',s_1}(s_2) P(s_1, s_2, \ldots, s_N) \big] \times \ln \frac{w_{s_1,s_1'}(s_2) P(s_1', s_2, \ldots, s_N)}{w_{s_1',s_1}(s_2) P(s_1, s_2, \ldots, s_N)} \tag{3.3}$$

and

$$\mathrm{EPR}_N = \frac{\epsilon}{2} \sum_{s_1,\ldots,s_N} \sum_{s_N'} \big[w_{s_N,s_N'}(s_{N-1}) P(s_1, s_2, \ldots, s_N') - w_{s_N',s_N}(s_{N-1}) P(s_1, s_2, \ldots, s_N) \big] \times \ln \frac{w_{s_N,s_N'}(s_{N-1}) P(s_1, s_2, \ldots, s_N')}{w_{s_N',s_N}(s_{N-1}) P(s_1, s_2, \ldots, s_N)}, \tag{3.4}$$

where ϵ is the energy scale for various biophysical processes taking place inside a typical dendritic spine and related to molecular plasticity. Its value was estimated at about $\epsilon \approx 4.6 \cdot 10^5$ kT (or $2.3 \cdot 10^4$ ATP molecules), where k is the Boltzmann constant and T is the absolute brain temperature (see Karbowski, 2021). In a nutshell, these numbers can be understood by considering that a typical dendritic spine contains roughly 10^4 proteins, each with a few degrees of freedom corresponding to the number of phosphorylation sites (Sheng & Hoogenraad, 2007).

As before, we explore the pair approximation in equations 3.2 to 3.4, which in the case of N spines takes the form

$$\begin{aligned} &P(s_1, \ldots, s_{i-1}, s_i', s_{i+1}, \ldots, s_N) \\ &\approx \frac{P(s_1, s_2) \ldots P(s_{i-1}, s_i') P(s_i', s_{i+1}) \ldots P(s_{N-1}, s_N)}{P(s_1) \ldots P(s_{i-1}) P(s_i') P(s_{i+1}) \ldots P(s_{N-1})}. \end{aligned} \tag{3.5}$$

This is a straightforward generalization of formula 2.6, which can be easily verified. Application of equation 3.5 leads to simplification of the ratio of probabilities under the logarithm in equation 3.2 as

$$\frac{P(s_1, \ldots, s_{i-1}, s_i', s_{i+1}, \ldots, s_N)}{P(s_1, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_N)} \approx \frac{P(s_{i-1}, s_i') P(s_i', s_{i+1}) P(s_i)}{P(s_{i-1}, s_i) P(s_i, s_{i+1}) P(s_i')}, \tag{3.6}$$

which allows us to perform summation over almost all states $s_1, s_2, \ldots, s_N$ except the few in equations 3.2 to 3.4. This step produces the final expression for the approximate total entropy production rate $\mathrm{EPR}(s_1, \ldots, s_N)$ of our interacting spines:

$$\mathrm{EPR}(s_1, \ldots, s_N) = \sum_{i=1}^{N} \mathrm{EPR}_i, \tag{3.7}$$

where

$$\begin{aligned} \mathrm{EPR}_i \approx \frac{\epsilon}{2} \sum_{s_{i-1}, s_{i+1}} \sum_{s_i, s_i'} \Bigg[& w_{s_i, s_i'}(s_{i-1}, s_{i+1}) \frac{P(s_{i-1}, s_i') P(s_i', s_{i+1})}{P(s_i')} - w_{s_i', s_i}(s_{i-1}, s_{i+1}) \\ & \times \frac{P(s_{i-1}, s_i) P(s_i, s_{i+1})}{P(s_i)} \Bigg] \ln \frac{w_{s_i, s_i'}(s_{i-1}, s_{i+1}) P(s_{i-1}, s_i') P(s_i', s_{i+1}) P(s_i)}{w_{s_i', s_i}(s_{i-1}, s_{i+1}) P(s_{i-1}, s_i) P(s_i, s_{i+1}) P(s_i')} \end{aligned} \tag{3.8}$$

for $i = 2, \ldots, N-1$, and for the boundary terms

$$\begin{aligned} \mathrm{EPR}_1 \approx \frac{\epsilon}{2} \sum_{s_1, s_2} \sum_{s_1'} & \left[w_{s_1, s_1'}(s_2) P(s_1', s_2) - w_{s_1', s_1}(s_2) P(s_1, s_2) \right] \\ & \times \ln \frac{w_{s_1, s_1'}(s_2) P(s_1', s_2)}{w_{s_1', s_1}(s_2) P(s_1, s_2)} \end{aligned} \tag{3.9}$$

and

$$\mathrm{EPR}_N \approx \frac{\epsilon}{2} \sum_{s_{N-1},s_N} \sum_{s'_N} \left[w_{s_N,s'_N}(s_{N-1}) P(s_{N-1}, s'_N) - w_{s'_N,s_N}(s_{N-1}) P(s_{N-1}, s_N) \right] \times \ln \frac{w_{s_N,s'_N}(s_{N-1}) P(s_{N-1}, s'_N)}{w_{s'_N,s_N}(s_{N-1}) P(s_{N-1}, s_N)}. \tag{3.10}$$

Note that the total entropy production of all spines $\mathrm{EPR}(s_1, \ldots, s_N)$ is determined exclusively in terms of the two types of probabilities considered in equations 2.7 and 2.8, that is, one- and two-point probabilities. It is also interesting to mention that the form of the approximated EPR in equation 3.8 can be also deduced instantly from the form of the approximated dynamics for probabilities in equation 2.7. This is possible if one realizes that the expression in the bracket on the right in equation 2.7 represents a probability flux.

The total energy E used by all spines for LTP induction and its maintenance up to recovery (during synaptic stimulation and post stimulation) is the energy needed to keep the memory trace above the threshold. E is the total energy cost of LTP, and it is defined as

$$E = \int_0^{\overline{T}_M} dt \ \mathrm{EPR}(t), \tag{3.11}$$

where $t = 0$ relates to the onset of stimulation, $\mathrm{EPR}(t)$ is the total entropy production rate given by equations 3.7 to 3.10 and $\overline{T}_M = T_M + \tau_1 + \tau_2$, where T_M is the memory time, and $\tau_1 + \tau_2$ is the duration of the stimulation (learning phase; see equation 2.3). The energy used solely for LTP induction and maintenance is denoted as E_{ltp}, and it is given by $E_{ltp} = E - \mathrm{EPR}_0 \overline{T}_M$, where EPR_0 is the baseline entropy production rate of all spines.

4 Definition of Memory Trace and Memory Time

We define the signal associated with dendritic spine activation as

$$S = \frac{1}{N} \sum_{i=1}^{N} s_i. \tag{4.1}$$

The signal at a steady state (baseline) is denoted as S_{ss}. Memory trace MT is defined as the average normalized signal-to-noise ratio. The normalized signal is simply its deviation from the steady state or baseline. Consequently, the memory trace takes the form

$$MT = \frac{(\langle S\rangle - \langle S_{ss}\rangle)}{\sqrt{\langle S^2\rangle - \langle S\rangle^2}}, \tag{4.2}$$

where $\langle S^n\rangle = \sum_{s_1} \dots \sum_{s_N} S^n P(s_1, \dots, s_N)$ for $n = 1, 2$. The explicit forms for the signals and variance of the signal are

$$\langle S\rangle = \frac{1}{N}\sum_{i=1}^{N}\langle s_i\rangle, \tag{4.3}$$

$$\langle S\rangle_{ss} = \frac{1}{N}\sum_{i=1}^{N}\langle s_i\rangle_{ss}, \tag{4.4}$$

$$\langle S^2\rangle - \langle S\rangle^2 = \frac{1}{N^2}\sum_{i=1}^{N}\left[\langle s_i^2\rangle - \langle s_i\rangle^2\right] + \frac{1}{N^2}\sum_{i\neq j}\sum_{j=1}^{N}\left[\langle s_i s_j\rangle - \langle s_i\rangle\langle s_j\rangle\right], \tag{4.5}$$

where $\langle s_i^n\rangle = \sum_{s_i} s_i^n P(s_i)$, $\langle s_i\rangle_{ss} = \sum_{s_i} s_i P_{ss}(s_i)$, and $\langle s_i s_j\rangle = \sum_{s_i}\sum_{s_j} s_i s_j P(s_i, s_j)$, where $P_{ss}(s_i)$ is the probability distribution at baseline state. The first sum on the right in equation 4.5 is the sum of variances of individual spines, while the second sum is the total cross-correlation of spines. In the pair approximation for the probability, the last sum associated with the correlations simplifies, as only the cross-correlations between neighboring spines provide nonzero contributions, since generally $P(s_i, s_j) \approx P(s_i)P(s_j)$ for $|i - j| \geq 2$.

Memory time T_M is defined as the time t after stimulation for which memory trace MT is in a declining phase and assumes value 1 (Fusi et al., 2005; Leibold & Kempter, 2008; Karbowski, 2019). This is the moment in time when a normalized signal becomes comparable to its noise component.

5 Derivation of Information Gain for Interacting Spines

Information gain I, for all spines, right after the end of LTP (i.e., when memory trace decays to the noise level $MT = 1$), is defined as the Kullback-Leibler divergence at time $t = \overline{T}_M$, that is, $\mathrm{KL}(t = \overline{T}_M)$, between the baseline steady-state initial probability $P(s_1, \dots, s_N)_{ss}$ at time $t = 0$ (before LTP stimulation) and final probability at time $t = \overline{T}_M$, that is, $P(s_1, \dots, s_N; t = \overline{T}_M)$. Its form is given by

$$\begin{aligned} I &\equiv \mathrm{KL}\left(P(s_1, \dots, s_N; t = \overline{T}_M)||P(s_1, \dots, s_N)_{ss}\right) \\ &= \sum_{s_1,\dots,s_N} P(s_1, \dots, s_N; t = \overline{T}_M) \ln \frac{P(s_1, \dots, s_N; t = \overline{T}_M)}{P(s_1, \dots, s_N)_{ss}}. \end{aligned} \tag{5.1}$$

For the pair approximation for the probabilities (see equation 3.5), information gain I takes the form

$$I = \sum_{i=1}^{N-2} \sum_{s_i} \sum_{s_{i+1}} P(s_i, s_{i+1}) \ln \frac{P(s_i, s_{i+1}) P_{ss}(s_{i+1})}{P_{ss}(s_i, s_{i+1}) P(s_{i+1})} + \sum_{s_{N-1}} \sum_{s_N} P(s_{N-1}, s_N) \ln \frac{P(s_{N-1}, s_N)}{P_{ss}(s_{N-1}, s_N)}, \tag{5.2}$$

which is used in the computations.

The rate of information gain is equivalent to the rate of Kullback-Leibler (KL) divergence at arbitrary time t, that is, KL(t) (given by equation 5.1, but for arbitrary t). We are interested in the rate of information gain, since we want to compare it directly to the entropy production rate, which has a similar information-theoretic meaning. The rate of KL, which we denote as KLR, is given by KLR $= d\mathrm{KL}(t)/dt$. We find

$$\mathrm{KLR} = \sum_{s_1, \ldots, s_N} \frac{dP(s_1, \ldots, s_N)}{dt} \ln \frac{P(s_1, \ldots, s_N)}{P_{ss}(s_1, \ldots, s_N)}, \tag{5.3}$$

where we used the fact that $\sum_{s_1, \ldots, s_N} dP(s_1, \ldots, s_N)/dt = 0$. The next step is to substitute equation 2.1 for $\frac{dP(s_1, \ldots, s_N)}{dt}$ and to perform summations over almost all states $s_1, \ldots, s_N$, except the few, similar to the calculation for EPR. Finally, we use the pair approximation. As a result, we obtain the total rate of information gain for all spines as

$$\begin{aligned} \mathrm{KLR} = & \sum_{s_1, s_1'} \sum_{s_2} \left[w_{s_1, s_1'}(s_2) P(s_1', s_2) - w_{s_1', s_1}(s_2) P(s_1, s_2) \right] \ln \frac{P(s_1, s_2)}{P_{ss}(s_1, s_2)} \\ & + \sum_{s_N, s_N'} \sum_{s_{N-1}} \left[w_{s_N, s_N'}(s_{N-1}) P(s_{N-1}, s_N') - w_{s_N', s_N}(s_{N-1}) P(s_{N-1}, s_N) \right] \\ & \times \ln \frac{P(s_{N-1}, s_N)}{P_{ss}(s_{N-1}, s_N)} + \sum_{i=2}^{N-1} \sum_{s_i, s_i'} \sum_{s_{i-1}, s_{i+1}} \left[w_{s_i, s_i'}(s_{i-1}, s_{i+1}) \frac{P(s_{i-1}, s_i') P(s_i', s_{i+1})}{P(s_i')} \right. \\ & \left. - w_{s_i', s_i}(s_{i-1}, s_{i+1}) \frac{P(s_{i-1}, s_i) P(s_i, s_{i+1})}{P(s_i)} \right] \ln \frac{P(s_{i-1}, s_i) P(s_i, s_{i+1}) P_{ss}(s_i)}{P_{ss}(s_{i-1}, s_i) P_{ss}(s_i, s_{i+1}) P(s_i)}. \end{aligned} \tag{5.4}$$

As can be seen, the rate of information gain KLR depends, similar to EPR, on the transition rates between states in all spines. In all figures, we plot KLR

per spine: KLR/N. Finally, note that the information gain I is the temporal integral of KLR from $t = 0$ to $t = \overline{T}_M$: $I = \int_0^{\overline{T}_M} dt\ \mathrm{KLR}(t)$.

6 Numerical Results on Memory Trace, Information Gain, and Energy Cost during LTP

We divide the global dynamics of our system of interacting dendritic spines into two stages. The first stage relates to approaching and reaching a steady state, starting from random initial conditions (with $f(t) = 0$ for each spine). This steady state is a thermodynamic nonequilibrium steady state that uses some small but nonzero energy (small but nonzero EPR) and can be thought of as the state in which some background information is written in synapses. After reaching the steady state, we start a second stage associated with spine stimulation, which we call the LTP (long-term potentiation) phase. This stage consists of a brief stimulation of the synaptic system by amplifying the transition rates by the function $f(t)$ present in equations 2.2 and 2.3, and then observation of the system recovery to the steady state, with simultaneous recording of the most important observables. Stimulation of synapses is done by turning on the amplifier function $f(t)$, which amplifies the transition rates between synaptic states. We call this stimulation phase the learning phase and the recovery phase as the memory phase.

6.1 Dynamics of Memory Trace, Information, and Energy Rates, and Synaptic Size Associated with LTP Induction. Following LTP induction (starting at time $t = 0$ in the $f(t)$ function) memory trace MT contained in synapses behaves differently from the rates of information gain (KLR) and energy (EPR) (see Figure 4). Initially all three quantities increase sharply, similar to $f(t)$, but later their dynamics diverge. Specifically, memory trace exhibits a long temporal tail; it decays much more slowly than the stimulation function $f(t)$, with a longer tail for positive cooperation ($\gamma > 0$) between neighboring spines than for negative cooperation (see Figure 4B). Both KLR and EPR (per synapse) decay extremely quickly to their baseline values, much faster than $f(t)$ (see Figures 4C and 4D). This strongly suggests that keeping memory trace high does not require large rates of energy. Moreover, the dynamics of KLR and EPR are very similar in shape, and their ratio is positive only initially, when synaptic stimulation increases in time (see Figure 4E). This behavior indicates that the amount of information written at synapses increases sharply only at the beginning of LTP (learning); at later stages (memory), it weakly decreases (note the negative values of KLR/EPR). The rates of information gain (KLR) and energy (EPR) depend weakly, and in the opposite way, on the sign of cooperativity γ; the peak of KLR is greater for negative γ, while the peak of EPR is greater for positive γ (see Figures 4C and 4D). At its peak, during stimulation, energy is consumed at rate $4.6 \cdot 10^5$ kT/min per spine, and information is gained

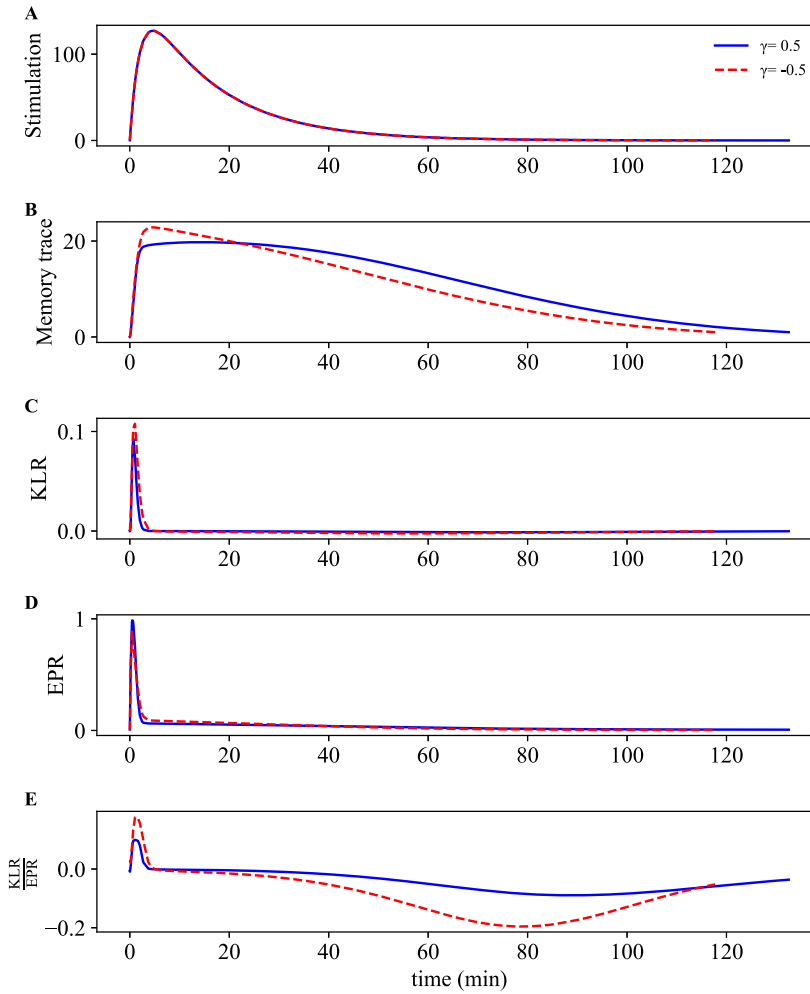

Figure 4: Dynamics of memory trace and the rates of information gain and energy. Temporal dependence of (A) stimulation function $f(t)$ (amplifier of transitions between the states), (B) memory trace MT, (C) information gain rate per spine (KLR/N), (D) entropy production rate (energy rate) per spine (EPR/N in units of ϵ/min), and of (E) the ratio of information gain rate to entropy production rate per spine (KLR/EPR). The learning phase, equivalent to stimulation phase, lasts up to 20 min (A). The memory phase, quantified by memory trace, starts after the end of stimulation and lasts up to about 120 minutes (B). Note that with stimulation, the rates of information gain and entropy production (KLR and EPR) achieve peaks quickly, but they also decay quickly. In contrast, memory trace lasts much longer. The EPR value at its peak is $\sim\epsilon$/min $\approx 4.6 \cdot 10^5$ kT/min per spine, which is about two to three orders of magnitude larger than at baseline (before or long after the stimulation).

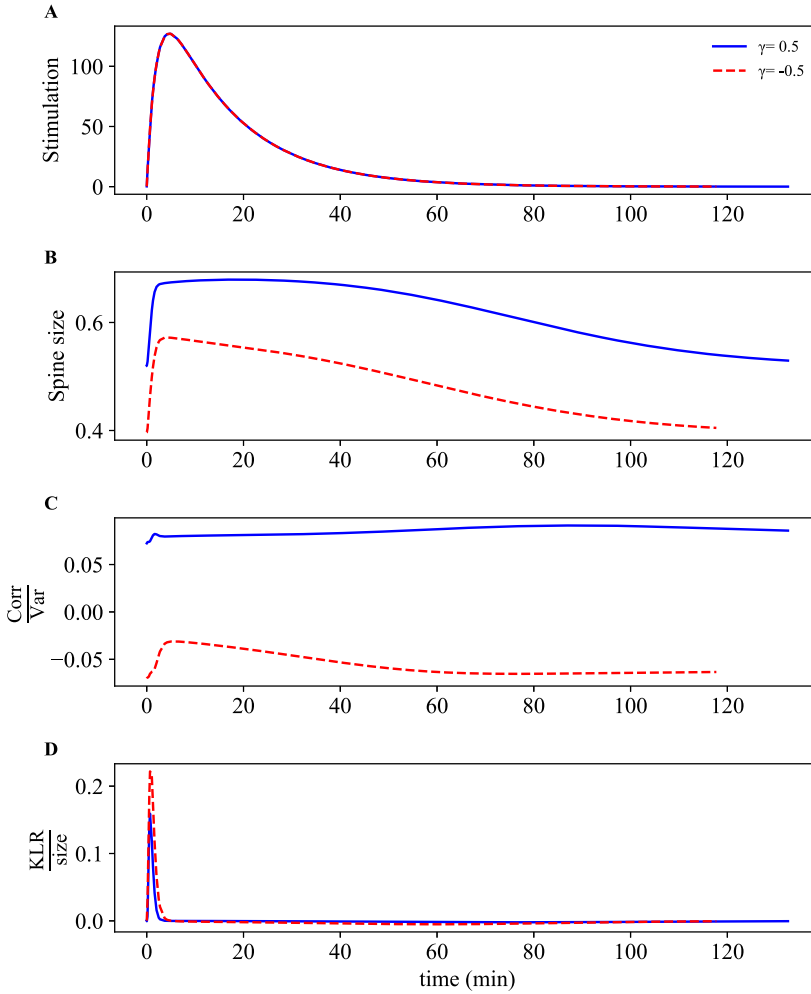

Figure 5: Dynamics of spine size, correlations, and information gain per spine size. Temporal dependence of (A) stimulation function $f(t)$, (B) average spine size $\langle S \rangle$, (C) normalized correlations between spines, and (D) the ratio of KLR per spine to average spine size $\langle S \rangle$.

at rate 0.1 bits/min per spine, which indicates that acquiring 1 bit at that moment in time is very expensive and costs about $4.6 \cdot 10^6$ kT per spine.

The dynamics of spine sizes following LTP induction are similar to the behavior of memory trace, except that sizes stabilize at some finite level (see Figure 5). Positive cooperativity among synapses ($\gamma > 0$) generally indicates positive correlations between them, and vice versa (see Figure 5C).

Consequently, positive correlations lead to higher mean spine sizes than negative correlations (see Figure 5B). An opposite effect is seen for the ratio of KLR and spine size; higher peaks are observed for negative cooperativity between synapses (see Figure 5D). Interestingly, normalized correlations are more variable during learning and memory phases for negative cooperativity than for positive one (see Figure 5C).

6.2 Memory Time, Spine Size, Information Gain after LTP, and Their Energy Costs as Functions of Synaptic Cooperativity. Memory time T_M grows monotonically but very weakly with cooperativity γ up to a point where γ is close to its maximal value 1 (see Figure 6A). In that regime of very high positive cooperativity, T_M increases sharply with γ. The opposite dependence on γ is present for information gain I and average spine size (see Figures 6B and 6E). I generally decreases with γ (for negative γ, the decay is stronger than for positive γ; see Figure 6B), whereas mean spine size monotonically increases with γ (the increase is stronger for γ close to 1; see Figure 6E). Total energy E consumed during LTP and its part E_{ltp} related solely to LTP depends nonmonotonically on cooperativity γ, exhibiting broad maxima for $\gamma = 0$ (see Figures 6C and 6D). For γ close to its maximal value 1, E and E_{ltp} behave in opposite ways: the former increases while the latter decreases with γ. This suggests that the cost of LTP alone drops for very high cooperativity.

How do these results translate to energy and structural efficiency of memory lifetime and information gain? Figure 7 provides the answers. The ratios of memory time and the energies, T_M/E and T_M/E_{ltp}, are essentially constant for γ up to $\sim$0.8 (see Figure 7A), suggesting that memory time and both of these energies grow proportionally with cooperativity for a wide range of γ. For larger γ, these ratios grow significantly with γ, especially T_M/E_{ltp}, indicating that energy efficiency of memory time is enhanced in the regime of very high, positive synaptic cooperativity (see Figure 7A).

Energy efficiency of information gain is more complex (see Figure 7B). Generally the ratio of information to total energy during LTP I/E decreases monotonically with γ, meaning that the efficiency of I is maximal for negative cooperativity between synapses. On the other hand, the ratio of information to energy solely to LTP, that is, I/E_{ltp} as a function of γ, has a U-shape, with large values for both high negative and high positive cooperativity. The latter means that information efficiency in that energy currency has two regimes of higher values (see Figure 7B). However, it should be emphasized that the overall energy efficiency of information gain is rather low, at $(5-10)\cdot 10^{-4}$ bits/ϵ or $(1-2)\cdot 10^{-9}$ bits/kT, that is, 1 bit of stored memory (after the decline of LTP, that is, after time $\sim T_M$) costs about 10^9 kT (see also below).

Structural efficiency (or energy efficiency of transmission) of memory time and information gain exhibit different behavior as functions of synaptic cooperativity (see Figures 7C and 7D). The ratio of information gain to

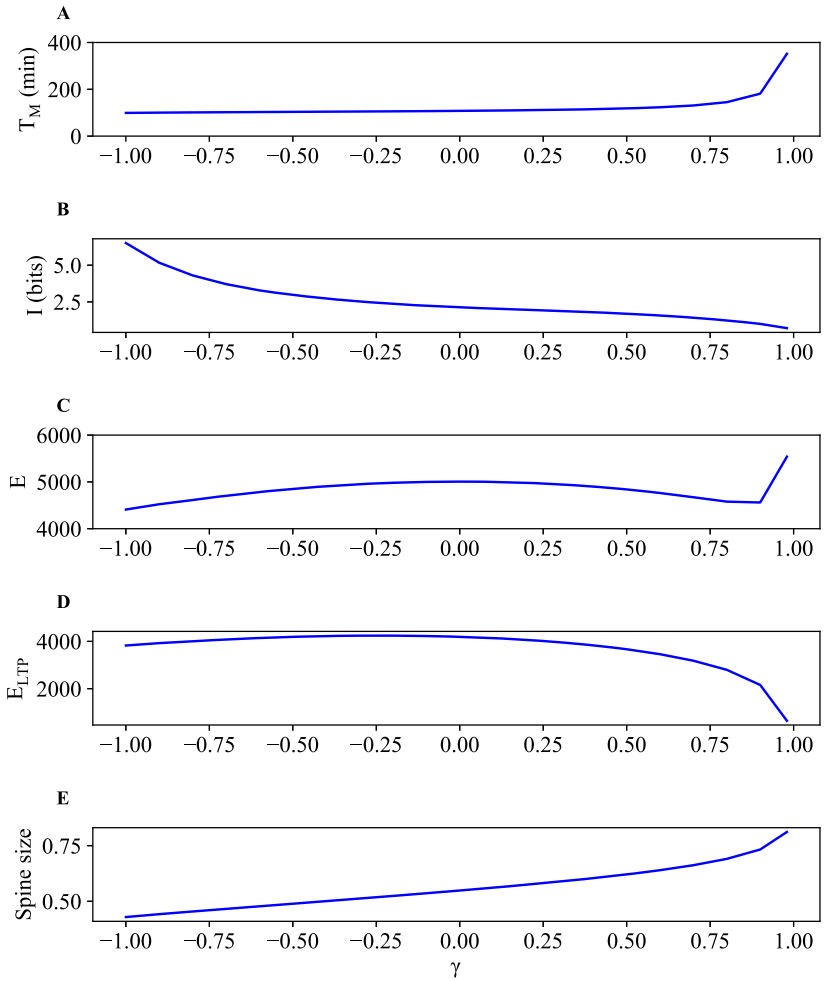

Figure 6: Dependence of memory time, information gain, energy cost, and spine size on the synaptic cooperativity. (A, B) Monotonic but opposite dependence of memory time T_M and total information gain I on cooperativity γ. (C, D) Total energy consumption of all spines E during LTP and its part E_{ltp}, related solely to LTP induction and maintenance, exhibit nonmonotonic dependence on γ (energy units are in ϵ). During the full LTP, a typical spine used about $5\epsilon \approx 2.3 \cdot 10^6$ kT of energy (all $N = 1000$ spines used $5000\epsilon \approx 2.3 \cdot 10^9$ kT). (E) Average spine size increases monotonically with γ.

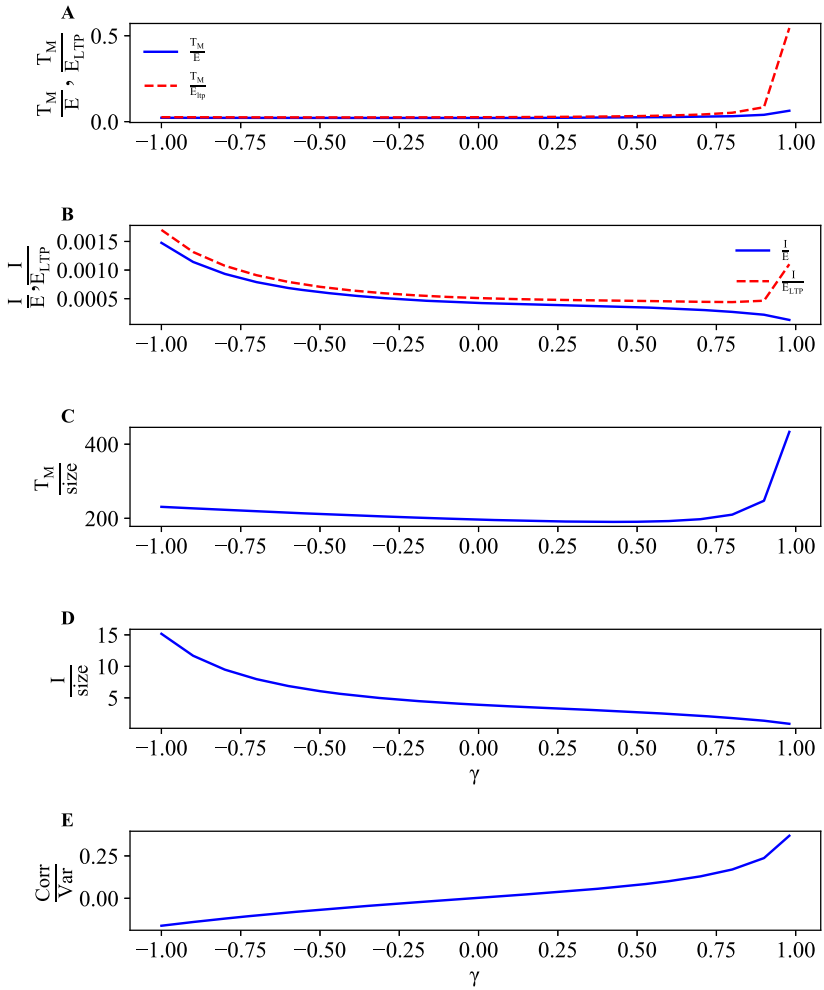

Figure 7: Energetic and structural efficiencies of memory time and information gain in comparison to spine correlations. Ratios of (A) memory time T_M and (B) information gain I to two energies E and E_{ltp} exhibit two opposite dependence. Similar behavior for the ratios of (C) memory time and (D) information gain to average spine size $\langle S \rangle$ as functions of γ. (E) Normalized correlations always increase with γ, reaching $\sim$0.3 for $\gamma \mapsto 1$.

mean spine size $I/\langle S \rangle$ decreases monotonically with γ, which indicates that structural efficiency of information is the largest for negative cooperativity (see Figure 7D), similar to the (total) energy efficiency of I (see Figure 7B). The ratio of memory time to mean spine size $T_M/\langle S \rangle$ decreases slightly with increasing cooperativity from negative values of γ, but for $\gamma \approx 0.4$, this ratio

increases sharply with γ (see Figure 7C). This result shows that the structural efficiency of memory time is the highest for strong, positive synaptic cooperativity (see Figure 7C), where spine normalized correlations reach values of 0.2 to 0.3 (see Figure 7E).

6.3 Sparse Representations of Synaptic Memory and Information Are More Energy Efficient. Next we investigate how memory time, information gain, average spine size, and energy cost depend on the fraction p_{act} of stimulated synapses by presynaptic neurons (see Figure 8). Memory time T_M, energy cost solely due to LTP E_{ltp}, and mean spine size grow monotonically with p_{act}, though the first one saturates for larger p_{act} (see Figures 8A, 8C, and 8D). In contrast, information gain I and total energy cost E display nonmonotonic behavior: the former has a maximum, while the latter a minimum for a small fraction of active synapses (see Figures 8B and 8C). In terms of energy efficiency, the ratios of T_M/E, T_M/E_{ltp} and I/E, I/E_{ltp} have maxima at around the same small fraction p_{act}, regardless of the sign of synaptic cooperativity γ (see Figures 9A and 9B). This means that there exist, an optimal percentage of activated synapses on a dendrite that yields the highest information gain and memory lifetime per invested energy (whether total or only due to LTP). For that optimal p_{act}, the energy cost of 1 bit of stored information is about 10^7 kT, which is much lower (and thus more efficient) than for values p_{act} away from the optimality. Interestingly, the normalized correlations between spines are essentially independent of p_{act} (see Figure 9E).

Structural efficiency of information gain and memory time is qualitatively similar to their energy efficiency (see Figure 9). The information per spine size $I/\langle S\rangle$ has a similar sharp peak as I/E for low p_{act} (see Figure 9D). However, memory time per spine size $T_M/\langle S\rangle$ has a much broader maximum at higher values of p_{act} (see Figure 9C).

Taken together, these results indicate that energetic and structural efficiency of information and its duration in synapses can be optimized for low fractions of activated synapses during LTP. In other words, acquiring and storing of synaptic information can be most efficient by using sparse synaptic representations, regardless of the nature of synaptic cooperativity.

6.4 Memory Time, Spine Size, Information Gain, and Their Energy Costs as Functions of Strength and Duration of Stimulation. In Figures 10 and 11, we show how the main observables depend on the duration of stimulation τ_1 (decay time of the stimulation). Memory time T_M and its energy costs E, E_{ltp} both grow proportionally with τ_1 (see Figures 10A and 10C), such that their ratios T_M/E and T_M/E_{ltp} are almost constant, although with a weak increasing trend (see Figure 11A). Information gain I as well as its energy efficiencies I/E, I/E_{ltp} decrease with τ_1 for small τ_1, but for larger τ_1, all of these quantities saturate at some small level (see Figures 10B and 11B). Average spine size shows a similar saturation effect after a small

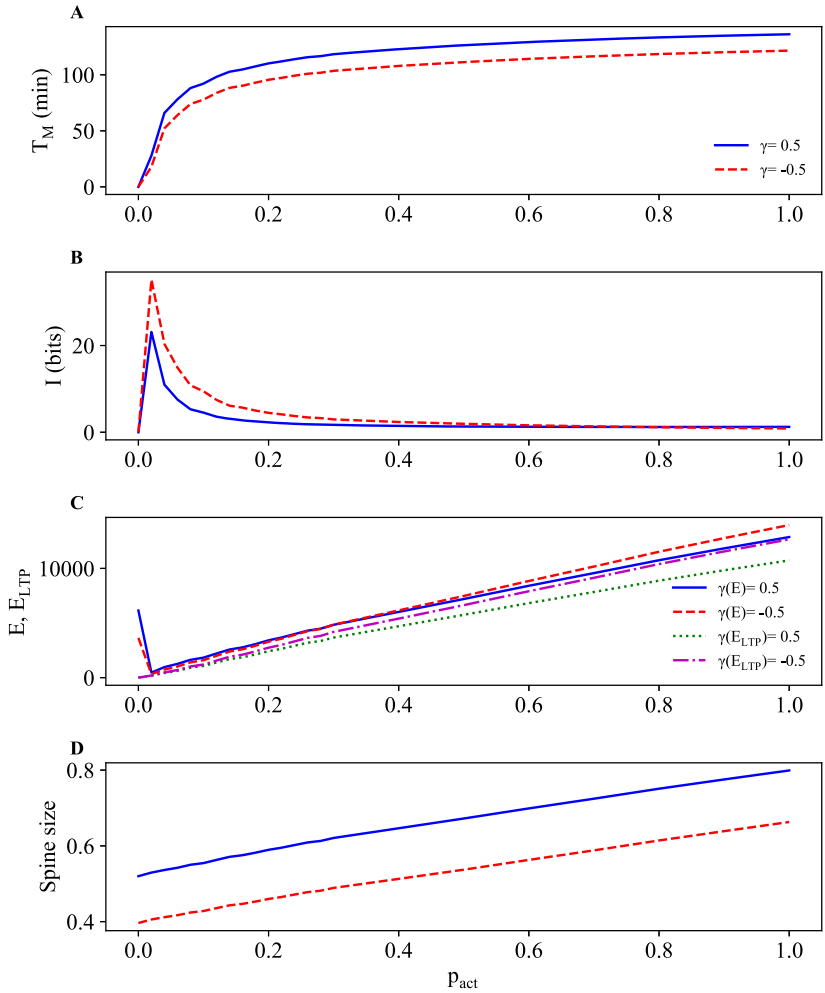

Figure 8: Dependence of memory time, information gain, energy cost, and spine size on probability of synaptic stimulation. (A) Memory lifetime T_M saturates for large p_{act}. (B) Information gain exhibits a sharp peak for very small p_{act}. (C, D) Energy solely due to LTP and average spine size increases linearly with p_{act}.

initial increase (see Figure 10D). In terms of structural efficiency, $T_M/\langle S\rangle$ grows linearly with τ_1 (see Figure 11C), and $I/\langle S\rangle$ first decreases with τ_1 and then saturates (see Figure 11D). The results in Figure 11 indicate that longer stimulation times are not particularly beneficial for energy efficiency of memory lifetime and information gain (see Figures 11A and 11B).

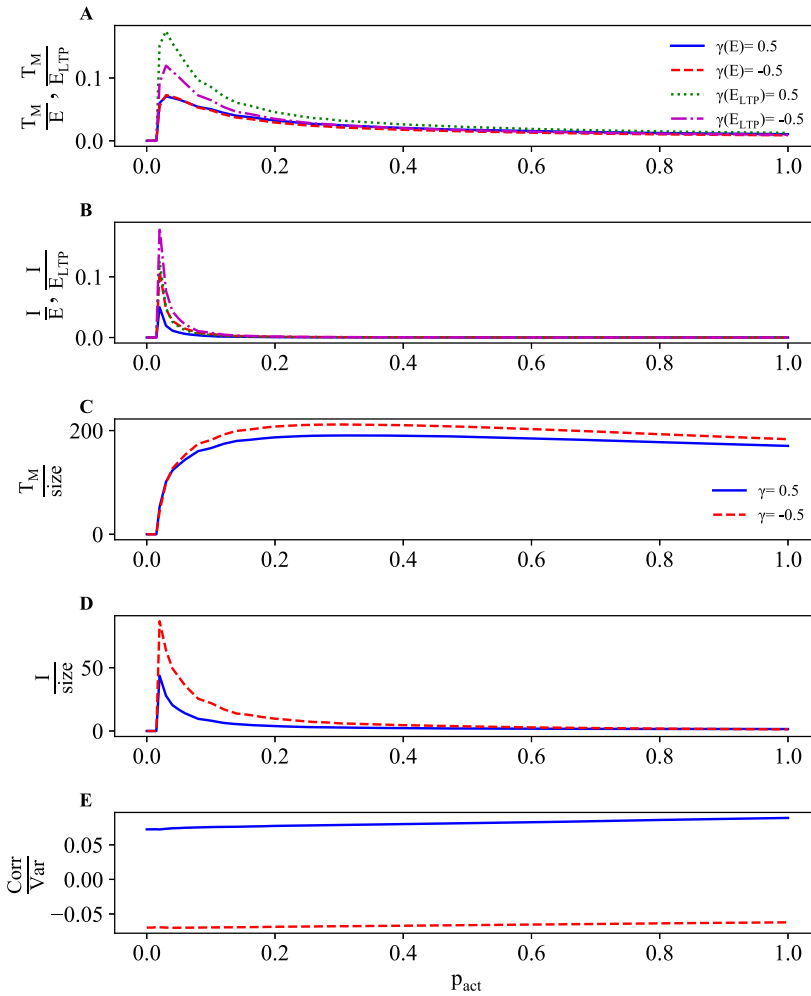

Figure 9: Energetic and structural efficiencies of memory time and information gain as functions of probability of synaptic stimulation (A–D). All the ratios of memory time and information gain to energies and to spine size $\langle S \rangle$ exhibit maxima. At the peak, the ratio I/E is $\sim$0.05 bit/ϵ and I/E_{ltp} is $\sim$0.15 bit/ϵ, which means that 1 bit of stored information after LTP degradation requires about 10^7 kT of energy (A–E).

However, longer stimulation may be advantageous for structural efficiency of memory duration (see Figure 11C), though not for information I (see Figure 11D).

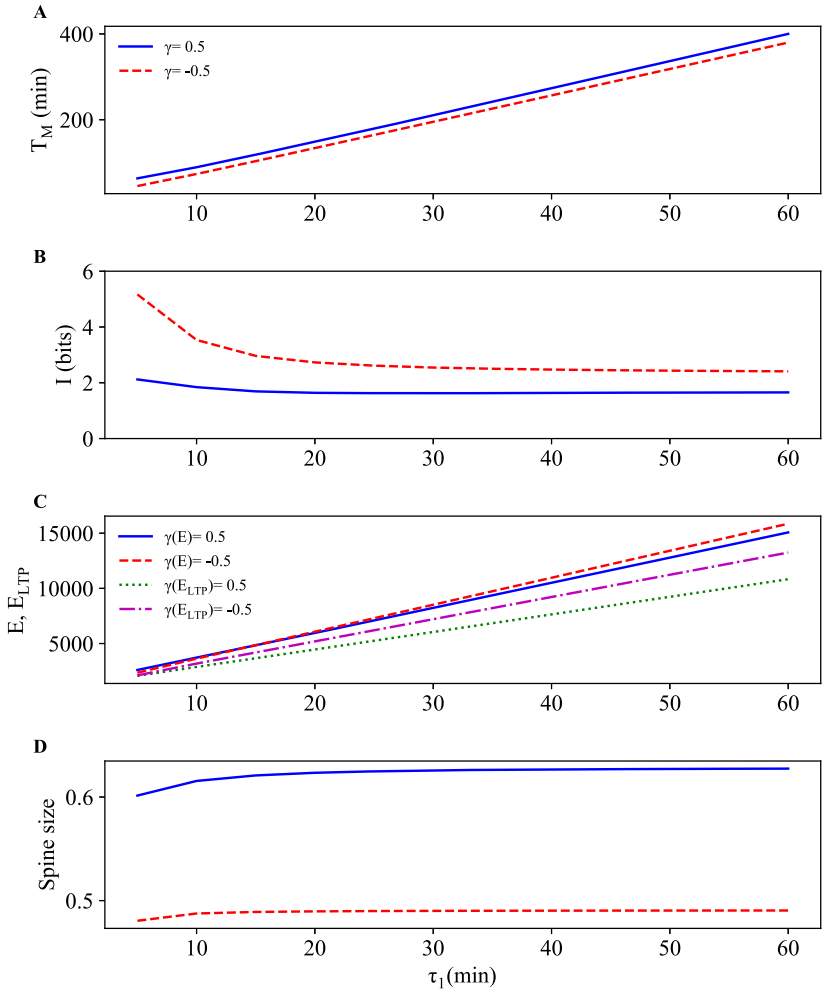

Figure 10: Dependence of memory time, information gain, energy cost, and spine size on duration of stimulation τ_1 (A–D).

In Figures 12 and 13, we present the dependence of memory time, information gain, energy cost, and mean spine size on the amplitude of stimulation A. Memory time T_M and energy costs both grow weakly but saturate with A (see Figures 12A and 12C). On the other hand, I and spine size $\langle S \rangle$ stay almost constant (see Figures 12B and 12D). These results translate into very weak variability of energy and structural efficiencies of memory time and information gain on A, which are close to constancy, except the ratio $T_M/\langle S \rangle$ that exhibits an increasing trend but with a saturation

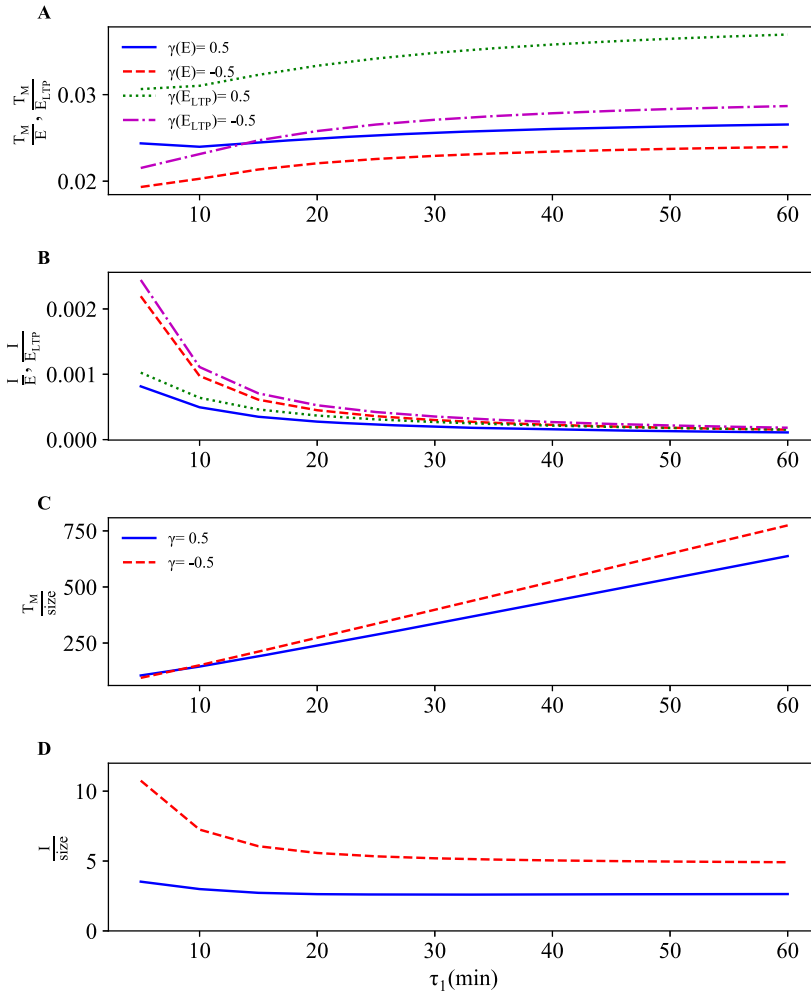

Figure 11: Energetic and structural efficiencies of memory time and information gain as functions of duration of stimulation τ_1 (A–D). Note that energetic efficiency of information gain drops with increasing the duration of stimulation (B). Structural efficiency of memory lifetime grows linearly with τ_1 (C).

(see Figure 13). This suggests that stimulations that are too strong are also not advantageous over weaker stimulations for the efficiency of information gain and its duration.

6.5 Influence of Synaptic Number on the Efficiencies of Memory Lifetime and Information Gain. In Figure 14, we show that energy efficiency of

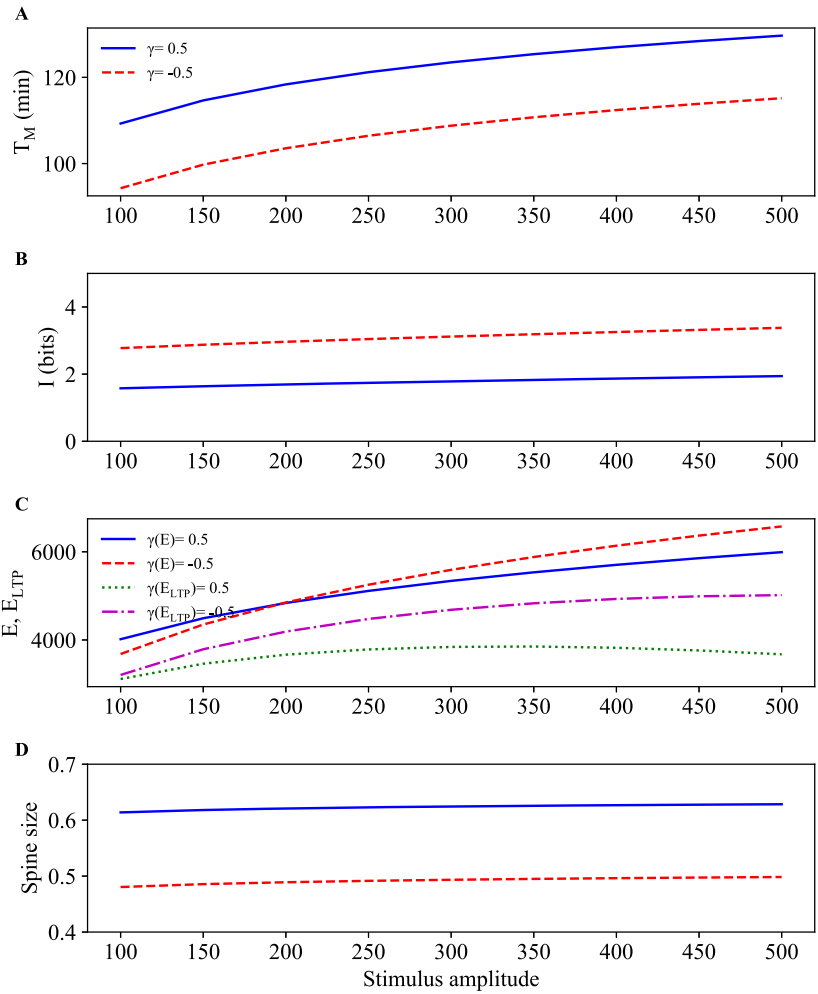

Figure 12: Dependence of memory time, information gain, energy cost, and spine size on strength of stimulation A (A–D). Note that information gain (B) and mean spine size (D) are essentially independent of A.

both memory lifetime T_M and information gain I exhibit a decreasing trend with increasing the number of dendritic spines N. The biggest drop in efficiency is for changing N from 10 to approximately 2000 (see Figures 14A and 14B). For higher values of N, the rate of decline is much slower. This result suggests that having a large number of spines on a dendrite is generally highly inefficient in terms of energy for storing information. For example, for $N = 10$, we obtain $T_M/E \approx 1$ min$/\epsilon$ and $I/E \approx 4 \cdot 10^{-2}$ bits$/\epsilon$ (i.e.,

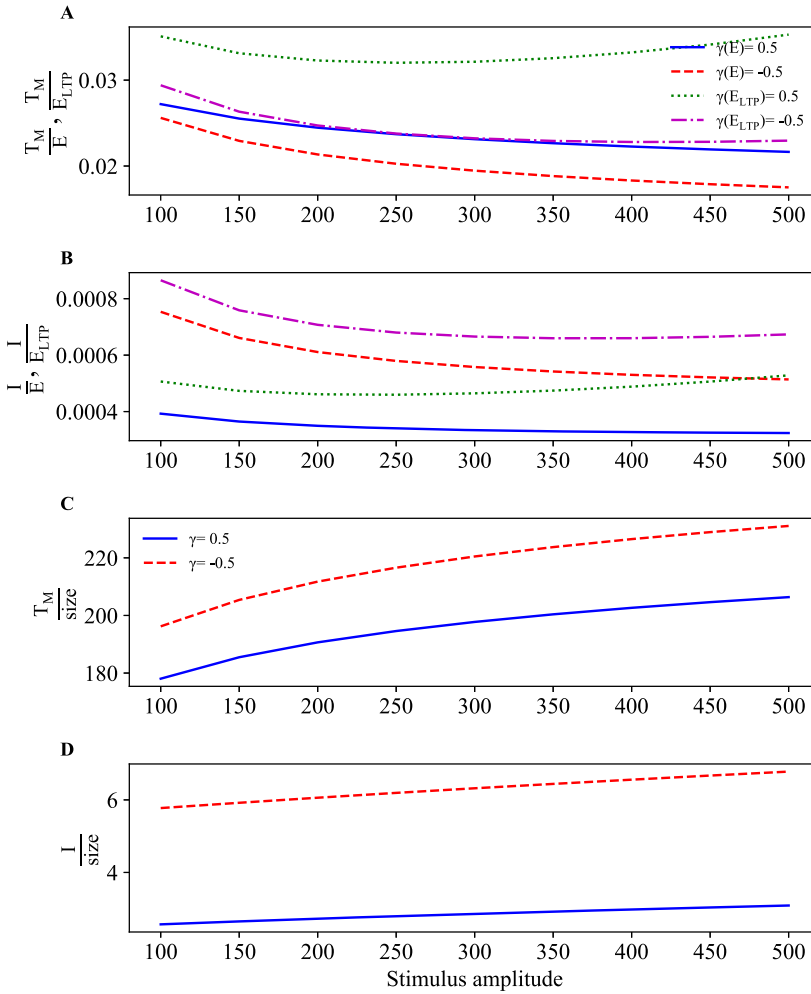

Figure 13: Energetic and structural efficiencies of memory time and information gain as functions of strength of stimulation A (A–D). Note that stimulation amplitudes that are too large are not beneficial for the energy and structural efficiencies (saturation effects).

storing 1 minute of memory in the spines costs about $\epsilon \approx 4.6 \cdot 10^5$ kT of energy, and storing 1 bit after the degradation of LTP costs $25\epsilon \approx 10^7$ kT. In contrast, for $N = (8 - 10) \cdot 10^3$, we have $T_M/E \approx 2 \cdot 10^{-3}$ min/ϵ and $I/E \approx 10^{-4}$ bits/ϵ, which means that in this case, storing 1 minute of memory costs $\sim 500\epsilon \approx 2 \cdot 10^8$ kT and storing 1 bit in all N spines after LTP degradation costs $10^4\epsilon \approx 4.6 \cdot 10^9$ kT.

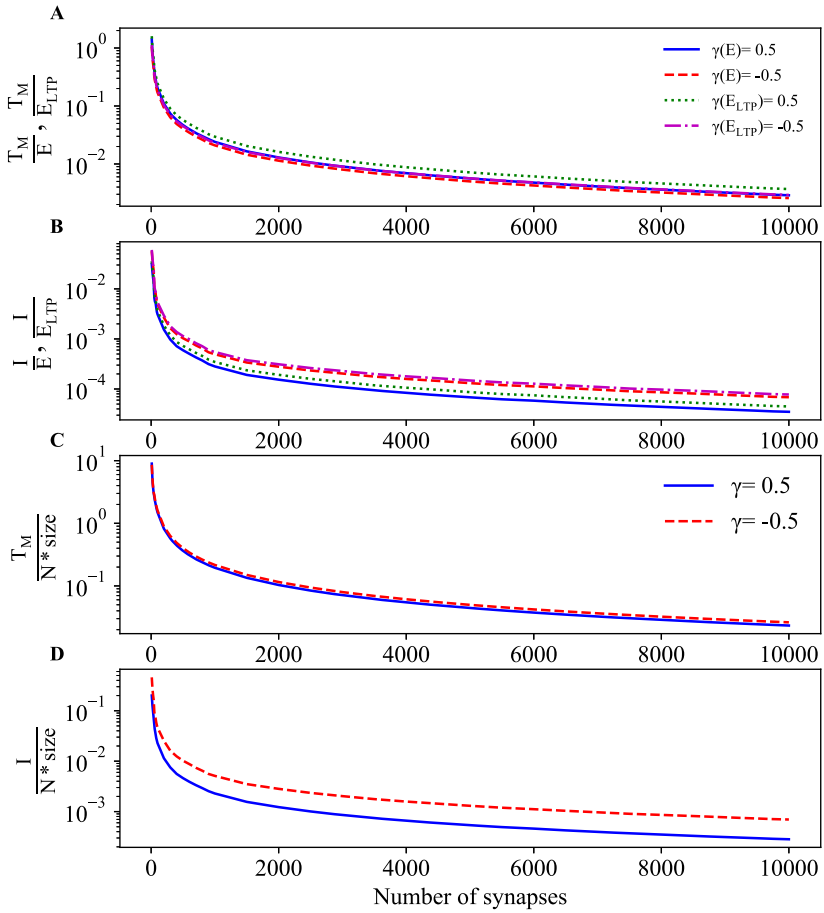

Figure 14: Energy and structural efficiencies of memory lifetime and information gain drop with increasing synaptic numbers N. (A, B) Energy efficiency of T_M and I decrease dramatically, by two orders of magnitude, with increasing the number of spines from $N = 10$ to $N = 2000$, and much slower with increasing the number of spines from $N = 2000$ to $N = 10{,}000$. (C, D) Essentially the same declining effect is observed for the structural efficiency of T_M and I, that is, $T_M/(N\langle S\rangle)$ and $I/(N\langle S\rangle)$.

Since memory lifetime and information gain are collective variables, to determine structural efficiencies of these variables as functions of N, we have to divide T_M and I by the whole structural cost of all spines, that is, by $N\langle S\rangle$. Thus, the structural efficiencies in this case are the ratios $T_M/(N\langle S\rangle)$ and $I/(N\langle S\rangle)$ (see Figures 14C and 14D), and they drop significantly with

increasing N, in a similar manner as the corresponding energy efficiencies (in Figures 14A and 14B).

Taken the two types of the efficiency together, it is clear that having too many synapses on a dendrite is not beneficial for information storage and its duration in terms of energy and biochemical resources.

7 Summary and Discussion

We formulated the probabilistic approach to global dynamics of interacting synapses (on a single dendrite) and to their nonequilibrium thermodynamics to study information processing in synaptic internal degrees of freedom. In order to make the high-dimensional system of interacting synapses computationally tractable, we introduced the so-called pair approximation, which effectively reduces the dimensionality and number of equations describing the system dynamics in a closed form by considering only the probabilities of singlets and doublets of dendritic spines. We verified on a simplified example that this approximation provides very good accuracy of the exact dynamics, as well as entropy production rate. We also gave the analytical condition for the applicability of the pair approximation in the equilibrium Ising model in appendix B (for a more general discussion about the accuracy of the pair approximation, see Matsuda et al., 1992, and Van Baalen, 2000).

The master equation approach combined with stochastic thermodynamics allows us to treat information contained in synaptic states on equal footing with its energy cost by relating them to, respectively, the rates of Kullback-Leibler divergence and entropy production. Both of these quantities depend on state probabilities as well as on transitions between the states, and we provided explicit formulas for their calculations. The formalism makes it clear that every plastic transition in a synaptic system is associated with some information processing (flow) and some energy expenditure (entropy production). Even stationary states out of thermodynamic equilibrium (baseline states) require some energy (usually small) in order to maintain them, which essentially means that keeping information always costs some energy for nonequilibrium systems even in stationary conditions. Physically, this energy dissipation (entropy production) in a stationary state is a consequence of breaking the so-called detailed balance in probability flows, and it is a generic feature of systems out of thermodynamic equilibrium (in the context of physics, see Maes et al., 2000, and Seifert, 2012 and for a dendritic spine, see Karbowski, 2019).

Our main results are, first, that the learning phase of a signal (spine stimulations) involves high levels of both information and energy rates, which are much larger than their values during a memory phase (see Figure 4). This indicates that keeping information (memory) is relatively cheap in comparison to acquiring it (learning). This result on a level of many interacting synapses is qualitatively similar to the result in Karbowski (2019),

where it was shown that a memory trace resulting from molecular transitions in a single synapse (protein phosphorylation) decouples from the energy rate after the stimulation phase, leading to a relatively cheap long memory trace of protein configurations. In our model, the maximal energy rate of spine plasticity taking place during stimulation is about $4.6 \cdot 10^5$ kT/min (for $p_{act} = 0.3$, but it depends on A). Second, memory lifetime and its energy efficiency can significantly increase their values for very strong, positive synaptic cooperativity, while the opposite is observed for information gain right after LTP drops to its noise level (see Figure 7). This result suggests that strong, local positive correlations between neighboring spines can be beneficial for memory storage, especially in the range of 0.3 (see Figure 7E), that is, for values reported experimentally (Makino & Malinow, 2011). This conclusion supports the so-called synaptic clustering hypothesis, which was proposed as a mechanism for producing synaptic memory (Govindarajan et al., 2006) and enhancing its capacity (Poirazi & Mel, 2001; Kastellakis & Poirazi, 2019). Third, there exists an optimal fraction of stimulated synapses during LTP for which energy efficiency of both memory lifetime and information gains exhibit maxima (see Figure 9). This means that sparse representations of learning and memory are much better in terms of energy efficiency, and thus might be preferable by actual synapses. This is also true for the structural efficiency of information gain (see Figure 9). Fourth, energy and structural efficiencies of memory lifetime and information gain after degradation of LTP, both drop dramatically with increasing the number of spines (see Figure 14). For example, storing 1 bit after LTP is over costs "only" about 10^7 kT for 10 spines and at least two orders of magnitude more: $\sim 4.6 \cdot 10^9$ kT for $\sim 10^4$ spines. In terms of the cost per one spine, these numbers are comparable to the energy cost of transmitting 1 bit through a chemical synapse (Laughlin et al., 1998).

Our model is based on the assumption that a dendritic spine can be treated as the system with discrete states, which is compatible with some morphological observations (Bourne & Harris, 2008; Montgomery & Madison, 2004; Bokota et al., 2016; Urban et al., 2020). In this respect, it is similar in architecture to some previous discrete models of synapses or dendritic spines (Fusi et al., 2005; Leibold & Kempter, 2008; Barrett et al., 2009; Benna & Fusi, 2016). However, these models treat synaptic states quite arbitrarily and abstractly, and consider mostly unidirectional transitions between the states, which makes these models thermodynamically inconsistent (e.g., entropy production rate, equivalent to plasticity energy rate, is ill defined and yields infinities for unidirectional transitions). In contrast, our model takes as a basis well-defined morphological synaptic states, with bidirectional transitions between them that are estimated based on empirical data (Bokota et al., 2016; Basu et al., 2018; Urban et al., 2020). The latter feature, bidirectional transitions, makes our model thermodynamically consistent (with finite entropy production), as explained in a previous model

of metabolic molecular activity in a single spine within the framework of cascade models of learning and memory (Karbowski, 2019).

In this letter, we consider two types of costs. The first is energy cost (related to entropy production) associated with stochastic "plastic" transitions between different synaptic states. The second is structural cost, defined here as proportional to average spine size and related to the biochemical cost of building a synapse. This structural cost is also proportional to the rate of electric energy of synaptic transmission (Attwell & Laughlin, 2001; Karbowski, 2009, 2012), as spine size is proportional to synaptic electric conductance or, more commonly, synaptic weight (Kasai et al., 2003). For standard cortical conditions (i.e. for low firing rates $\sim$1 Hz), the energy cost of synaptic transmission is much larger than the energy cost related to plasticity processes inside the spine (Karbowski, 2019). However, these two costs can become comparable for very large firing rates of about $\sim$100 Hz (Karbowski, 2021). There is some confusion in the literature about these two types of energy costs, and some researchers associate the structural cost related to synapse size or weight with "plasticity energy cost," by assuming that the larger synaptic weight leads the higher plastic energy cost (Li & van Rossum, 2020). However, this does not have to be so, and we should make a distinction between the two energy costs. Bigger synapses do not necessarily require larger amounts of plasticity-related energy than smaller synapses because the transitions in bigger synapses could be generally much slower than in smaller synapses (as is in fact reported in some experiments; Kasai et al., 2003). Thus, what mostly matters for the plasticity energy rate are the rates of transition between internal synaptic states. On the other hand, synaptic size or weight is always a good indicator of electric energy rate related to fast synaptic transmission (Attwell & Laughlin, 2001; Karbowski, 2009, 2012).

The model can be extended in several ways, for example, by introducing heterogeneity in spine interactions, that is, by allowing random signs of the cooperativity parameter γ. However, we suspect that such modifications would not alter the general qualitative conclusions. Finally, our model considers only the early phase of LTP, the so-called e-LTP, which generally lasts up to a few hours and does not involve protein synthesis inside spines. Inclusion of protein synthesis, associated with the process of memory consolidation and thus the late phase of LTP (so-called l-LTP), would require some modifications in our model, the most important of which is inclusion of additional variables in the probabilities characterizing spine states, which are related to internal degrees of freedom (e.g., proteins, actin). This clearly would make the model much more complex, and thus it remains a major challenge at this time. However, we hope that our approach of stochastic thermodynamics provides some insight into attempting to model the interplay of information and energy during the late phase of LTP and memory consolidation for interacting spines.

Appendix A: Stochastic Model of Morphological States in Dendritic Spines

Our data on dendritic spines come from cultures of rat hippocampus (Bokota et al., 2016; Basu et al., 2018). We assume that each dendritic spine can be in four different morphological states: nonexistent (lack of spine), stubby, filopodia, and mushroom. These four states constitute the minimal number of states that can be classified and quantified on a mesoscopic level (Bokota et al., 2016; Basu et al., 2018; Urban et al., 2020). Each state has a typical size, which can be characterized by several geometric parameters. We focus on one particular parameter, spine head surface area, as an indicator of both spine structure and function. Spine head surface area is proportional to synaptic weight (as measured by the number of AMPA receptors; Kasai et al., 2003), which relates to spine neurophysiological function (synaptic transmission and information storage in molecular structure). Spine area is also a measure of its structural and metabolic (synaptic transmission) costs (a larger area means larger both costs). To estimate spine areas in each of the three states (for the nonexistent state, the size is 0), we used the data on minimal and maximal spine head diameters from Bokota et al. (2016) and Urban et al. (2020), which gave us the following numbers: $d(0) = 0$ for nonexistent, $d(1) = 0.496\ \mu\text{m}^2$ for stubby, $d(2) = 0.786\ \mu\text{m}^2$ for filopodia/thin, and $d(3) = 1.045\ \mu\text{m}^2$ for mushroom. Values of the intrinsic transition matrix are given in Table 1. The values of global parameters are presented in Table 2.

We assume that global spine dynamics can be described as a Markov chain model, in which there are stochastic transitions between spine internal states. The general model of this kind is given by the following master equation (Glauber, 1963):

$$\frac{dP(s_1,\ldots,s_N)}{dt} = \sum_{s'_1,\ldots,s'_N} W(s_1,\ldots,s_N|s'_1,\ldots,s'_N)P(s'_1,\ldots,s'_N) - P(s_1,\ldots,s_N)\sum_{s'_1,\ldots,s'_N} W(s'_1,\ldots,s'_N|s_1,\ldots,s_N), \quad \text{(A.1)}$$

where $W(s_1,\ldots,s_N|s'_1,\ldots,s'_N)$ is the multidimensional transition matrix of the whole system of N dendritic spines. We assume that transitions between the states take place in only one of the spines at any given time unit (the rest of states in other spines do not change in that brief time step). This means that the multidimensional matrix $W(s_1,\ldots,s_N|s'_1,\ldots,s'_N)$ can be decomposed as

$$\begin{aligned} W(s_1,\ldots,s_N|s'_1,\ldots,s'_N) = {} & w_{s_1,s'_1}(s_2)\delta_{s_2s'_2}\ldots\delta_{s_Ns'_N} + \ldots \\ & + w_{s_i,s'_i}(s_{i-1},s_{i+1})\delta_{s_1s'_1}\ldots\delta_{s_{i-1}s'_{i-1}}\delta_{s_{i+1}s'_{i+1}}\ldots\delta_{s_Ns'_N} + \ldots \\ & + w_{s_N,s'_N}(s_{N-1})\delta_{s_1s'_1}\ldots\delta_{s_{N-1}s'_{N-1}}, \end{aligned} \quad \text{(A.2)}$$

where $w_{s_i,s_i'}(s_{i-1}, s_{i+1})$ are the transition rates at individual spines, which are dependent on the neighboring spines. After insertion of the form of multidimensional transition matrix W in equation A.1 and performing summations, we obtain equation 2.1 in the main text.

By using the above pattern of transition probabilities, we assume that at a sufficiently short time step, only one transition can take place; simultaneous transitions in different spines are much less likely, and thus can be neglected. Indeed, since the local basic transitions between mesoscopic states in individual spines are of the order of several minutes (Urban et al., 2020; see Table 1), the likelihood that two or more such transitions in two or more spines take place simultaneously in a short period of time (much smaller than a minute) is small. This type of locality of explicit synaptic interactions allows us to analyze the dynamics of global system of N interacting spines.

Appendix B: Validity of the Pair Approximation for Analytically Solvable Model

In this section, we provide conditions that must be satisfied for applying the pair approximation in a case that can be treated analytically, which is a simplified Ising model in thermal equilibrium. We consider three interacting units ($i = 1, 2, 3$) forming a linear ordered chain, similar as in Figure 1, but each unit having only two states, $s_i = -1$ or $s_i = 1$. Additionally, in this model, nearest neighbors interact strongly with the coupling J, whereas remote units (1 and 3) interact weakly with the coupling κ, which is much smaller than J. Our goal is to check how accurate the pair approximation is as we increase the strength of remote coupling κ in relation to J.

The equilibrium probability of finding a given configuration of units s_1, s_2, s_3 has the form (Feynman, 1972)

$$P(s_1, s_2, s_3) = Z^{-1} e^{-J(s_1 s_2 + s_2 s_3) - \kappa s_1 s_3}, \tag{B.1}$$

where Z^{-1} is the normalization factor. The two-point marginal probabilities are given by

$$P(s_1, s_2) = 2Z^{-1} e^{-J s_1 s_2} \cosh(J s_2 + \kappa s_1), \tag{B.2}$$

$$P(s_2, s_3) = 2Z^{-1} e^{-J s_2 s_3} \cosh(J s_2 + \kappa s_3). \tag{B.3}$$

The one-point marginal probability for the middle unit is

$$P(s_2) = 2Z^{-1}\left[e^{J s_2} \cosh(J s_2 - \kappa) + e^{-J s_2} \cosh(J s_2 + \kappa)\right]. \tag{B.4}$$

These four probabilities are all we need to quantify the accuracy of the pair approximation, which in our case is represented by $P(s_1, s_2, s_3) \approx$

$P(s_1, s_2)P(s_2, s_3)/P(s_2)$. The numerical accuracy of this approximation can be assessed by defining the ratio R_3 as

$$R_3 \equiv \frac{P(s_1, s_2)P(s_2, s_3)}{P(s_2)P(s_1, s_2, s_3)}, \tag{B.5}$$

and looking at how much R_3 deviates from unity. Since R_3 depends on configurations s_1, s_2, s_3, it is good to determine the mean value of R_3, averaged over all these states: $\langle R_3 \rangle = \frac{1}{2^3} \sum_{s_1,s_2,s_3} R_3$.

After some straightforward algebra, we can find $\langle R_3 \rangle$ as

$$\langle R_3 \rangle = \frac{e^{\kappa}\left[1 + \frac{1}{4}(e^{-4\kappa} - 1)\right] + e^{-\kappa}\left[1 + \frac{1}{4}(e^{4\kappa} - 1)\right]\cosh(2J)}{e^{\kappa} + e^{-\kappa}\cosh(2J)}. \tag{B.6}$$

From this formula, it is clear that for $\kappa \mapsto 0$, we get $\langle R_3 \rangle \mapsto 1$, regardless of the value of J. For large coupling J ($J \gg 1$), we obtain $\langle R_3 \rangle \approx 1 + \frac{1}{4}(e^{4\kappa} - 1)$, which means that $\langle R_3 \rangle$ is essentially close to 1 for small κ. For example, for $\kappa = 0.1$, we get $\langle R_3 \rangle = 1.12$; for $\kappa = 0.3$, we get $\langle R_3 \rangle = 1.58$; and higher values of κ increase $\langle R_3 \rangle$ even further, which breaks the pair approximation. For intermediate values of J (e.g., $J = 1$), the ratio $\langle R_3 \rangle$ achieves value 1.55 for $\kappa = 0.4$, which is a slightly larger value than for the strong coupling case.

Supplementary Information

The code for performed computations is provided in the supplementary material.

Acknowledgments

This work was supported by the Polish National Science Centre (NCN) grant 2021/41/B/ST3/04300 (to J.K.).

References

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the gray matter of the brain. *Journal of Blood Flow and Cerebral Metabolism*, *21*, 1133–1145. 10.1097/00004647-200110000-00001

Balasubramanian, V., Kimber, D., & Berry, M. J (2001). Metabolically efficient information processing. *Neural Computation*, *13*, 799–815. 10.1162/089976601300014358

Barrett, A. B., Billings, G. O., Morris, R. G. M., & van Rossum, M. C. W. (2009). State based model of long-term potentiation and synaptic tagging and capture. *PLOS Computational Biology*, *5*, e1000259. 10.1371/journal.pcbi.1000259

Basu, S., Saha, P. K., Roszkowska, M., Magnowska, M., & Baczynska, E., Das, M., . . . Wlodarczyk, J. (2018). Quantitative 3-D morphometric analysis of individual dendritic spines. *Scientific Reports*, *8*, 3545. 10.1038/s41598-018-21753-8

Benna, M. K., & Fusi, S. (2016). Computational principles of synaptic memory consolidation. *Nature Neuroscience*, *19*, 1697–1706. 10.1038/nn.4401

Bennett, C. H. (1982). The thermodynamics of computation: A review. *International Journal of Theoretical Physics*, *21*, 905–940. 10.1007/BF02084158

Berut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R., & Lutz, E. (2012). Experimental verification of Landauer's principle linking information and thermodynamics. *Nature*, *483*, 187–190. 10.1038/nature10872

Bokota, G., Magnowska, M., Kusmierczyk, T., Lukasik, M., Roszkowska, M., & Plewczynski, D. (2016). Computational approach to dendritic spine taxonomy and shape transition analysis. *Frontiers in Computational Neuroscience*, *10*, 140. 10.3389/fncom.2016.00140

Bonhoeffer, T., & Yuste, R. (2002). Spine motility: Phenomenology, mechanisms, and function. *Neuron*, *35*, 1019–1027. 10.1016/S0896-6273(02)00906-6

Bourne, J. N., & Harris, K. M. (2008). Balancing structure and function at hippocampal dendritic spines. *Annual Review of Neuroscience*, *31*, 47–67. 10.1146/annurev.neuro.31.060407.125646

Chaudhuri, R., & Fiete, I. (2016). Computational principles of memory. *Nature Neuroscience*, *19*, 394–403. 10.1038/nn.4237

Choquet, D., & Triller, A. (2013). The dynamic synapse. *Neuron*, *80*, 691–703. 10.1016/j.neuron.2013.10.013

Feynman, R. P. (1972). *Statistical mechanics: A set of lectures.* Westview Press.

Fusi, S., Drew, P. J., & Abbott, L. F. (2005). Cascade models of synaptically stored memories. *Neuron*, *45*, 599–611. 10.1016/j.neuron.2005.02.001

Gardiner, C. W. (2004). *Handbook of stochastic methods.* Springer.

Glauber, R. J. (1963). Time-dependent statistics of the Ising model. *Journal of Mathematical Physics*, *4*, 294–307. 10.1063/1.1703954

Govindarajan, A., Kelleher, R. J., & Tonegawa, S. (2006). A clustered plasticity model of longterm memory engrams. *Nature Reviews Neuroscience*, *7*, 575–583. 10.1038/nrn1937

Holtmaat, A. J., Trachtenberg, J. T., Wilbrecht, L., Shepherd, G. M., Zhang, X., Knott, G., & Svoboda, K. (2005). Transient and persistent dendritic spines in the neocortex in vivo. *Neuron*, *45*, 279–291. 10.1016/j.neuron.2005.01.003

Kandel, E. R., Dudai, Y., & Mayford, M. R. (2014). The molecular and systems biology of memory. *Cell*, *157*, 163–186.

Karbowski, J. (2009). Thermodynamic constraints on neural dimensions, firing rates, brain temperature and size. *Journal of Computational Neuroscience*, *27*, 415–436. 10.1007/s10827-009-0153-7

Karbowski, J. (2012). Approximate invariance of metabolic energy per synapse during development in mammalian brains. *PLOS One*, *7*, e33425. 10.1371/journal.pone.0033425

Karbowski, J. (2019). Metabolic constraints on synaptic learning and memory. *Journal of Neurophysiology*, *122*, 1473–1490. 10.1152/jn.00092.2019

Karbowski, J. (2021). Energetics of stochastic BCM type synaptic plasticity and storing of accurate information. *Journal of Computational Neuroscience*, *49*, 71–106. 10.1007/s10827-020-00775-0

Karbowski, J., & Urban, P. (2022). *Information encoded in volumes and areas of dendritic spines is nearly maximal across mammalian brains.* www.biorxiv/2021.12.30.474505 (Scientific Reports, in press).

Kasai, H., Matsuzaki, M., Noguchi, J., Yasumatsu, N., & Nakahara, H. (2003). Structure-stability-function relationships of dendritic spines. *Trends in Neuroscience*, *26*, 360–368. 10.1016/S0166-2236(03)00162-0

Kastellakis, G., & Poirazi, P. (2019). Synaptic clustering and memory formation. *Frontiers in Molecular Neuroscience*, *12*, 300. 10.3389/fnmol.2019.00300

Kennedy, M. B. (2000). Signal-processing machines at the postsynaptic density. *Science*, *290*, 750–754. 10.1126/science.290.5492.750

Laughlin, S. B., de Ruyter van Steveninck, R. R., & Anderson, J. C. (1998). The metabolic cost of neural information. *Nature Neuroscience*, *1*, 36–40. 10.1038/236

Leff, H. S., & Rex, A. F. (1990). *Maxwell's demon: Entropy, information, computing*. Princeton University Press.

Leibold, C., & Kempter, R. (2008). Sparseness constrains the prolongation of memory lifetime via synaptic metaplasticity. *Cerebral Cortex*, *18*, 67–77. 10.1093/cercor/bhm037

Levy, W. B., & Baxter, R. A. (1996). Energy efficient neural codes. *Neural Computation*, *8*, 531–543. 10.1162/neco.1996.8.3.531

Levy, W. B., & Baxter, R. A. (2002). Energyefficient neuronal computation via quantal synaptic failures. *Journal of Neuroscience*, *22*, 4746–4755. 10.1523/JNEUROSCI.22-11-04746.2002

Levy, W. B., & Calvert, V. G. (2021). Communication consumes 35 times more energy than computation in the human cortex, but both costs are needed to predict synapse number. *Proceedings of the National Academy of Sciences USA*, *118*, e2008173118. 10.1073/pnas.2008173118

Li, H. L., & van Rossum, M. C. W. (2020). Energy efficient synaptic plasticity. *eLife*, *9*, e50804. 10.7554/eLife.50804

Loewenstein, Y., Kuras, A., & Rumpel, S. (2011). Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. *Journal of Neuroscience*, *31*, 9481–9488. 10.1523/JNEUROSCI.6130-10.2011

Maes, C., Redig, F., & van Moffaert, A. (2000). On the definition of entropy production, via examples. *Journal of Mathematical Physics*, *41*, 1528. 10.1063/1.533195

Makino, H., & Malinow, R. (2011). Compartmentalized versus global synaptic plasticity on dendrites controlled by experience. *Neuron*, *72*, 1001–1011. 10.1016/j.neuron.2011.09.036

Matsuda, H., Ogita, N., Sasaki, A., & Sato, K. (1992). Statistical mechanics of population: The lattice Lotka-Volterra model. *Progress of Theoretical and Experimental Physics*, *88*, 1035–1049. 10.1143/ptp/88.6.1035

Meyer, D., Bonhoeffer, T., & Scheuss, V. (2014). Balance and stability of synaptic structures during synaptic plasticity. *Neuron*, *82*, 430–443. 10.1016/j.neuron.2014.02.031

Miller, P., Zhabotinsky, A. M., Lisman, J. E., & Wang, X.-J. (2005). The stability of a stochastic CaMKII switch: Dependence on the number of enzyme molecules and protein turnover. *PLOS Biology*, *3*, e107. 10.1371/journal.pbio.0030107

Montgomery, J. M., & Madison, D. V. (2004). Discrete synaptic states define a major mechanism of synaptic plasticity. *Trends in Neuroscience*, *27*, 744–750. 10.1016/j.tins.2004.10.006

Nicolis, G., & Prigogine, I. (1977). *Self-organization in nonequilibrium systems*. Wiley.

Niven, B., & Laughlin, S. B. (2008). Energy limitation as a selective pressure on the evolution of sensory systems. *Journal of Experimental Biology*, *211*, 1792–1804. 10.1242/jeb.017574

Parrondo, J. M. R., Horowitz, J. M., & Sagawa, T. (2015). Thermodynamics of information. *Nature Physics*, *11*, 131–139. 10.1038/nphys3230

Peliti, L., & Pigolotti, S. (2021). *Stochastic thermodynamics: An introduction.* Princeton University Press.

Poirazi, P., & Mel, B. W. (2001). Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron*, *29*, 779–796. 10.1016/S0896-6273(01)00252-5

Poo, M.-m., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T., Martin, K. C., . . . Stevens, C. (2016). What is memory? The present state of the engram. *BMC Biology*, *14*, 40. 10.1186/s12915-016-0261-6

Rieke, F., Warland, D., de Ruyter, R., & Bialek, W. (1999). *Spikes: Exploring the neural code.* MIT Press.

Schnakenberg, J. (1976). Network theory of microscopic and macroscopic behavior of master equation systems. *Reviews of Modern Physics*, *48*, 571–585. 10.1103/RevModPhys.48.571

Seifert, U. (2012). Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, *75*, 126001. 10.1088/0034-4885/75/12/126001

Sheng, M., & Hoogenraad, C. C. (2007). The postsynaptic architecture of excitatory synapses: A more quantitative view. *Annual Review of Biochemistry*, *76*, 823–847. 10.1146/annurev.biochem.76.060805.160029

Statman, A., Kaufman, M., Minerbi, A., Ziv, N. E., & Brenner, N. (2014). Synaptic size dynamics as an effective stochastic process. *PLOS Computational Biology*, *10*, e1003846. 10.1371/journal.pcbi.1003846

Takeuchi, T., Duszkiewicz, A. J., & Morris, R. G. M. (2014). The synaptic plasticity and memory hypothesis: Encoding storage and persistence. *Philosophical Transactions of the Royal Society B*, *369*, 20130288. 10.1098/rstb.2013.0288

Urban, P., Tabar, V. R., Denkiewicz, M., Bokota, G., Das, N., Basu, S., & Plewczynski, D. (2020). The mixture of autoregressive hidden Markov models of morphology for dendritic spines during activation process. *Journal of Computational Biology*, *27*, 1471–1485. 10.1089/cmb.2019.0383

Van Baalen, M. (2000). Pair approximation for different spatial geometries. In U. Dieckmann, R. Law, & J. Metz (Eds.), *The geometry of ecological interactions: Simplifying spatial complexity* (pp. 359–387). Cambridge University Press.

Van den Broeck, C., & Esposito, M. (2015). Ensemble and trajectory thermodynamics: A brief introduction. *Physica A*, *418*, 6–16. 10.1016/j.physa.2014.04.035

Volgushev, M., Kudryashov, I., Chistiakova, M., Mukovski, M., Niesmann, J., & Eysel, U. (2004). Probability of transmitter release at neocortical synapses at different temperatures. *Journal of Neurophysiology*, *92*, 212–220. 10.1152/jn.01166.2003

Winnubst, J., Lohmann, C., Jontes, J., Wang, H., & Niell, C. (2012). Synaptic clustering during development and learning: The why, when, and how. *Frontiers in Molecular Neuroscience*, *5*, 70. 10.3389/fnmol.2012.00070

Yadav, A., Gao, Y. Z., Rodriguez, A., Dickstein, D. L., Wearne, Luebke, J., . . . Weaver, C. (2012). Morphologic evidence for spatially clustered spines in apical

dendrites of monkey neocortical pyramidal cells. *Journal of Comparative Neurology*, *520*, 2888–2902. 10.1002/cne.23070

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

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

Received February 4, 2023; accepted October 4, 2023.