Spontaneous activity emerging from an inferred network model captures complex temporal dynamics of spiking data

Cristiano Capone*,[1,2,a)] Guido Gigante*,[1,3] and Paolo Del Giudice[1,2]
[1)] *Istituto Superiore di Sanità, Rome, Italy*
[2)] *INFN, Sezione di Roma 1, Italy*
[3)] *Mperience s.r.l., Rome, Italy*

The combination of new recording techniques in neuroscience and powerful inference methods recently held the promise to recover useful effective models, at the single neuron or network level, directly from observed data. The value of a model of course should critically depend on its ability to reproduce the dynamical behavior of the modeled system; however, few attempts have been made to inquire into the dynamics of inferred models in neuroscience, and none, to our knowledge, at the network level. Here we introduce a principled modification of a widely used generalized linear model (GLM), and learn its structural and dynamic parameters from *ex-vivo* spiking data. We show that the new model is able to capture the most prominent features of the highly non-stationary and non-linear dynamics displayed by the biological network, where the reference GLM largely fails. Two ingredients turn out to be key for success. The first one is a bounded transfer function that makes the single neuron able to respond to its input in a saturating fashion; beyond its biological plausibility such property, by limiting the capacity of the neuron to transfer information, makes the coding more robust in the face of the highly variable network activity, and noise. The second ingredient is a super-Poisson spikes generative probabilistic mechanism; this feature, that accounts for the fact that observations largely undersample the network, allows the model neuron to more flexibly incorporate the observed activity fluctuations. Taken together, the two ingredients, without increasing complexity, allow the model to capture the key dynamic elements. When left free to generate its spontaneous activity, the inferred model proved able to reproduce not only the non-stationary population dynamics of the network, but also part of the fine-grained structure of the dynamics at the single neuron level.

Dynamic models, in neuroscience and in general, embody in a mathematical form the causal relationships between variables deemed essential to describe the system of interest; of course, the value of the model is measured both by its ability to match observations, and by its predictive power. Typically, along this route, parameters appearing in the model are assigned through a mix of insight from experiments and trial-and-error and, in turn, the study of the model dynamics helps understanding the relevance of each parameter in determining the dynamic regimes accessible to the system. Another approach, initially quite detached from dynamic modeling, is rooted in the domain of statistical inference, and a seminal example in neuroscience was offered by the application of maximum-entropy inference of Ising-like models to multi-electrode recordings of neural activity; in this case, as well as in the later extension to kinetic Ising-like models, the Montecarlo/Glauber dynamics of the model is only meant as a means to sample the probability distribution of interest and it is not claimed to offer a detailed model of the actual dynamics at work in the system. Recently, Generalized Linear Models (GLM) (which incorporate kinetic Ising models as a special case) have been recognized as flexible and powerful inference models[1,2]. In time, efforts have been made to make contact between the two approaches, e.g., in the case of neuroscience, by endowing the coupling structure of the inference model with features motivated by biological plausibility[3–5]; it has also been recognized that a GLM is close to a stochastic version of the spike-response model. Recently, the repertoire of the driven dynamics of GLM models of single neurons has been explored[6]; however, to our knowledge, a largely open issue is to endow inference models with predictive power in terms of the system dynamics. Our approach to this problem is to explore the free, spontaneous dynamics of the inferred model in its relation with the one of the biological system generating the data and, in doing this, to identify the role of different elements of the inference model in determining the spontaneous dynamics of the neuronal network. Indeed, an obvious but persisting problem in the application of inference models to neuroscience has been how to assess the meaning and value of the inferred parameters; a recurring example is offered by the inferred synaptic couplings[3]: in the literature, a cautionary remark is usually included, recognizing that the inferred couplings (whether in the form of synaptic efficacies of more complicated synaptic kernels) are to be meant as 'effective', leaving of course open the problem of what exactly this means (especially in the face of the dramatic subsampling of the underlying biological network). The value of the inferred model cannot be assessed directly by a detailed correspondence between its elements and corresponding elements of the biological system; using the output of a GLM to decode input stimuli is an interesting recent approach[7] Our attitude is that spontaneous activity is a good testing ground to assess whether the inferred model does indeed capture essential features of the biological system, and we choose a case in

[a)] Electronic mail: cristiano0capone@gmail.com
*authors contributed equally

which the spontaneous activity of the biological network is highly non-stationary and irregular: a cultured neuronal network generating spontaneously a wide spectrum of activity fluctuations, from population bursts, to neuronal avalanches and noisy oscillations[8–20]. Such richness is extremely hard to reproduce with a GLM, and to our knowledge no models proposed so far were able to cope with it. For this reason we chose population bursting activity as a challenging benchmark to test our model. In our GLM approach we introduced novel ingredients inspired by biological observation, such as activity dependent negative feedback over different time-scales for the single neuron[8] (spike frequency adaptation, SFA), a bounded transfer function, and a generative probabilistic mechanism for spike generation with super-Poisson fluctuations. We show that these ingredients are crucial to endow the system with non-stationary spontaneous activity and also to account for detailed dynamic features such as temporal correlations and spatio-temporal evolution of network bursts.

RESULTS

Bounded firing rate, super-Poisson spike generation, and their impact on the spontaneous activity of the inferred model

The model we introduce is an extension of GLMs as widely employed for inferring the structure of neuronal networks. With GLMs it shares the three fundamental assumptions that spike generation is a probabilistic process driven by the instantaneous current afferent to the neuron; that, secondly, such current is a linear function of the activity of the neurons in the network and, finally, that there is a non-linear transfer function linking the current to the probabilistic spike generation. The parameters of the model, in this framework, are usually obtained through the numerical maximization of the likelihood that the model itself generated the observed data. The main points of departure between more standard GLMs and the present model lay in the choice of the specific non-linear transfer function and probabilistic generative model; another notable difference is in the introduction, at the single neuron level, of an activity-dependent self-inhibition current, mimicking spike-frequency adaptation (SFA) effects widely observed in real neurons.

Fig. 1, panel A, gives a schematic view of the model. The current $H_i(t)$ instantaneously felt by neuron i sums spikes emitted by other neurons j in the recent past $(t-\Delta t)$ through synaptic kernels $k_{ij}(\Delta t)$, a constant external field h_i, and a set of 5 SFA signals (collectively indicated as $c_i(t)$), that integrate, over different time-scales, the spiking activity S_i of neuron i itself. This current provides input to a non-linear transfer-function $f[H_i]$, which in turn determines the expected firing rate $\lambda_i(t)$ for the spike generation model at the next time bin $t + \mathrm{dt}$. The transfer function and generative model most commonly employed in the literature are, respectively, the exponential function and the Poisson distribution; in the following we will label this choice 'Exp-Poisson'; our model, on the other hand, makes use of a sigmoid function (see Materials and Methods) and of a Negative-Binomial distribution with shape parameter r_{NB} set to 0.2, and will be referred to in the following as 'Sig-NegBin' model.

The synaptic kernels $k_{ij}(\Delta t)$, following[4], are parametrized as a weighted sum of 4 'raised cosine' functions peaking at times spanning a range between 10 ms (equal to the chosen time bin dt) and 30 ms, with support extending in the past for $15\,\mathrm{dt} = 150\,\mathrm{ms}$ ($k_{ij}(\Delta t) \equiv 0$ for $\Delta t > 150\,\mathrm{ms}$). The sigmoid transfer function has two free parameters: the maximum firing rate attainable by the neuron and an asymmetry parameter that determines how fast this maximum is approached for high currents compared to the other asymptote at 0 for strongly negative currents. Finally each of the 5 SFA signals c_i, that are set equal for each neuron, has associated a characteristic time-scale τ_c for the integration of the spike train $S_i(t)$ and a weight g_c with which $c_i(t)$ enters the sum of the current $H_i(t)$. These latter parameters, the weights defining each of the $k_{ij}(\Delta t)$ functions, and the free parameters of the sigmoid function are all determined through the likelihood maximization procedure on data (see Materials and Methods for further details). In the case of the Exp-Poisson model, the only optimized parameters are the $k_{ij}(\Delta t)$ weights, both for $i \neq j$ and $i = j$ ('post-spike filter'); the k_{ii} functions are parametrized as a weighted sum of 10 raised cosines peaking at times spanning from 10 ms to 50 ms and again extending to 150 ms in the past; the Sig-NegBin model does not include post-spike filters; thus the total number of free parameters for the Sig-NegBin is actually lower than that for the Exp-Poisson model.

Once the parameters are fitted by maximizing the likelihood, both models can be simulated to generate new spike trains according to two basic simulation modalities. The first one ('driven' mode) feeds the model with the actual spike trains observed in the data and at each time t uses the model's output to make a 1-step prediction of the network activity at step $t + dt$; thus the generated activity never re-enters the network dynamics; the probabilistic distance between the 1-step prediction and the actual recorded network activity is ultimately what is minimized during the optimization procedure. In the second modality ('free' mode) the models are left free to evolve and the generated spike trains are fed back into the network, driving its dynamics at subsequent times.

We trained both the Sig-NegBin and the Exp-Poisson model on data from multi-electrode recordings of *ex-vivo* cultured cortical neurons[10], that show a clear non-stationary dynamics (Fig. 1B, top row) in the rastergrams and in the whole network activity (red line), with large and irregular bursts of synchronous activity interleaved by periods of low, noisy activity ('inter-burst in-

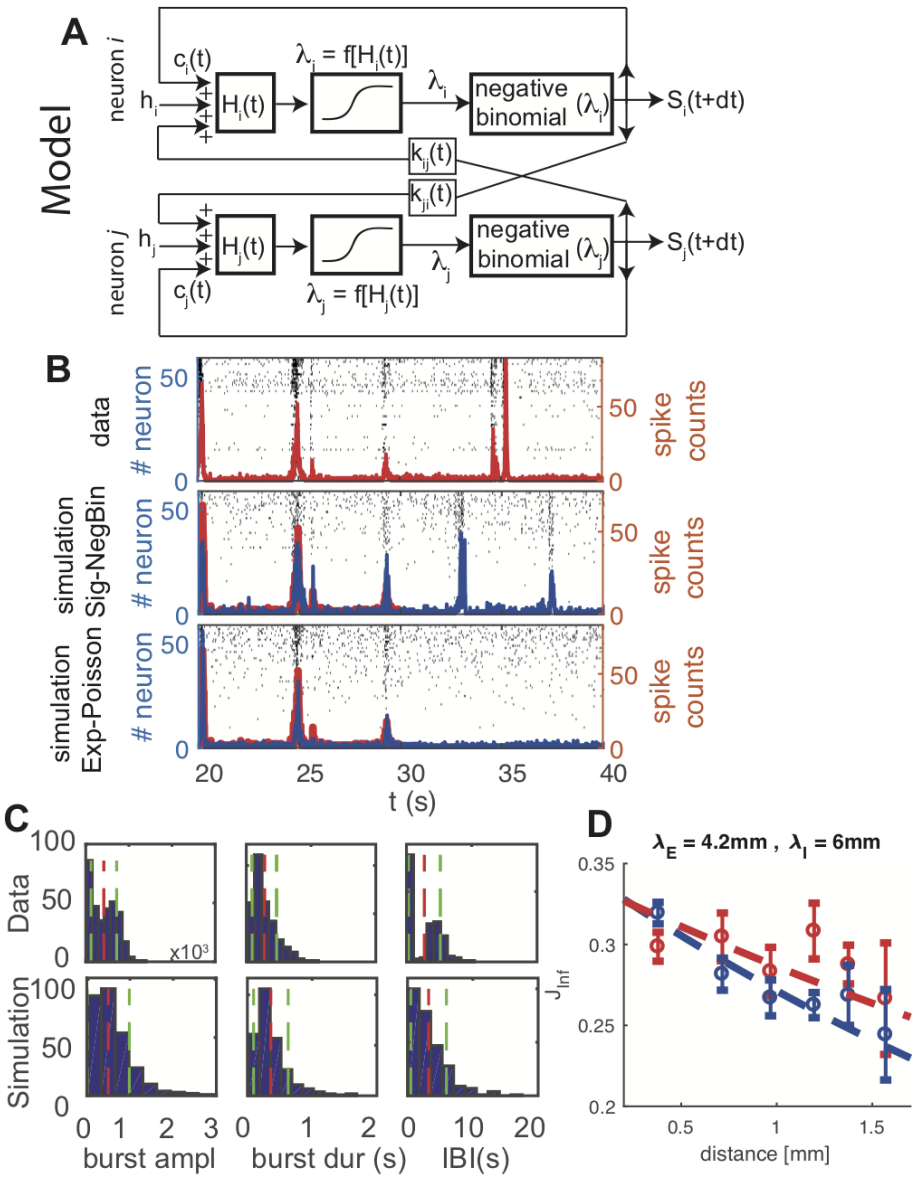

FIG. 1: Spontaneous activity of the inferred model reproduces non-stationary network dynamics. **A** Schematic representation of the model. The current $H_i(t)$ felt by neuron i sums spikes emitted by other neurons j in the recent past $(t - \Delta t)$ through synaptic kernels $k_{ij}(\Delta t)$, a constant external field h_i, and spike-frequency adaptations signals that integrate, over different time-scales, the spiking activity S_i of neuron i itself. This current, fed into a non-linear transfer-function $f[H_i]$, determines the expected number $\lambda_i(t)$ of spikes $S_i(t + \mathrm{d}t)$ that the neuron will probabilistically generate at the next time bin. **B** Rastergrams of *ex-vivo* cultured cortical neurons show a strongly non-stationary dynamics (top row; red line: whole network activity). The proposed Sig-NegBin model (central row, blue line) closely follows the data used to drive its dynamics (red line, t between 20 and 30 s); when left free to run autonomously at $t = 30$ s, the model displays large irregular bursts of activity, qualitatively mimicking the behavior observed in the recordings. The standard Exp-Poisson model (bottom row, blue line), on the other hand, whilst still able to follow the data in driven mode, fails to produce any large activity fluctuation, settling instead into a noisy, stationary state. **C** Comparison between burst amplitudes (total spike counts, left), bursts durations (center) and IBI durations (right) histograms both for the data (top row) and the autonomous activity of the Sig-NegBin model (bottom row). Mean and standard deviation of each distribution are reported with red and green lines respectively. Results for the standard Exp-Poisson model are not shown because of the extreme sparsity of burst events in that case. **D** Average area under the inferred synaptic kernels (absolute value) as a function of the distance separating the electrodes; blue and red circles are for excitatory and inhibitory kernels respectively; dashed lines: exponential fits; the strength (area under the kernel) of inferred excitatory synapses decays faster than that of the inhibitory ones.

tervals' or IBIs). At the end of the training, both models show, in driven mode, an activity that follows quite closely the experimental spike trains (Fig. 1B, middle and bottom rows; red line: data; blue line and rastergrams: simulation). However, when the simulation is switched to free mode at time t = 30 s, the behavior of the two models clearly diverges: whilst the Sig-NegBin is able to autonomously sustain a highly non-stationary activity, the Exp-Poisson only exhibits stationary fluctuations. Note that, in free mode, the activity of the Sig-NegBin model does not closely follow the activity of the real network anymore (top row, last 10 s); this is not surprising: the choice of a probabilistic model as a GLM stems from the assumption that the network dynamics is, to a large extent, influenced by self-induced random fluctuations and, thus, non-deterministic; it is therefore expected that, at least after a transient phase, the model and the data will drift away. Yet the activity spontaneously generated by the model clearly resembles the one seen in data in an intuitive statistical way, showing irregular bursts interspersed with interval of relative quiescence.

Such resemblance is made more quantitative in Fig. 1C. Bursts of the network global activity are detected through the algorithm developed in[8], both for the data (top row) and the autonomous activity of the Sig-NegBin model (bottom row). The comparison between burst amplitudes (total spike counts, left), bursts durations (center) and IBI durations (right) histograms shows a semi-quantitative agreement between the dynamics of the biological network and the inferred model (mean and standard deviation of each distribution are reported with red and green lines respectively). The main discrepancy between model and data is visible in the IBI distribution; "doublets" of bursts in the data, clearly separated by the rest of the IBI distribution, are absent in the model. Based on previous modeling work on similar data[21], the discrepancy is likely due to elements, we did not include in our minimal model, such as short-term synaptic facilitation and depression.

We remark that the analogous distributions for the Exp-Poisson model are not reported because of the extreme sparsity of burst events in that case. Moreover we observe that, despite the major difference in the exhibited dynamic behavior, the inferred synaptic kernels for our model and the Exp-Poisson model are correlated ($c_{\text{Pearson}} = 0.49$ with a $P < 0.05$; correlation between the total areas under each synaptic kernel, as a measure of 'synaptic strength'). The inferred synaptic kernels also show a correlation with the distance separating the electrodes where the activity of the pre- and post-synaptic neurons where recorded (Fig. 1D; blue and red circles are for excitatory and inhibitory kernels respectively; dashed lines: exponential fits). It is worth noting that the connections' decay with distance is faster for excitatory than for inhibitory inferred synapses, consistently with what is found in activity correlations in the neocortex by using 2D multielectrode arrays[22].

The choice of a sigmoidal, saturating transfer function, besides seeming just natural in a GLM meant to model neural activities, turns out to be key in achieving the above results. We repeated the inference procedure with a hybrid Sig-NegBin model, in which the sigmoidal transfer function is replaced with an exponential (Fig. 2A). As above, when during the first 10 s the model is simulated (blue line and rastergrams) in driven mode, it shows a good agreement with the driving data (red line); when the driving is turned off, instead, the model settles in a quiescent state, absent any large bursts of activity. In free mode, the saturation of the sigmoidal function $f[\cdot]$ (Fig. 2B, top-right panel) allows a wide distribution of input currents $H_i(t)$ to the single neuron across different times (Fig. 2B, bottom panel) to give rise to a distribution of firing rates $\lambda_i(t) = f[H_i(t)]$ (Fig. 2B, top-left panel) that is strongly bimodal (notice the logarithmic scale on the y-axis). On the other hand, the fast growth of the exponential function (Fig. 2C, top-right panel) pushes the inference procedure towards somewhat narrower, more conservative distribution of currents (Fig. 2C, bottom panel), resulting in a unimodal distribution of firing rates. The comparison of the two models in driven and free modes suggests that the exponential transfer function makes the system very selectively susceptible to the fluctuations experienced during learning, while the sigmoid transfer function allows for susceptibility to a wider range of fluctuations, which we believe is the root of the ability to spontaneously generate network bursts.

The recorded neurons from the cultured network constitute of course a dramatic subsampling of the network (Fig. 3A, leftmost sketch; though, in fact, the activity of several neurons will be collected by each electrode). The effects of the unobserved neurons on the behavior of the observed part of the network could be in principle very complex[23]; yet we argue that the adoption of a Negative Binomial distribution for the spike counts generation, instead of a Poisson distribution as commonly assumed, captures a relevant part of those effects and is in fact critical to account for the non-stationarity of the spontaneous activity of the network.

One way to approach the problem is to assume that at each time t the generated spike count of a generic neuron $S_i(t)$ is determined by both the field $H_i(t) = h_i + \sum_\tau \sum_j \alpha_{ij}(\tau)\, S_j(t-\tau)$ (the sum of external fields and the one due to the activity of the other observed neurons), and the field due to inputs from unobserved, 'hidden' neurons, h_i^{hid} (Fig. 3A), such that the observed spike count probability distribution would be obtained by integrating over the (unknown) distribution of h_i^{hid}:

$$p(S_i|H_i) = \int dh_i^{\text{hid}} p(S_i|H_i, h_i^{\text{hid}})\; p(h_i^{\text{hid}}) \qquad (1)$$

where we still assume that $p(S_i(t)|H_i(t), h^{\text{hid}})$ is a Poisson distribution, with mean given by $\lambda_i(t) = f\big[H_i(t) + h_i^{\text{hid}}\big]$.

It is intuitive that $p(S_i(t)|H_i(t))$ will now generate super-Poissonian fluctuations; and it turns out that

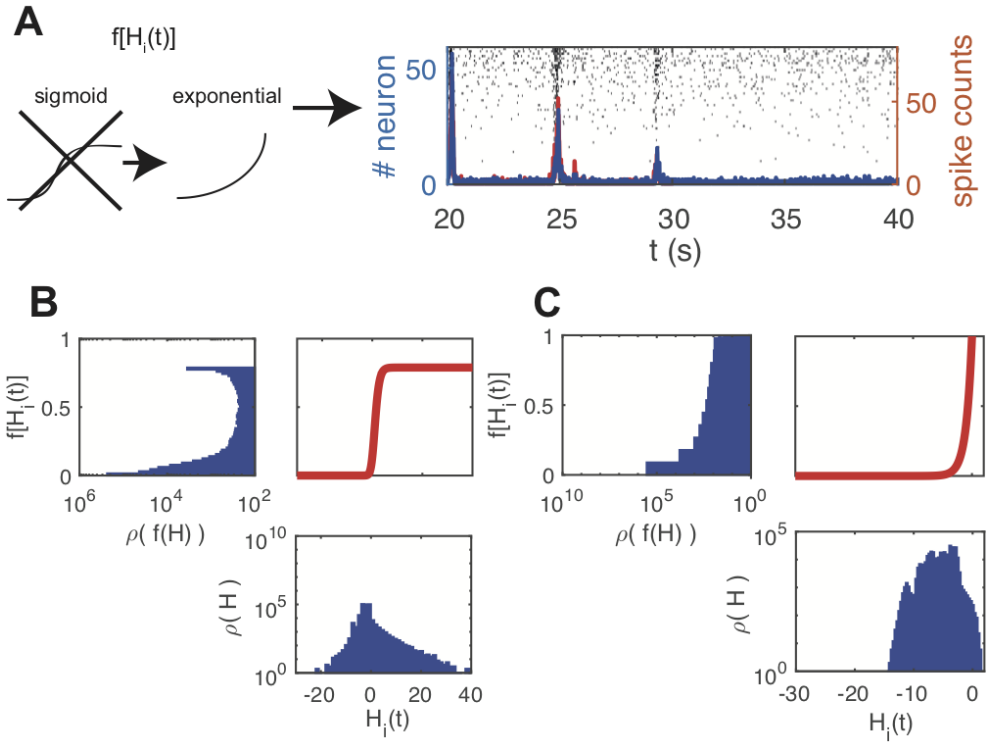

FIG. 2: Bounded spike rate is necessary to generate non-stationary spontaneous activity in the inferred model. **A** Population activity (blue line) and rastergram generated by the model where the bounded transfer function is replaced by an exponential and inference procedure is carried out from scratch. Red line represent the data driving the dynamics for the first 10 s ($20 \leq t \leq 30$ s); afterwards, when left free to generate its spontaneous activity, the model is unable to sustain any non-stationary bursting activity and settles in a quiescent state. **B-C** transfer function (respectively sigmoid and exponential) $f[\cdot]$ (top-right panel) distribution of input currents $H_i(t)$ to the single neuron across different times (bottom panel) and distribution of firing rates $\lambda_i(t) = f[H_i(t)]$ (top-left panel)

larger fluctuations are critical for the ability of the Sig-NegBin model to spontaneously generate bursting activity, as shown above (Fig. 3B upper panel; blue line: data; red line: Sig-NegBin model in free mode). In fact, an inferred Sig-Poisson model shows very sparse bursting (Fig. 3B, middle panel). But it is not just that a broader spike generation distribution favors larger fluctuations in the network activity. We simulated the Sig-NegBin model replacing the Negative Binomial with a Poisson spike generator (Fig. 3B, bottom panel); such 'Poisson replay' of the inferred Sig-NegBin model generates frequent bursts, each lasting for long time on average. Thus the inferred synaptic couplings of the Sig-NegBin model, unlike the ones for the Sig-Poisson, appear to be compatible with a bistable network dynamics, between UP and DOWN states, as also suggested by the results for the sigmoid transfer function (Fig. 2B). Therefore, the main role of a super-Poisson spike generation distribution seems to be to destabilize the UP state, making the inferred model once again more robust to the variability of the free network's activity and thus able to self-sustain a highly non-stationary activity.

Taking together the results in Fig. 2 and Fig. 3, the suggested picture is that on the one hand the sigmoid gain function allows the inference procedure to explore and use ranges of couplings that are effectively forbidden for an exponential gain function, and are essential to spontaneously generate large sudden increase of activity (burst onset); on the other hand, the NegBin generator plays the dynamic role of efficiently destabilizing, with large fluctuations of the spike counts, the high-activity states which are spontaneously generated by the large recurrent excitation. The two ingredients are robustly coupled to spontaneously produce bursting activity.

Albeit representing just one of the forms that $p(S_i(t)|H_i(t))$ can take, the NegBin distribution has been suggested to adequately model fluctuations in observed neural activity[24,25], which makes it a natural candidate. Yet, a NegBin would result from Eq. 1 assuming an exponential transfer function and a log-Gamma distribution for $h_i^{\rm hid}$: the first condition is patently in contradiction with our choice of a saturating transfer function and the second one is not expected to hold in general. We nevertheless tested the adequacy of the NegBin assumption for the Sig-NegBin model as follows. We assumed that $p(h_i^{\rm hid})$ in Eq. 1 is a scaled and shifted version of the distribution of the observed fields H; in turn we estimated such distribution from the inferred synaptic cou-

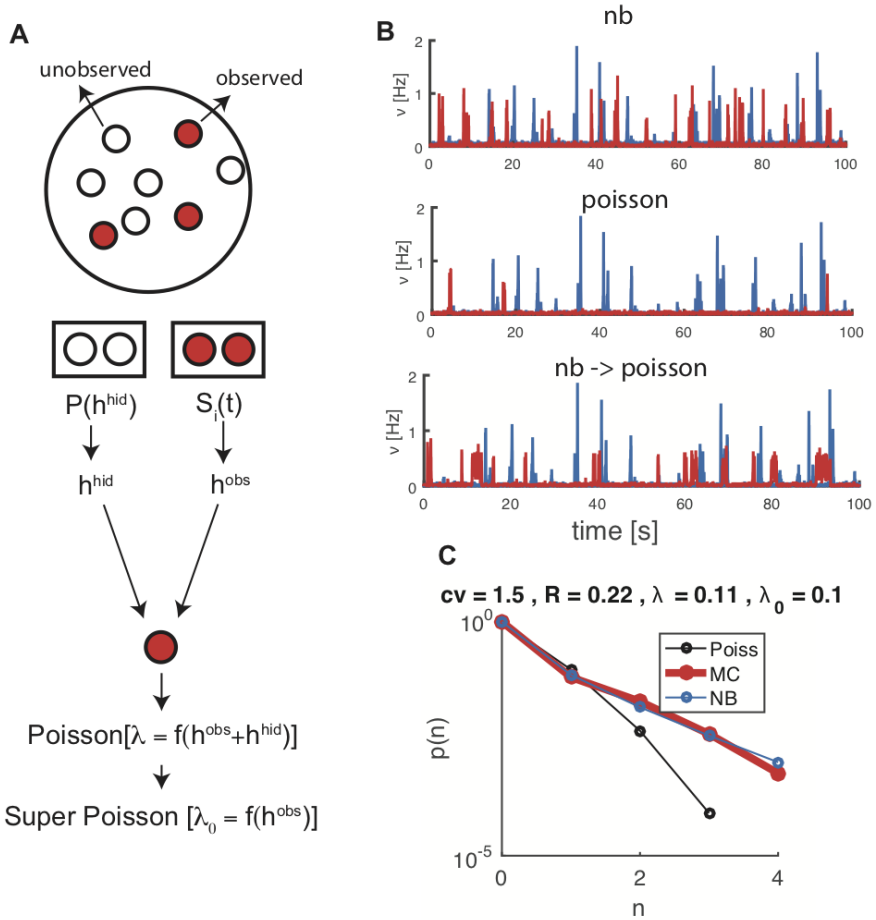

FIG. 3: Rationale and implications of choosing a negative binomial as generative model of spike counts. **A** Sketch showing the effects of the unobserved neurons on the fluctuations in the observed part of the network. **B** Spontaneous bursting activity generated by different models (blue line: data; red line: model in free mode). Upper panel: Sig-NegBin model; middle panel: Sig-Poisson; bottom panel: Sig-NegBin model after replacing the Negative Binomial with a Poisson spike generator. **C** Comparison between the spike counts generated by a negative binomial distribution ($r = 0.22$) with a mean value equal to the average spike count of the neuron (blue line), the Monte Carlo sampled $p(S_i|H_i)$ with best matching values for the first two moments of spike counts (red line) and Poisson distribution for the same average spike count (black line).

plings and the observed experimental spike counts, for the neuron with maximal average activity, for which we expected the differences in the spike count distribution would both show up more clearly, and matter more for the dynamics. From this, we performed a Monte Carlo estimate of $p(S_i|H_i)$ for different values of the mean and the standard deviation defining $p(h_i^{\rm hid})$. Fig. 3C shows a NegBin that is very close to the one used in the Sig-NegBin model ($r = 0.22$) for a mean value equal to the average spike counts of the neuron (blue line); the sampled $p(S_i|H_i)$ with best matching values for the first two moments is reported in red; the two distributions are very similar, especially if compared to the Poisson distribution for the same average spike count (black line). Thus spike counts compatible with a NegBin distribution are naturally accounted for by a plausible assumption on the effect of sub-sampling underlying Eq. 1, with marked deviations from a Poisson distribution. We repeated the above procedure sampling $h_i^{\rm hid}$ from a Gaussian distribution, finding similar results; this suggests that, provided suitable values for the mean and variance are chosen, the precise shape of the $p(h_i^{\rm hid})$ has a minor effect.

The inferred model captures detailed temporal and spatial aspects of the neural dynamics

Although the spike-frequency mechanism with which the Sig-NegBin model is endowed allows for 5 independent time-scales, it is interesting to note that the inferred values consistently aggregate around just two values, at about 100 ms and 2 s respectively, for different temporal segments of the same network's activity (Fig. 4A). This result is consistent with previous work on the same data aiming to infer the main time-scales of the bursting dynamics via a completely different approach[8].

To assess how relevant the adaptation mechanism is for the model dynamics, we performed inference with and

without SFA. Then we simulated the model in driven mode until just after the end of a burst of large amplitude, where adaptation effects are expected to be more pronounced, leaving the model free to spontaneously evolve afterwards. In Fig.4B the SFA and non-SFA dynamics are reported (top and bottom, respectively); black line is the model activity averaged over 500 simulations of the dynamics, while gray shading marks the plus/minus one standard deviation range; red line is the population activity from data. For each realization we collected the next IBI, that is the time interval to the next burst spontaneously generated by the model. The IBI histogram (inset) attains a maximum very close to zero when SFA is not present; such maximum shifts toward higher values with SFA; the two distributions also have different averages (2.3 s *vs* 3.2 s) and coefficients of variation (0.98 *vs* 0.65). Thus SFA significantly increases the average interval to the next burst, reducing at the same time its variability.

We also found in data, consistently with the results reported in[19], a positive correlation between burst amplitude and the length of the previous IBI, an effect that could be possibly attributed to some adaptation mechanism, such as SFA. We estimated this correlation from simulations of the models inferred, with and without SFA, on different recordings; Fig. 4C compares the result of the two models with that found in their experimental counterparts. It is seen that the model without SFA is unable to recover correlations that are significantly different from zero, while the model with SFA produces correlations that are consistent with that measured in the data.

We also addressed the ability of the model to account for the temporal structure and spatial organization of the network activity. We first considered the average order of activation of different spatial locations within a burst (see Methods for details): time-rank is the index of the time bin when a neuron spikes first since the burst onset. By evaluating the average rank for both each neuron in the model and each electrode in the data, we found them strongly correlated ($c = 0.9$). In Fig. 5A we report the spatial distribution of average time-ranks for data (left) and simulations (right).

Looking farther into the burst temporal structure, we asked to what extent the time-ranks in the model and in the data are comparable on a single burst basis. For each burst in the data, we selected and compare the simulated burst with the closest time-rank pattern. Fig 5B and C show two examples of such comparison. To provide a global comparison, capturing the correlation in time-ranks between data and simulations, we proceeded as follows: independently for each neuron, we shuffled the vector of its time-ranks across all the simulated bursts, thereby destroying spatial correlations in the time-ranks, while preserving the average ranks (Fig. 5A, right).

For each burst in the data we took the closest simulated burst and the closest shuffled simulated one, computed the corresponding distances (d^{min}_{shuff} and d^{min}_{sim}) and the distribution of their differences (Fig. 5D). It is seen that, for the large majority of bursts, the simulated burst is closer to the data burst than the surrogate. We therefore can conclude that inferred model captures most of the spatial development of neural activity at the single burst level.

To inquire into the relationship between the inferred synaptic structure and the ensuing network dynamics, we also asked whether for a neuron with low average time-rank (early spiking), the efficacy of its outgoing synapses correlates with the time-ranks of its post-synaptic neurons. Fig. 5E shows indeed a high negative correlation between the efficacy of outgoing synapses and the time-rank of post-synaptic neurons ($c = -0.8$).

DISCUSSION

Several criteria can be identified to evaluate a model, yet the ability to reproduce the behavior of the system under analysis is certainly among the most relevant. We have seen that the addition of few computational ingredients can enable a probabilistic, generative network to autonomously sustain a rich, highly non-stationary dynamics, as observed in the experimental data employed to infer the model's parameters.

In recent years, the field of statistical inference applied to neuronal dynamics has mainly focused on devising models and procedures that could reliably recover the real or effective synaptic structure of a network of neurons, under different dynamical and stimulation conditions[26–28], even though, more often than not, an experimental ground truth was not readily available.

Although, quite surprisingly, the study of the dynamics of the inferred models has been largely neglected, we are not the first to address the issue. In[4] the authors demonstrated, on an in-vitro population of ON and OFF parasol ganglion cells, the ability of a GLM to accurately reproduce the dynamics of the network;[7] studied the response properties of lateral intraparietal area neurons at the single trial, single cell level; the capability of GLM to capture a broad range of single neuron response behaviors was analyzed in[6]. In all these works, however, the focus was on the response of neurons to stimuli of different spatio-temporal complexity; even where network interactions were accounted for, they proved to be, for the overall dynamics displayed by the ensemble, an important, yet not decisive correction. To our knowledge, no published study to date has focused on the autonomous dynamics of GLM networks inferred from neuronal data.

Important progress has also been made on the ability of GLMs to predict single neuron spiking in an ensemble of neurons, with accuracy higher than that provided by methods based on the instantaneous state of the ensemble itself[2] and the PSTH[4]. In the lingo of the present paper, though, these attempts were based on simulations of the inferred network in driven mode, with at most the activity of a single neuron free to reenter just that sin-

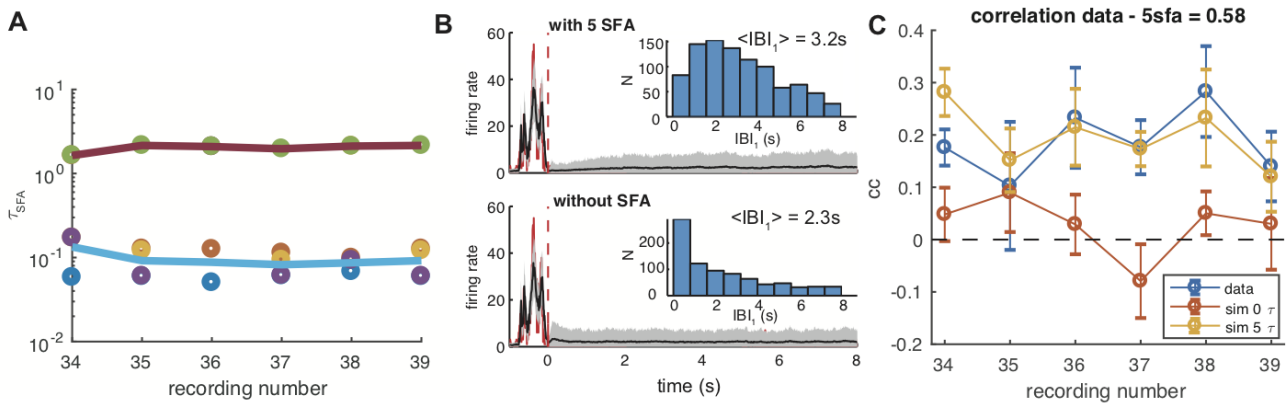

FIG. 4: Effect of inferred time-scales and SFA currents. **A** Inferred τ_{SFA} for different recordings for the same preparation; the five inferred time-scales cluster around two values for all recordings, one about $100ms$, the other around $2s$. **B** Comparison between the model dynamics with and without SFA. In black, and gray-shading, are represented the average and variance over 500 simulations. Red line is the average activity from data during a large chosen burst. Insets: histograms of the time to the subsequent burst. **C** Correlation between the burst amplitude and the time since the preceding burst.

gle neuron's dynamics. To our knowledge, our work for the first time has attempted a direct, microscopic comparison between the activity autonomously generated by an inferred GLM network and the one observed in the biological data. The free mode of simulation opens the way, in principle, to multi-time-step, whole network activity predictions. Being the model probabilistic and the data noisy, it is of course expected that the ability of the model, in free simulation mode after having been driven by the data for some time, to follow the experimentally observed spike trains should quickly deteriorate. In fact, this is what we found with our model: a burst becomes practically unforeseeable going some 100-200 ms in the past. This result is in contrast with the impressive findings reported in[29], in which the authors, on recordings similar to the ones used in this work, through a new model-free method based on state-space reconstruction, were able to predict the occurrence of a burst even 1 s in advance. Whilst it is not clear why this is the case, we applied the state-reconstruction method on our data with negative results, in agreement with what we found with our model.

We presented several semi-quantitative comparisons between the spike trains generated by the model network, without any time-varying external input, and the data from which the model's parameters have been inferred. Such tests, albeit successful, represent of course just an example of the many possible ways to compare two bursting dynamics. Yet we stress how a similar model, that has been widely and successfully applied to the analysis of neuronal data and has a slightly greater number of parameters too, failed in large part to reproduce the most prominent features of the dynamics displayed by the data.

Two interrelated questions pertain to how representative the chosen experimental system is of neuronal network dynamics in general and the capability of the presented framework to generalize to other kinds of non-stationary neural activities. As for the first question, our primary concern in selecting an *ex-vivo* bursting culture has been to study a system that, absent any external stimuli, would show a non-trivial collective behavior. As discussed above, the capability of GLMs to mimic the response properties of single neurons or a network to stimuli has already been well investigated. On the other hand, numerous studies in literature witness to the complexity of spatial and temporal behavior of neuronal cultures[8–20]. As for the second question, we remark that our primary goal was to devise a 'proof of concept' extension to show that a GLM can autonomously generate a complex non-stationary activity, and not to establish general conditions for such capability. We therefore chose a bursting regime as a testbed for the model not because it is exemplary, but because it represented a clear challenge: a collective, non-linear phenomenon that demands the model to adapt to a broad spectrum of time- and space-scales. There were of course sensible alternatives to choose from, the most relevant being, *e.g.*, the UP and DOWN dynamics in brain slices; although we have not tackled the issue directly, our results (see Fig. 3B, bottom plot) hint at a inherent capability of the model to sustain a noisy bistable dynamics, as more in general suggested in[30]. Nevertheless, the extent to which generative models are able to adapt to different regimes, and how dynamic network effects and stimulation characteristics interact in a generative model like the one studied here are open issues, that will probably attract much interest in the near future.

Given our goal to establish the capability of a GLM to reproduce a complex, non-linear dynamics, we found two ingredients - a non-symmetric sigmoidal transfer function and a Negative Binomial spike-generation model - that proved to be minimal: absent one of them, the model's performances were drastically impaired. Even if we do not claim these specific ingredients to be necessary in general, we are persuaded they represent instances of mech-

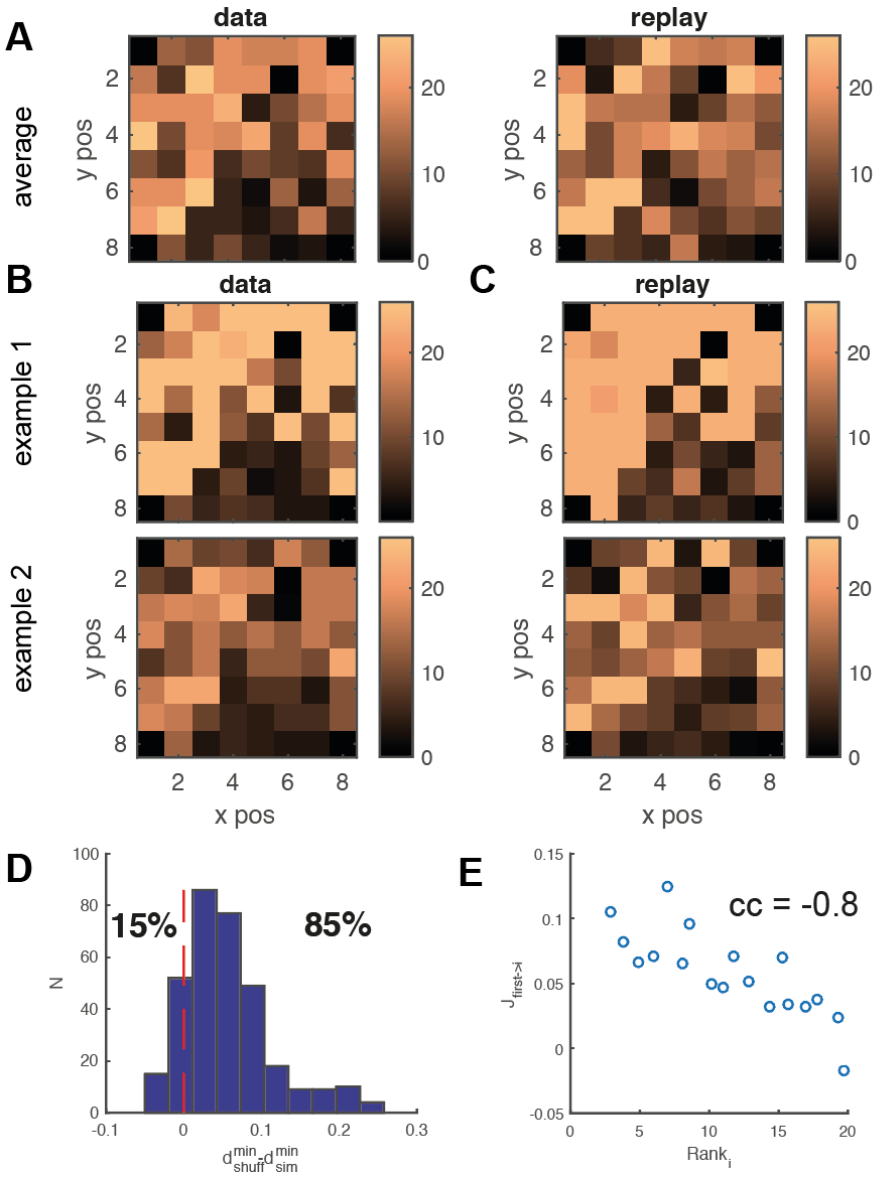

FIG. 5: Spatial microscopic network dynamics, data *vs* model. **A** Spatial distribution of the average time-rank of neuron activations during bursts for data and model. Averages are computed over all bursts in the data and a simulation of comparable length. **B** Spatial distribution of time-rank for 2 single bursts in data. **C** Spatial distribution of time-rank for the 2 bursts from simulation, most similar to those showed in B. **D** Distribution of the difference of the similarity between bursts from data, and both bursts from simulation and bursts from surrogate simulated data. **E** Scatterplot of post-synaptic efficacies of the neuron with smallest time-rank *vs* the rank of its post-synaptic targets.

anisms that are both biologically plausible and computationally desirable for the model. Although a sigmoidal, saturating transfer function would appear to be a natural choice for a model meant to reproduce neural activities, surprisingly this option has never been explored, to our knowledge, to perform network inference on neuronal recordings. While the use of an exponential transfer function in GLMs is grounded in general statistical requirements[31], in our case an asymmetric sigmoid appears to be the single most important factor explaining the success of the proposed model. As already noted, the saturation of the sigmoidal function allows naturally for a bimodal distribution of firing rates and thus makes model's behavior more robust in the face of the intrinsic fluctuations of network activity. Therefore, if the specific form we chose is just one possible instance among many plausible ones, a saturating behavior is expected to benefit the model beyond the immediate scope of the present work; a certain degree of asymmetry also appears to be beneficial according to our results, allowing for differential sensitivity to high and low input levels.

It is interesting to note how this finding contrast with

the success of non-saturating transfer functions in deep learning literature, where the introduction of rectified linear units, in place of the standard logistic ones, has represented one of the major breakthroughs of recent years[32]. Such units exhibit 'intensity equivariance', that is the ability to readily generalize to data points that differ only for a scale factor; whilst such property is clearly valuable when dealing with data such natural images and sounds, when applying machine learning techniques to very noise and sparse data, such as in the case studied here, bounded transfer functions are probably beneficial exactly for the opposite reason: they filter out most of the incoming information to gain a poorer but more robust coding in the output.

We have seen that, if the role of a sigmoid gain function seems to facilitate a bistable behavior of the network, the negative binomial's one is to efficiently destabilize, with large fluctuations of the spike counts, the high-activity state. Recent experimental evidence has emerged supporting the Negative Binomial distribution as a candidate for the spike counts variability in real neurons[25]; moreover, a negative binomial spike generation has already been adopted in a model combining GLM with flexible graph-theoretic priors for the connectivity[33]. Although, then, our second ingredient already finds experimental and theoretical support in the literature, we provide a new hypothesis on why the super-Poisson statistics arises: as the effect of input fluctuations generated by the activity of neurons that have not been recorded. And our hypothesis directly hints to the Negative Binomial as just one possible way to model such effect, where over-dispersed spike counts are, instead, a necessary signature of it; it is this more general feature, in our opinion, that proved to be so important for the success of the proposed model. Interestingly, our findings resonate with the recently proposed role of fluctuating unobserved variables in the emergence of criticality in a wide range of systems[34].

Our hypothesis presupposes that the recorded neurons have been chosen at random from the whole network; this is probably not the case in many instances, where a local set of units is sampled instead; such non-random sub-sampling can potentially lead to strong deviations from expected behavior[23]; this of course could produce systematic effects in the inference. Besides, our model cannot incorporate the effects of slow changes in single neuron excitability, as observed experimentally in[35,36], and analyzed in a statistical model in[25].

METHODS

Experimental data

As originally described in[10], cortical neurons were obtained from newborn rats within 24 hours after birth, following standard procedures. Briefly, the neurons were plated directly onto a substrate-integrated multielectrode array (MEA). The cells were bathed in MEM supplemented with heat-inactivated horse serum (5%), glutamine (0.5 mM), glucose (20 mM), and gentamycin (10 μg/ml) and were maintained in an atmosphere of 37° C, 5% CO2/95% air in a tissue culture incubator as well as during the recording phases. The data analyzed here was collected during the third week after plating, thus allowing functional and structural maturation of the neurons. MEAs of 60 Ti/Au/TiN electrodes, 30 μm in diameter, and spaced 200 μm from each other (Multi Channel Systems, Reutlingen, Germany) were used. The insulation layer (silicon nitride) was pretreated with poly-D-lysine. All experiments were conducted under a slow perfusion system with perfusion rates of $\sim$100 μl/h. A commercial 60-channel amplifier (B-MEA-1060; Multi Channel Systems) with frequency limits of 1-5000 Hz and a gain of 1024$\times$ was used. The B-MEA-1060 was connected to MCPPlus variable gain filter amplifiers (Alpha Omega, Nazareth, Israel) for additional amplification. Data was digitized using two parallel 5200a/526 analog-to-digital boards (Microstar Laboratories, Bellevue, WA). Each channel was sampled at a frequency of 24000 Hz and prepared for analysis using the AlphaMap interface (Alpha Omega). Thresholds (8$\times$ root mean square units; typically in the range of 10-20 μV) were defined separately for each of the recording channels before the beginning of the experiment. The electrophysiological data are freely available from S. Marom (`http://marom.net.technion.ac.il/`).

ACKNOWLEDGEMENTS

We are grateful to Shimon Marom for having shared his data, and for several illuminating discussions along the way.

[1] W. Truccolo, U. T. Eden, M. R. Fellows, J. P. Donoghue, and E. N. Brown, "A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects," Journal of neurophysiology **93**, 1074–1089 (2005).

[2] W. Truccolo, L. R. Hochberg, and J. P. Donoghue, "Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes," Nature Neuroscience **13**, 105–111 (2010).

[3] C. Capone, C. Filosa, G. Gigante, F. Ricci-Tersenghi, and P. Del Giudice, "Inferring synaptic structure in presence of neural interaction time scales," PloS one **10**, e0118412 (2015).

[4] J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. Chichilnisky, and E. P. Simoncelli, "Spatio-temporal correlations and visual signalling in a complete neuronal population," Nature **454**, 995–999 (2008).

[5] I. H. Stevenson, J. M. Rebesco, N. G. Hatsopoulos, Z. Haga, L. E. Miller, and K. P. Kording, "Bayesian inference of functional connectivity and network structure from spikes," IEEE Transactions on Neural Systems and Rehabilitation Engineering **17**, 203–213 (2009).

[6] A. I. Weber and J. W. Pillow, "Capturing the dynamical repertoire of single neurons with generalized linear models," Neural Computation **29**, 3260–3289 (2017).

[7] I. M. Park, M. L. Meister, A. C. Huk, and J. W. Pillow, "Encoding and decoding in parietal cortex during sensorimotor decision-making," Nature neuroscience **17**, 1395–1403 (2014).

[8]G. Gigante, G. Deco, S. Marom, and P. Del Giudice, "Network events on multiple space and time scales in cultured neural networks and in a stochastic rate model," PLoS computational biology **11**, e1004547 (2015).

[9]T. Baltz, A. Herzog, and T. Voigt, "Slow oscillating population activity in developing cortical networks: models and experimental results." J Neurophysiol **106**, 1500–1514 (2011).

[10]D. Eytan and S. Marom, "Dynamics and effective topology underlying synchronization in networks of cortical neurons," The Journal of Neuroscience **26**, 8465–8476 (2006).

[11]M. Giugliano, P. Darbon, M. Arsiero, H.-R. Lüscher, and J. Streit, "Single-neuron discharge properties and network activity in dissociated cultures of neocortex," Journal of Neurophysiology **92**, 977–996 (2004).

[12]T. Gritsun, J. le Feber, J. Stegenga, and W. L. Rutten, "Experimental analysis and computational modeling of interburst intervals in spontaneous activity of cortical neuronal culture," Biological Cybernetics **105**, 197–210 (2011).

[13]I. P. I. Park, D. X. D. Xu, T. B. DeMarse, and J. C. Principe, "Modeling of Synchronized Burst in Dissociated Cortical Tissue: An Exploration of Parameter Space," (2006).

[14]D. A. Wagenaar, J. Pine, and S. M. Potter, "An extremely rich repertoire of bursting patterns during the development of cortical cultures," BMC neuroscience **7**, 11 (2006).

[15]J. M. Beggs and D. Plenz, "Neuronal avalanches in neocortical circuits," The Journal of Neuroscience **23**, 11167–11177 (2003).

[16]T. Petermann, T. C. Thiagarajan, M. A. Lebedev, M. A. Nicolelis, D. R. Chialvo, and D. Plenz, "Spontaneous cortical activity in awake monkeys composed of neuronal avalanches," Proceedings of the National Academy of Sciences **106**, 15921–15926 (2009).

[17]D. Plenz and T. C. Thiagarajan, "The organizing principles of neuronal avalanches: cell assemblies in the cortex?" Trends in neurosciences **30**, 101–110 (2007).

[18]D. Plenz and H. G. Schuster, *Criticality in neural systems* (Wiley-VCH New York, NY, 2014).

[19]F. Lombardi, H. J. Herrmann, D. Plenz, and L. de Arcangelis, "Temporal correlations in neuronal avalanche occurrence," Scientific reports **6**, 24690 (2016).

[20]M. Yaghoubi, T. de Graaf, J. G. Orlandi, F. Girotto, M. A. Colicos, and J. Davidsen, "Neuronal avalanche dynamics indicates different universality classes in neuronal cultures," Scientific reports **8**, 3417 (2018).

[21]T. Masquelier and G. Deco, "Network bursting dynamics in excitatory cortical neuron cultures results from the combination of different adaptive mechanism," PloS one **8**, e75824 (2013).

[22]A. Peyrache, N. Dehghani, E. N. Eskandar, J. R. Madsen, W. S. Anderson, J. A. Donoghue, L. R. Hochberg, E. Halgren, S. S. Cash, and A. Destexhe, "Spatiotemporal dynamics of neocortical excitation and inhibition during human sleep," Proceedings of the National Academy of Sciences **109**, 1731–1736 (2012).

[23]A. Levina and V. Priesemann, "Subsampling scaling," Nature communications **8**, 15140 (2017).

[24]A. Onken, S. Grünewälder, M. H. Munk, and K. Obermayer, "Analyzing short-term noise dependencies of spike-counts in macaque prefrontal cortex using copulas and the flashlight transformation," PLoS computational biology **5**, e1000577 (2009).

[25]R. L. Goris, J. A. Movshon, and E. P. Simoncelli, "Partitioning neuronal variability," Nature neuroscience **17**, 858–865 (2014).

[26]Y. Roudi, B. Dunn, and J. Hertz, "Multi-neuronal activity and functional connectivity in cell assemblies," Current opinion in neurobiology **32**, 38–44 (2015).

[27]J. Tyrcha, Y. Roudi, M. Marsili, and J. Hertz, "The effect of nonstationarity on models inferred from neural data," Journal of Statistical Mechanics: Theory and Experiment **2013**, P03005 (2013).

[28]O. M. A. D. U. F. Trang-Anh Nghiem, Bartosz Telenczuk, "Maximum entropy models reveal the correlation structure in cortical neural activity during wakefulness and sleep," ArXiv (2018).

[29]S. Tajima, T. Mita, D. J. Bakkum, H. Takahashi, and T. Toyoizumi, "Locally embedded presages of global network bursts," arXiv preprint arXiv:1703.04176 (2017).

[30]V. Rostami, P. P. Mana, S. Grün, and M. Helias, "Bistability, non-ergodicity, and inhibition in pairwise maximum-entropy models," PLOS Computational Biology **13**, e1005762 (2017).

[31]J. K. Lindsey, *Applying generalized linear models* (Springer Science & Business Media, 2000).

[32]V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th international conference on machine learning (ICML-10)* (2010) pp. 807–814.

[33]S. Linderman, R. P. Adams, and J. W. Pillow, "Bayesian latent structure discovery from multi-neuron recordings," in *Advances in Neural Information Processing Systems* (2016) pp. 2002–2010.

[34]D. J. Schwab, I. Nemenman, and P. Mehta, "Zipfs law and criticality in multivariate data without fine-tuning," Physical review letters **113**, 068102 (2014).

[35]A. Gal, D. Eytan, A. Wallach, M. Sandler, J. Schiller, and S. Marom, "Dynamics of excitability over extended timescales in cultured cortical neurons," The Journal of Neuroscience **30**, 16332–16342 (2010).

[36]A. Gal and S. Marom, "Self-organized criticality in single-neuron excitability," Physical Review E **88**, 062717 (2013).