# Extracting neuronal functional network dynamics via adaptive Granger causality analysis

Alireza Sheikhattar[a], Sina Miran[a], Ji Liu[b], Jonathan B. Fritz[c], Shihab A. Shamma[a,c], Patrick O. Kanold[b,c], and Behtash Babadi[a,c,1]

[a]Department of Electrical & Computer Engineering, University of Maryland, College Park, MD 20742; [b]Department of Biology, University of Maryland, College Park, MD 20742; and [c]The Institute for Systems Research, University of Maryland, College Park, MD 20742

Quantifying the functional relations between the nodes in a network based on local observations is a key challenge in studying complex systems. Most existing time series analysis techniques for this purpose provide static estimates of the network properties, pertain to stationary Gaussian data, or do not take into account the ubiquitous sparsity in the underlying functional networks. When applied to spike recordings from neuronal ensembles undergoing rapid task-dependent dynamics, they thus hinder a precise statistical characterization of the dynamic neuronal functional networks underlying adaptive behavior. We develop a dynamic estimation and inference paradigm for extracting functional neuronal network dynamics in the sense of Granger, by integrating techniques from adaptive filtering, compressed sensing, point process theory, and high-dimensional statistics. We demonstrate the utility of our proposed paradigm through theoretical analysis, algorithm development, and application to synthetic and real data. Application of our techniques to two-photon Ca²⁺ imaging experiments from the mouse auditory cortex reveals unique features of the functional neuronal network structures underlying spontaneous activity at unprecedented spatiotemporal resolution. Our analysis of simultaneous recordings from the ferret auditory and prefrontal cortical areas suggests evidence for the role of rapid top-down and bottom-up functional dynamics across these areas involved in robust attentive behavior.

Granger causality | adaptive filtering | functional network dynamics | point processes | sparsity

Converging lines of evidence in neuroscience, from neuronal network models and neurophysiology (1–8) to resting-state imaging (9–11), suggest that sophisticated brain function results from the emergence of distributed, dynamic, and sparse functional networks underlying the brain activity. These networks are highly dynamic and task-dependent, which allows the brain to rapidly adapt to abrupt changes in the environment, resulting in robust function. To exploit modern-day neuronal recordings to gain insight into the mechanisms of these complex dynamic functional networks, computationally efficient time series analysis techniques capable of simultaneously capturing their dynamicity, sparsity, and statistical characteristics are required.

Historically, various techniques such as cross-correlogram (12) and joint peristimulus time histogram (13) analyses have been used for inferring the statistical relationship between pairs of spike trains (12–14). Despite being widely used, these methods are unable to provide reliable estimates of the underlying directional patterns of causal interactions among an ensemble of interacting neurons due to the intrinsic deficiencies in identification of directionality, low sensitivity to inhibitory interactions (15), and susceptibility to the indirect interactions and latent common inputs.

Methods based on Granger causality (GC) analysis have shown promise in addressing these shortcomings and have thus been used for inferring functional interactions from neural data of different modalities (16–19). The rationale behind GC analysis is based on two principles: the temporal precedence of cause over effect and the unique information of cause about the effect. Given two time series $\{X_t, Y_t\}_{t=1}^{T}$, if including the history of $Y_t$ can improve the prediction of $X_{t+1}$, it is implied that the history of $Y_t$ contains unique information about $X_t$, not captured by other covariates. In this case, we say that $Y_t$ has a G-causal link to $X_t$.

Numerous efforts have been dedicated to extending the bivariate GC measure to more general settings, such as the conditional form of GC in ref. 20 for multivariate setting, and several frequency-domain variants of GC (21–23). Despite significant advances in time series analysis using GC and its variants, when applied to neuronal data, the existing methods exhibit several drawbacks.

First, most existing methods for causality inference provide static estimates of the causal influences associated with the entire data duration. Although suitable for the analysis of stationary neural data, they are not able to capture the rapid task-dependent changes in the underlying neural dynamics. To address this challenge, several time-varying measures of causality have been proposed in the literature based on Bayesian filtering and wavelets (24–30). Second, there are very few causal inference

## Significance

**Probing functional interactions among the nodes in a network is crucial to understanding how complex systems work. Existing methodologies widely assume static network structures or Gaussian statistics or do not take account of likely sparse interactions. They are therefore not well-suited to neuronal spiking data with rapid task-dependent dynamics, binary statistics, and sparse functional dependencies. We develop an inference framework for extracting functional network dynamics from neuronal data by integrating techniques from adaptive filtering, compressed sensing, point processes, and high-dimensional statistics. We derive efficient estimation algorithms and precise statistical inference procedures. We apply our proposed techniques to experimentally recorded neuronal data to probe the neuronal functional networks underlying attentive behavior. Our techniques provide substantial gains in computation, resolution, and statistical robustness.**

approaches to take into account the sparsity of the functional networks (31–33). As an example, authors in ref. 31 introduced a method for sparse identification of functional connectivity patterns from large-scale functional imaging data. Despite their success in inferring sparse connectivity patterns, these techniques assume static connectivity structures.

Third, most existing approaches are tailored for continuous-time data, such as electroencephalography (EEG) and local field potential recordings, which limits their utility when applied to binary neuronal spike recordings. These methods are generally based on multivariate autoregressive (MVAR) modeling, with a few nonparametric exceptions (30, 34). Some efforts have been made to adapt the MVAR modeling to neuronal spike trains (17, 35, 36). For instance, the binary spikes were preprocessed in refs. 17 and 35 via a smoothing kernel, which significantly distorts the temporal details of the neuronal dynamics. In addition, the frequency-domain GC analysis techniques implicitly assume that the data have rich oscillatory dynamics. Although this assumption is valid for steady-state EEG responses or resting-state recordings, spike trains recorded from cortical neuronal ensembles often do not exhibit any oscillatory behavior.

To address the third challenge, point process modeling and estimation have been successfully used in capturing the stochastic dynamics of binary neuronal spiking data (37, 38). This framework has been particularly used for inferring functional interactions in neuronal ensembles from spike recordings (32, 38–42). A maximum likelihood (ML)-based approach was introduced in ref. 38 based on a network likelihood formulation of the point process model; a model-based Bayesian approach based on point process likelihood models with sparse priors on the connectivity pattern was introduced in ref. 32. Among the more recent results, an information-theoretic measure of causality is proposed in ref. 41; a static GC measure based on point process likelihoods is proposed in ref. 40. However, a modeling and estimation framework to simultaneously take into account the dynamicity and sparsity of the G-causal influences as well as the statistical properties of binary neuronal spiking data is lacking.

In this paper, we close this gap by developing a dynamic measure of GC by integrating the forgetting-factor mechanism of recursive least squares (RLS), point process modeling, and sparse estimation. To this end, we first exploit the prevalent parsimony of neurophysiological time constants manifested in neuronal spiking dynamics, such as those in sensory neurons with sharp tunings, as well as the potential low-dimensional structure of the underlying functional networks. These features can be captured by point process models in which the cross-history dependence of the neurons is described by sparse vectors. We then use an exponentially weighted log-likelihood framework (43) to recursively estimate the model parameters via sparse adaptive filtering, thereby defining a dynamic measure of GC, which we call the adaptive GC (AGC) measure.

The significance of sparsity in our approach is twofold. First, while the functional networks may not be truly sparse, they can often be parsimoniously described by a sparse set of significant functional links. Our models can indeed capture these significant links through sparse cross-history dependence. Second, sparsity enables stable estimation in the face of limited data. This is particularly important for adaptive estimation, where the goal is to reliably estimate a large number of cross-history parameters using short, effective observation windows.

We next develop a statistical inference framework for the proposed AGC measure by extending classical results on the analysis of deviance to our sparse dynamic point process setting. We provide simulation studies to evaluate the identification and tracking capabilities of our proposed methodology, which reveal remarkable performance gains compared with existing techniques, in both detecting the existing G-causal links and avoiding false alarms, while capturing the dynamics of the G-causal interactions

in a neuronal ensemble. We finally apply our techniques to two experimentally recorded datasets: two-photon imaging data from the mouse auditory cortex under spontaneous activity and simultaneous single-unit recordings from the ferret primary auditory (A1) and prefrontal cortices (PFC) under a tone-detection task. Our analyses reveal the temporal details of the functional interactions between A1 and PFC under attentive behavior as well as among the auditory neurons under spontaneous activity at unprecedented spatiotemporal resolutions. In addition to their utility in analyzing neuronal data, our techniques have potential application in extracting functional network dynamics in other domains beyond neuroscience, such as social networks or gene regulatory networks, thanks to the plug-and-play nature of the algorithms used in our inference framework.

## Theory and Algorithms

**Preliminaries and Notations.** We use point process modeling to capture neuronal spiking statistics. A point process is a stochastic sequence of discrete events occurring at random points in continuous time. When adapted to the discrete time domain, point process models have proven to be successful in capturing the statistics of neuronal spiking (37, 44–46). Our analysis in this paper is based on discrete point process models, in which the observation interval $\mathcal{T}$ is discretized to $T := \lceil \mathcal{T}/\Delta \rceil$ bins of length $\Delta$. By choosing $\Delta$ small enough, the resulting neuronal data transform into a binary sequence $\{n_t\}_{t=1}^{T}$. The statistics of this sequence can be fully characterized by its conditional intensity function (CIF) denoted by $\lambda_t$, representing the neuron's instantaneous firing rate at time bin $t$ conditional on all of the data and covariates up to time $t$. The binary spiking sequence can be modeled by a conditionally independent Bernoulli process with success probability of $\lambda_t \Delta$.

Suppose that at time bin $t$ the effective neural covariates are collected in a vector $\mathbf{x}_t \in \mathbb{R}^M$. Such covariates include the neuron's spiking history, the history of the activity of other neurons, and extrinsic stimuli. A dynamic generalized linear model (GLM) with logistic link function for this neuron's CIF (43) is given by

$$\lambda_t \Delta := \text{logit}^{-1}\left(\boldsymbol{\omega}_t' \mathbf{x}_t\right), \qquad [1]$$

where $\text{logit}^{-1}(z) := \frac{1}{1+\exp(-z)}$, for $z \in \mathbb{R}$, is the logistic function, and $\boldsymbol{\omega}_t$ denotes the time-varying parameter vector of length $M$ at time $t$, characterizing the dynamics of the underlying neuronal encoding process. We consider a multiscale window-based model for $\boldsymbol{\omega}_t$ with piecewise constant dynamics within windows of length $W \geq 1$ samples. To this end, we segment the spiking activity $\{n_t\}_{t=1}^{T}$ to $K := \lceil T/W \rceil$ windows of length $W$ bins and assume that $\boldsymbol{\omega}_t = \boldsymbol{\omega}_k$ for all $t \in [(k-1)W + 1, \ldots, kW]$, where $k = 1, \ldots, K$. For notational convenience, we denote the spiking activity associated with time window $k$ by a vector $\mathbf{n}_k := [n_{(k-1)W+1}, n_{(k-1)W+2}, \ldots, n_{kW}]'$, the CIF vector by $\boldsymbol{\lambda}_k := [\lambda_{(k-1)W+1}, \lambda_{(k-1)W+2}, \ldots, \lambda_{kW}]'$, and the matrix of covariates by $\mathbf{X}_k := [\mathbf{x}_{(k-1)W+1}, \mathbf{x}_{(k-1)W+2}, \ldots, \mathbf{x}_{kW}]'$ with rows corresponding to the covariate vectors. We assume that the parameter vectors $\{\boldsymbol{\omega}_k\}_{k=1}^{K}$ are sparse.

To capture the adaptivity manifested in the spiking dynamics, we use the forgetting factor mechanism of RLS algorithms (47) and combine the data log-likelihoods up to time $k$ using an exponential weighting scheme (43):

$$\ell_k^\beta(\boldsymbol{\omega}) := (1-\beta) \sum_{i=1}^{k} \beta^{k-i} \left(\mathbf{n}'_i \mathbf{X}_i \boldsymbol{\omega} - \mathbf{1}'_W \log(\mathbf{1}_W + \exp(\mathbf{X}_i \boldsymbol{\omega}))\right),$$

$$[2]$$

where $\boldsymbol{\omega}$ denotes a generic parameter vector, $0 < \beta \leq 1$ is the forgetting factor parameter, and $\mathbf{1}_W$ is the vector of all-ones

of length $W$. When applied to vectors, the functions $\log(\cdot)$ and $\exp(\cdot)$ are understood to act in an elementwise fashion.

The parameter vectors $\boldsymbol{\omega}_k$ for $k = 1, 2, \ldots, K$ can be efficiently estimated from the data using a sparse adaptive point process filter, referred to as $\ell_1$–PPF$_1$ (43). The $\ell_1$–PPF$_1$ algorithm estimates the sparse time-varying parameter vectors from point process observations in an online fashion by recursively maximizing a sequence of $\ell_1$-regularized exponentially weighted log-likelihoods via a proximal algorithm:

$$\widehat{\boldsymbol{\omega}}_k = \underset{\boldsymbol{\omega}_k}{\operatorname{argmax}} \left\{ \ell_k^\beta(\boldsymbol{\omega}_k) - \gamma \|\boldsymbol{\omega}_k\|_1 \right\}, \qquad [3]$$

where $\gamma \geq 0$ is a regularization parameter and can be selected analytically or through cross-validation (43). Statistical confidence regions for the estimates $\widehat{\boldsymbol{\omega}}_k$ can be computed using a recursive nodewise regression procedure (43). Throughout the rest of paper, we use the $\ell_1$–PPF$_1$ algorithm for adaptive parameter estimation.

**The AGC Measure.** Consider simultaneous spike recordings from an ensemble of $C$ neurons indexed by $c = 1, 2, \cdots, C$, denoted by $\{\{n_t^{(c)}\}_{t=1}^T\}_{c=1}^C$ over the time bins $t = 1, \cdots, T$. At time $t$, the spiking statistics of each neuron $(c)$ are modeled via the CIF formulation of Eq. 1 using a sparse modulation parameter vector $\boldsymbol{\omega}_t^{(c)} = [\mu_t^{(c)}, \boldsymbol{\omega}_t^{(c,1)\prime}, \boldsymbol{\omega}_t^{(c,2)\prime}, \cdots, \boldsymbol{\omega}_t^{(c,C)\prime}, \boldsymbol{\theta}_t^{(c)}]\prime$ consisting of a scalar baseline firing parameter $\mu_t^{(c)}$, a collection of sparse history dependence parameter vectors $\{\boldsymbol{\omega}_t^{(c,\tilde{c})}\}_{\tilde{c}=1}^C$ of size $M_H$, in which $\boldsymbol{\omega}_t^{(c,\tilde{c})}$ represents the contribution of the spiking history of neuron $(\tilde{c})$ to the CIF of neuron $(c)$, and $\boldsymbol{\theta}_t^{(c)}$ accounts for the stimulus modulation vector (e.g., receptive field). Let $h_{t,i}^{(c)} := \sum_{j=t-1-b_i}^{t-1-b_{i-1}} n_j^{(c)}$ be the spike count of neuron $(c)$ within the $i$-th spike counting window of length $W_{H,i}$, where $b_i := \sum_{j=1}^i W_{H,j}$ for $i = 1, 2, \cdots, M_H$ and $b_0 = 0$. The covariates associated with the ensemble activity are given by $\mathbf{x}_t := [1, \mathbf{h}_t^{(1)\prime}, \mathbf{h}_t^{(2)\prime}, \cdots, \mathbf{h}_t^{(C)\prime}, \mathbf{s}_t\prime]\prime$, where $\mathbf{h}_t^{(c)} := [h_{t,1}^{(c)}, h_{t,2}^{(c)}, \cdots, h_{t,M_H}^{(c)}]\prime$ denotes the history of spike counts of neuron $(c)$ within nonoverlapping windows of $W_H = [W_{H,1}, \ldots, W_{H,M_H}]$ up to a lag of $L_H := \sum_{i=1}^{M_H} W_{H,i}$, and $\mathbf{s}_t \in \mathbb{R}^{M_s}$ is the vector of neural stimuli in effect at bin $t$. We refer to this model, where the history of all of the neurons in the ensemble is taken into account, as the full model. Fig. 1 shows an example of the neuronal ensemble and the corresponding covariates for $C = 3$.

To assess the G-causal influences, a likelihood-based GC measure has been proposed in ref. 40 for point process models. Consider neuron $(c)$ as the target neuron with an observation
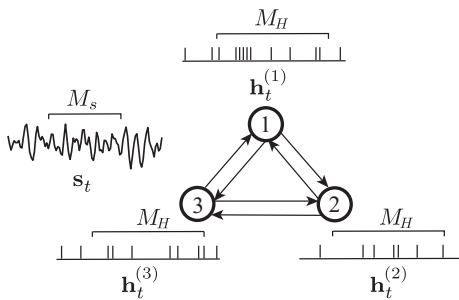
vector $\mathbf{n}^{(c)} := [n_1^{(c)}, n_2^{(c)}, \cdots, n_T^{(c)}]\prime$. Let $\mathcal{H}^{(c)}$ denote the history of the covariates of neuron $(c)$. The parameter vector and covariate history of neuron $(c)$ after excluding the effect of neuron $(\tilde{c})$ are denoted by $\boldsymbol{\omega}^{(c \smallsetminus \tilde{c})}$ and $\mathcal{H}^{(c \smallsetminus \tilde{c})}$, respectively, and compose the so-called reduced model. The log-likelihood ratio statistic associated with the G-causal influence of neuron $(\tilde{c})$ on neuron $(c)$ can be defined as

$$\mathcal{F}^{(\tilde{c} \mapsto c)} := s(\widehat{\boldsymbol{\omega}}^{(c,\tilde{c})}) \log \frac{\mathcal{L}(\widehat{\boldsymbol{\omega}}^{(c)} | \mathbf{n}^{(c)}, \mathcal{H}^{(c)})}{\mathcal{L}(\widehat{\boldsymbol{\omega}}^{(c \smallsetminus \tilde{c})} | \mathbf{n}^{(c)}, \mathcal{H}^{(c \smallsetminus \tilde{c})})}, \qquad [4]$$

where $\mathcal{L}(\widehat{\boldsymbol{\omega}} | \mathbf{n}, \mathcal{H})$ denotes the likelihood of estimated parameter vector $\widehat{\boldsymbol{\omega}}$ given the observation sequence $\mathbf{n}$ and the history of the covariates $\mathcal{H}$, and $s(\boldsymbol{\omega}) := \operatorname{sign}(\sum_l \omega_l)$. Based on this formulation, the GC effect from neuron $(\tilde{c})$ to neuron $(c)$ can be measured as the reduction in the point process log-likelihood of neuron $(c)$ in the reduced model as compared with the full model. Note that the signum function determines the effective excitatory or inhibitory nature of this influence.

Most existing formulations of GC leverage the MVAR modeling framework (20–29, 31, 35), which pertains to data with linear Gaussian statistics. The GC measure in Eq. 4, however, benefits from the likelihood-based inference methodology and covers a wide range of complex statistical models. Both the MVAR-based GC measure and its log-likelihood-based point process variant of ref. 40 assume that the underlying time series are stationary (i.e., the modulation parameters are all static). In many scenarios of interest, however, the underlying dynamics exhibit nonstationarity. An example of such a scenario is the task-dependent receptive field plasticity phenomenon (43, 48, 49). In addition, ML estimation used by these techniques does not capture the underlying sparsity of the parameters and often exhibits poor performance, when the data length is short or the number of neurons $C$ is large.

To account for possible changes in the ensemble parameters and their underlying sparsity, we introduce the AGC measure, which is capable of capturing the dynamics of G-causal influences in the ensemble. To this end, we make two major modifications to the classical GC measure. First, we leverage the exponentially weighted log-likelihood formulation of Eq. 2 to induce adaptivity into the GC measure. Second, we exploit the possible sparsity of the ensemble parameters. Replacing the standard data log-likelihoods in Eq. 4 by their sparse adaptive counterparts given in Eqs. 2 and 3, we define the AGC measure from neuron $(\tilde{c})$ to neuron $(c)$ at time window $k$ as

$$\mathcal{F}_{k,\beta}^{(\tilde{c} \mapsto c)} := s_k(\widehat{\boldsymbol{\omega}}_k^{(c,\tilde{c})}) \left( \ell_k^\beta(\widehat{\boldsymbol{\omega}}_k^{(c)}) - \ell_k^\beta(\widehat{\boldsymbol{\omega}}_k^{(c \smallsetminus \tilde{c})}) \right). \qquad [5]$$

Although these modifications bring about crucial advantages in capturing the functional network dynamics in a robust fashion, they require construction of a statistical inference framework in order for the proposed AGC measure to be useful. We address these issues in the forthcoming section.

**Statistical Inference of the AGC Measure.** Due to the stochastic and often biased nature of GC estimates, nonzero values of GC do not necessarily imply existence of G-causal influences. Hence, a statistical inference framework is required to assess the significance of the extracted G-causal interactions.

Consider two nested GLM models, referred to as full and reduced models, with parameters $\boldsymbol{\omega}^{(F)} := \boldsymbol{\omega}^{(c)}$ and $\boldsymbol{\omega}^{(R)} = \boldsymbol{\omega}^{(c \smallsetminus \tilde{c})}$, respectively, in which the latter is a special case of the former. To assess the significance of a GC link, one can test for the null hypothesis $H_0 : \boldsymbol{\omega} = \boldsymbol{\omega}^{(R)}$ against the alternative $H_1 : \boldsymbol{\omega} = \boldsymbol{\omega}^{(F)}$. The commonly used test statistic is referred



**Fig. 1.** An example of the neuronal ensemble model for $C = 3$ neurons and $W_{H,i} = 1, \forall i$. The CIF of neuron (2) can be expressed as $\lambda_t^{(2)} \Delta = \operatorname{logit}^{-1}(\mu_t^{(2)} + \boldsymbol{\omega}_t^{(2,1)\prime} \mathbf{h}_t^{(1)} + \boldsymbol{\omega}_t^{(2,2)\prime} \mathbf{h}_t^{(2)} + \boldsymbol{\omega}_t^{(2,3)\prime} \mathbf{h}_t^{(3)} + \boldsymbol{\theta}_t^{(2)\prime} \mathbf{s}_t)$.

to as the deviance difference of the two models and is defined as $D(\widehat{\boldsymbol{\omega}}^{(F)}; \widehat{\boldsymbol{\omega}}^{(R)}) := 2(\ell(\widehat{\boldsymbol{\omega}}^{(F)}) - \ell(\widehat{\boldsymbol{\omega}}^{(R)}))$, where $\ell(\cdot)$ is the log-likelihood and $\widehat{\boldsymbol{\omega}}^{(F)}$ and $\widehat{\boldsymbol{\omega}}^{(R)}$ denote the parameter estimates under the full and reduced models, respectively. The deviance difference for the likelihood-based GC is twice the right-hand side of Eq. **4**, modulo the signum function.

To perform the foregoing hypothesis test, the distributions of the deviance difference under the two hypotheses need to be characterized. Although these distributions are known for the classical GC measure (50–52), they cannot be readily extended to our AGC measure for two main reasons. First, the log-likelihoods are replaced by their exponentially weighted counterparts, which suppresses their dependence on the data length $N$ due to the forgetting factor mechanism. Second, unlike ML estimates, which are asymptotically unbiased, the $\ell_1$-regularized ML estimates are biased and hence violate the common asymptotic normality assumptions.

To address these challenges, inspired by recent results in high-dimensional regression (53, 54), we define the adaptive de-biased deviance as

$$D_{k,\beta}(\widehat{\boldsymbol{\omega}}_k; \boldsymbol{\omega}_k) := \frac{1+\beta}{1-\beta}\Big( 2\big(\ell_k^\beta(\widehat{\boldsymbol{\omega}}_k) - \ell_k^\beta(\boldsymbol{\omega}_k)\big) \\ - \dot{\ell}_k^\beta(\widehat{\boldsymbol{\omega}}_k)'\ddot{\ell}_k^\beta(\widehat{\boldsymbol{\omega}}_k)^{-1}\dot{\ell}_k^\beta(\widehat{\boldsymbol{\omega}}_k)\Big), \quad [6]$$

where $\dot{\ell}_k^\beta(\cdot)$ and $\ddot{\ell}_k^\beta(\cdot)$ are the gradient vector and Hessian matrix of the exponentially weighted log-likelihood function $\ell_k^\beta(\cdot)$, and $\boldsymbol{\omega}_k$ and $\widehat{\boldsymbol{\omega}}_k$ denote the true and estimated parameter vector at time window $k$, respectively. The adaptive de-biased deviance is composed of two main terms: The first term is twice the exponentially weighted log-likelihood ratio statistic, which is analogous to the standard deviance difference, whereas the second is a bias correction term. The bias correction term compensates for the effect of the $\ell_1$-regularization bias imposed in favor of enforcing sparsity in the estimate $\widehat{\boldsymbol{\omega}}_k$. The effect of forgetting factor mechanism appears in the form of the scaling $(1+\beta)/(1-\beta)$. Finally, we define a test statistic referred to as adaptive de-biased deviance difference:

$$D_{k,\beta}^{(\tilde{c} \mapsto c)} := D_{k,\beta}(\widehat{\boldsymbol{\omega}}_k^{(c)}; \boldsymbol{\omega}_k^{(c)}) - D_{k,\beta}(\widehat{\boldsymbol{\omega}}_k^{(c \smallsetminus \tilde{c})}; \boldsymbol{\omega}_k^{(c \smallsetminus \tilde{c})}). \quad [7]$$

In what follows, we will mainly work with $D_{k,\beta}^{(\tilde{c} \mapsto c)}$, as opposed to its biased version given by $\mathcal{F}_{k,\beta}^{(\tilde{c} \mapsto c)}$ in Eq. **5**. Note that $\mathcal{F}_{k,\beta}^{(\tilde{c} \mapsto c)} = \frac{1}{2} s_k(\widehat{\boldsymbol{\omega}}_k^{(c,\tilde{c})}) \Big( \frac{D_{k,\beta}^{(\tilde{c} \mapsto c)}}{1+\beta} + B_{k,\beta}^{(\tilde{c} \mapsto c)} \Big)$, where $B_{k,\beta}^{(\tilde{c} \mapsto c)}$ is the difference of the bias terms of the full and reduced models.

There are four major challenges in inferring a GC influence from $D_{k,\beta}^{(\tilde{c} \mapsto c)}$: (*i*) efficient computation of $D_{k,\beta}^{(\tilde{c} \mapsto c)}$ from the data, (*ii*) determining the distribution of $D_{k,\beta}^{(\tilde{c} \mapsto c)}$ under the absence and presence of a GC link, (*iii*) controlling the false discovery rate (FDR), and (*iv*) assessing the significance of detected GC links. We will address these challenges in the

remainder of this section. Fig. 2 shows a schematic depiction of the overall inference procedure, which we will discuss next.

***(i) Recursive computation of the AGC.*** The computation of the adaptive de-biased deviance differences $D_{k,\beta}^{(\tilde{c} \mapsto c)}$ for all of the possible $|\mathcal{C}|$ links and at all times $k$ is required for our statistical analysis. Therefore, in order for the analysis to scale favorably with the network size $C$ and the data length $K$, it is crucial to develop an efficient framework for the computation of the AGC measure. The RLS-inspired exponential weighting of the log-likelihoods in Eq. **2** indeed paves the way for the recursive computation of the AGC measure. The recursive procedure for computing $\ell_k^\beta(\widehat{\boldsymbol{\omega}}_k)$ for a generic estimate $\widehat{\boldsymbol{\omega}}_k$ is given in *Algorithm S3* in *SI Appendix*, section 3, from which the AGC measure of Eq. **5** can be computed. This step comprises the recursive AGC computation block in Fig. 2.

***(ii) Asymptotic distributional analysis of the AGC.*** Let $M^{(F)}$ and $M^{(R)}$ denote the dimensions of $\boldsymbol{\omega}^{(c)}$ and $\boldsymbol{\omega}^{(c \smallsetminus \tilde{c})}$, respectively. In *Theorem S1* in *SI Appendix*, section 1 we establish the following result under mild technical conditions: *As $\beta \to 1$,*

i)  *in the absence of a GC link from* $(\tilde{c})$ *to* $(c)$, $D_{k,\beta}^{(\tilde{c} \mapsto c)} \to \chi^2(M^{(d)})$, *and*

ii)  *in the presence of a GC link from* $(\tilde{c})$ *to* $(c)$, *if the corresponding cross-history coefficients scale at least as* $\mathcal{O}\left(\sqrt{\frac{1-\beta}{1+\beta}}\right)$, *then*
$$D_{k,\beta}^{(\tilde{c} \mapsto c)} \to \chi^2\big(M^{(d)}, \nu_k^{(\tilde{c} \mapsto c)}\big),$$

*where* $M^{(d)} := M^{(F)} - M^{(R)}$ *is the dimensionality difference of the two nested models, and* $\nu_k^{(\tilde{c} \mapsto c)} > 0$ *is the corresponding noncentrality parameter at time window $k$.*

*Theorem S1* has two main implications. First, it establishes that our proposed adaptive de-biased deviance difference statistic admits simple asymptotic distributional characterization. Given that these asymptotic distributions form the main ingredients of the forthcoming inference procedure, the second block in Fig. 2 serves to highlight the significance of adaptive de-biasing. As shown in *SI Appendix*, section 3, the bias $B_{k,\beta}^{(\tilde{c} \mapsto c)}$ can also be computed in a recursive fashion.

Second, given that for $\nu_k^{(\tilde{c} \mapsto c)} = 0$ the noncentral chi-squared distribution coincides with the chi-squared distribution, the noncentrality parameter plays a key role in separating the distributions under the null and alternative hypotheses: When the deviance difference is close to zero, the null hypothesis $H_0$ is likely to be true (i.e., no GC link). When the deviance difference is large, the alternative $H_1$ is likely to be true (i.e., a GC link exists) (See *Remark 2* in *SI Appendix*, section 1, for further discussion.) The noncentrality parameter $\nu_k^{(\tilde{c} \mapsto c)}$, however, is a complicated function of the true values of the parameters and cannot be directly observed. In the next two subsections, we initially assume that an estimate $\widehat{\nu}_k^{(\tilde{c} \mapsto c)}$ is at hand, and later on derive an algorithm for its estimation.

The output of the second block in Fig. 2 is the de-biased deviance differences corresponding to all pairs of neurons
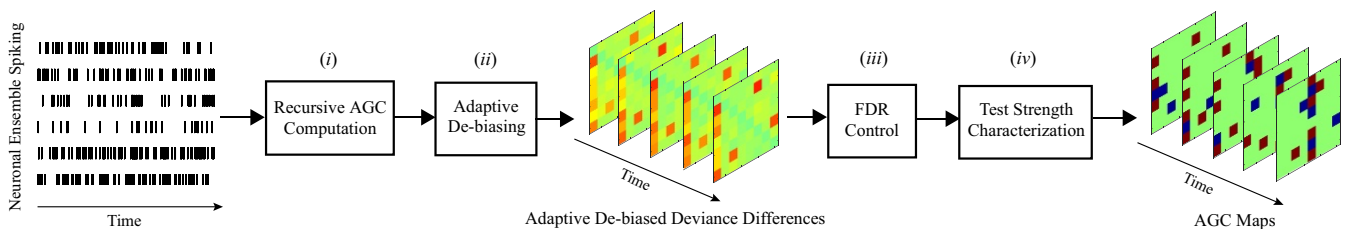


**Fig. 2.** Schematic depiction of the inference procedure for the AGC measure.

(shown in 2D as deviance difference maps). In the next two sub-sections we will show how to translate the deviance differences to statistically interpretable AGC links.

***(iii) FDR control.*** First, we use part *i* of the result of *Theorem S1* to control the FDR in a multiple hypothesis testing framework. To this end, we use the Benjamini–Yekutieli (BY) procedure (55). The BY procedure aims at controlling the FDR, which is the expected ratio of incorrectly rejected null hypotheses, or namely "false discoveries," at a desired significance level $\alpha$.

To identify significant GC interactions while avoiding spurious false positives, we conduct multiple hypothesis tests on the set of $|\mathcal{C}| := C \times (C-1)$ pairwise possible GC interactions $\mathcal{C} := \{(\tilde{c} \mapsto c) \mid \tilde{c}, c = 1, \ldots, C, \ c \neq \tilde{c}\}$ among the $C$ neurons at each time step $k$. The null hypothesis $H_{0,k}^{(\tilde{c} \mapsto c)}$ corresponds to lack of a GC link from neuron $(\tilde{c})$ to $(c)$ at time step $k$. Thus, rejection of the null hypothesis amounts to discovering a GC link $(\tilde{c} \mapsto c)$ at time step $k$. We first compute $D_{k,\beta}^{(\tilde{c} \mapsto c)}$ for all possible links in $\mathcal{C}$. Based on *Theorem S1*, under null hypothesis $H_{0,k}^{(\tilde{c} \mapsto c)}$ we have $D_{k,\beta}^{(\tilde{c} \mapsto c)} \to \chi^2(M^{(d)})$ as $\beta \to 1$. Hence, by virtue of convergence in distribution, for $\beta$ close to 1, thresholding the test statistic results in a consistent approximation to limiting the false positive rate: $H_{0,k}^{(\tilde{c} \mapsto c)}$ is rejected at a confidence level of $1-\alpha$, if $D_{k,\beta}^{(\tilde{c} \mapsto c)} > \mathrm{F}_{\chi^2(M^{(d)})}^{-1}(1-\alpha)$, where $\mathrm{F}_{\chi^2(M^{(d)})}^{-1}(\cdot)$ is the inverse CDF of a $\chi^2$ distribution with $M^{(d)}$ degrees of freedom. Using the BY procedure, we can thus control the mean FDR at a rate of $\bar{\alpha} := \frac{(|\mathcal{C}|+1)\alpha}{2|\mathcal{C}| \log |\mathcal{C}|}$ for all tests (Fig. 2, third block). *Algorithm S1* in *SI Appendix, section 2*, summarizes this procedure.

***(iv) Test strength characterization via noncentral $\chi^2$ filtering and smoothing algorithm.*** Next, we use part *ii* of the result of *Theorem S1* to assess the significance of the tests for the detected GC links. Under the alternative hypothesis, *Theorem S1* implies that $H_{1,k}^{(\tilde{c} \mapsto c)} : D_{k,\beta}^{(\tilde{c} \mapsto c)} \to \chi^2\big(M^{(d)}, \nu_k^{(\tilde{c} \mapsto c)}\big)$ as $\beta \to 1$. Hence, by virtue of convergence in distribution, the false negative rate can be estimated by $\eta_k^{(\tilde{c} \mapsto c)} := \mathrm{F}_{\chi^2(M^{(d)}, \widehat{\nu}_k)}\big(\mathrm{F}_{\chi^2(M^{(d)})}^{-1}(1-\alpha)\big)$, at a confidence level of $1-\alpha$, where $\mathrm{F}_{\chi^2(M^{(d)}, \widehat{\nu}_k)}(\cdot)$ represents the CDF of a noncentral $\chi^2$ distribution with $M^{(d)}$ degrees of freedom and the estimate $\widehat{\nu}_k$ of the noncentrality parameter $\nu_k^{(\tilde{c} \mapsto c)}$. To quantify the significance of an estimated GC link, we use the Youden's $J$-statistic, which is an effective measure often used for summarizing the overall performance of a diagnostic test. The $J$-statistic in our setting is given by

$$J_k^{(\tilde{c} \mapsto c)} := 1 - \alpha - \mathrm{F}_{\chi^2(M^{(d)}, \widehat{\nu}_k)}\big(\mathrm{F}_{\chi^2(M^{(d)})}^{-1}(1-\alpha)\big), \qquad \textbf{[8]}$$

for a fixed significance level $\alpha$. Note that the $J$-statistic can take values in $[0,1]$. The case of $J_k^{(\tilde{c} \mapsto c)}$ being close to one represents high sensitivity and specificity of the test statistic, which coincides with large values of noncentrality. One advantage of the $J$-statistic over the conventional $P$ value is that it accounts for both type I and type II errors. In the context of GC analysis, the $J$-statistic for each possible link can serve as a normalized indicator of how reliable the detected link is. For consistency, we assign a value of $J_k^{(\tilde{c} \mapsto c)} = 0$, when the null hypothesis is not rejected.

It remains to estimate the unknown noncentrality parameters $\nu_k^{(\tilde{c} \mapsto c)}$ given the observed deviance differences $D_{k,\beta}^{(\tilde{c} \mapsto c)}$. Under the assumption that $\nu_k^{(\tilde{c} \mapsto c)}$ changes smoothly in time, this can be carried out efficiently by a noncentral $\chi^2$ filtering and smoothing algorithm, which is given by *Algorithm S2* in *SI Appendix, section 2*. Given these estimates, the

$J$-statistics for the rejected nulls can be computed at $\bar{\alpha}$ (Fig. 2, fourth block), as summarized in *Algorithm S1* in *SI Appendix, section 2*.

---

***Algorithm 1*: AGC Inference from Ensemble Neuronal Spiking**

   **Input:** Spike trains $\{\{n_t^{(c)}\}_{t=1}^T\}_{c=1}^C$ and parameters $\Theta$.

1. **for** $c, \tilde{c} = 1, \ldots, C, \tilde{c} \neq c$ **do**
2. Recursively estimate the sparse time-varying modulation parameter vectors $\{\widehat{\omega}_k^{(c)}\}_{k=1}^K$ and $\{\widehat{\omega}_k^{(c \setminus \tilde{c})}\}_{k=1}^K$ corresponding to full and reduced GLMs using $\ell_1$–$\mathrm{PPF}_1$ (43),
3. Recursively compute the adaptive de-biased deviance differences $\{D_k^{(\tilde{c} \mapsto c)}\}_{k=1}^K$ (*Algorithm S3*),
4. Perform noncentral $\chi^2$-squared filtering and smoothing to estimate the noncentrality parameters $\{\widehat{\nu}_k^{(\tilde{c} \mapsto c)}\}_{k=1}^K$ from $\{D_{k,\beta}^{(\tilde{c} \mapsto c)}\}_{k=1}^K$ (*Algorithm S2*),
5. **for** $k = 1, \ldots, K$ **do**
6. Apply BY rejection rule to the ensemble set of GC tests to control FDR at rate $\alpha$ (*Algorithm S1*),
7. Compute AGC maps $\widehat{\Phi}_k \in [-1,1]^{C \times C}$ based on the $J$-statistics as $(\widehat{\Phi}_k)_{c,\tilde{c}} := s_k\big(\widehat{\omega}_k^{(c,\tilde{c})}\big) J_k^{(\tilde{c} \mapsto c)}$ (*Algorithm S1*).

   **Output:** AGC maps $\{\widehat{\Phi}_k\}_{k=1}^K$.

---

**Summary of Advantages over Existing Work.** *Algorithm 1* summarizes the overall AGC inference procedure. Choices of the parameters $\Theta$ involved in *Algorithm 1* and its computational complexity are discussed in *SI Appendix, sections 4–6*. Before presenting applications to synthetic and real data, we summarize the advantages of our methodology over existing work:

  *i)* Sparse dynamic GLM modeling provides more accurate estimates of the parameters (43), and hence more reliable detection of the GC links, as compared with existing static methods based on ML. We examine this aspect of our methodology in *SI Appendix, section 8*, using an illustrative simulation study;
 *ii)* Relating the noncentrality parameters to the test strengths of the detected GC links is not used by existing techniques. In light of *Theorem S1* and the need for estimating the noncentrality parameters, we devised a noncentral $\chi^2$ filtering and smoothing algorithm to exploit the entire observed data for obtaining reliable estimates;
*iii)* Exponential weighting of the log-likelihoods admits construction of adaptive filters for estimating the network parameters in a recursive fashion, which significantly reduces the computational complexity of our inference procedure; and
*iv)* Characterization of AGC via the $J$-statistic as a normalized measure of hypothesis test strength for each detected GC link can be further used for graph-theoretic analysis of the inferred functional networks. By viewing the $J$-statistic as a surrogate for link strength, the AGC networks can be refined by thresholding the $J$-statistics, and access to the distribution of the $J$-statistics in a network allows one to perform further hypothesis tests regarding the network function (56).

In the next section, we illustrate these advantages by comparing our methodology with two representative techniques for inferring functional network dynamics.

## Applications

**A Simulated Example.** We consider a network of $C = 8$ functionally interconnected neurons whose connectivity pattern evolves in time. As illustrated in Fig. 3*A*, the network dynamics undergo three main states, each covering one-third (40 s) of the simulation period: (*i*) the first static state, where neuron (1) plays a dominant role, causally influencing all other neurons; (*ii*) the

**Fig. 3.** Functional network dynamics inference from simulated spikes. (*A*) Three states of the functional network evolution, where neurons (vertices) are interacting through static (solid edges) or dynamic (dashed edges) causal links of inhibitory (open circles) or excitatory (filled circles) nature. (*B*) One realization of simulated spikes within windows of 1 s selected at $\{40, 60, 120\}$ s. (*C*) Estimated noncentrality $\widehat{\nu}_k$ across time corresponding to four selected GC links (color-coded in network map), along with the shifted deviance differences $D_k - M^{(d)}$ (black traces) and the 95% confidence regions for each estimated trace $\widehat{\nu}_k$ (colored traces). (*D*) Four panels of estimated $J$-statistics $J_k$ corresponding to the selected GC links. (*E*) Performance comparison of the causal inference methods: (*i*) the proposed AGC method (second row), (*ii*) the static GC method in ref. 40 (third row), and (*iii*) the functional connectivity method in ref. 38 (last row), along with the true causality maps (first row). Each panel represents the estimated $8 \times 8$ causality map at a specific time.

intermediate dynamic state, where neuron (1) loses the dominant role to neuron (5), as its causal influences smoothly decay, while a new set of causal interactions from neuron (5) to all of the other neurons emerge; and (*iii*) the final static state, where all of the causal links from neuron (1) are completely vanished and the links from neuron (5) are stabilized. The network also comprises three static causal links, for example, $(3 \mapsto 7)$, which remain constant throughout.

An observation period of $\mathcal{T} = 120$ s is discretized to $K = 120$ k bins of length $\Delta = 1$ ms. We use a point process model with Bernoulli spiking statistics to generate the binary spike trains for all neurons, where the CIF is modeled given the dynamic GLM of Eq. **1**. The details of the parameter selection and estimation are given in *SI Appendix*, section 5. Fig. 3*B* shows a realization of the simulated spike trains indicated by black vertical lines for all eight neurons within three sample windows of length 1 s, with end points at $\{40, 60, 120\}$ s, selected from

the three segments of the simulation. Note that the G-causal pattern of Fig. 3*A* is unknown to the estimator, and is to be inferred from the simulated spike trains. Fig. 3*C* shows the time course of the estimated noncentrality parameters $\widehat{\nu}_k$ and their 95% confidence intervals obtained by the noncentral $\chi^2$ filtering and smoothing algorithm associated with four selected GC links: (*i*) $(1 \mapsto 4)$ a dynamic weakening GC link (red), (*ii*) $(5 \mapsto 2)$ a dynamic strengthening GC link (blue), (*iii*) $(8 \mapsto 6)$ a static link (green), and (*iv*) $(8 \mapsto 2)$ a nonexisting GC link (magenta). Black traces show the shifted observed deviances $D_k - M^{(d)}$. Fig. 3*D* represents the time course of the estimated $J$-statistics plotted in four separate panels, where the FDR is controlled at a rate $\alpha = 0.1$.

In Fig. 3*C*, the estimates of $\widehat{\nu}_k$ corresponding to the three existing GC links take significant values, correctly identifying the G-causal interactions, while $\widehat{\nu}_k$ takes values close to zero for the nonexisting link, implying no significant G-causal interaction. The time course of changes for both dynamic links and the static link is closely tracked by the noncentrality parameters, albeit with an apparent delay. This delay is due to the choice of the effective window length and highlights the trade-off between estimation accuracy and delay. While it is possible to reduce this delay by choosing smaller effective windows, for the sake of accuracy of parameter estimation, and thereby robust detection of the AGC links, we have chosen the effective window length to be 10 s (a fraction of the 40-s transition period) to incur a tolerable delay. The aforementioned performance is echoed in the test strengths quantified by the $J$-statistics shown in Fig. 3*D*. Even though the noncentrality parameters in Fig. 3C track the changes of the network parameters much faster, the $J$-statistics may lag behind due to the conservative statistical thresholds set by the FDR control procedure. By choosing a higher FDR level, the $J$-statistics will capture the changes much faster, but at the expense of possibly more false discoveries. It is noteworthy that our proposed method distinguishes the direct GC links from the indirect ones, as it correctly detects the direct GC links $(8 \mapsto 6)$ and $(6 \mapsto 2)$ but rejects the existence of the corresponding indirect link $(8 \mapsto 2)$.

The top row in Fig. 3*E* shows the ground truth G-causal maps plotted at nine time instances (three per segment). Each map $\Phi$ represents an $8 \times 8$ color-coded array showing the excitatory, inhibitory, and no-GC links in red, blue, and green, respectively. The AGC maps estimated by our method are shown in the second row, where each entry $(\widehat{\Phi}_k)_{c,\tilde{c}}$ represents the $J$-statistic $J_k^{(\tilde{c} \mapsto c)}$ of the estimated GC link $(\tilde{c} \mapsto c)$, where the excitatory or inhibitory nature of the links is determined by the sign of the AGC measure, accounting for the aggregate cross-history contribution, and is not indicative of the morphological identity of the connections.

We compare the AGC maps with two other methods: the static GC method of ref. 40 (third row), and the functional connectivity analysis of ref. 38 (final row). To adapt these methods to the time-varying setting, we used nonoverlapping window segments whose length is chosen to match the effective window length $N_{eff} := \frac{W}{1-\beta}$ of our method. Within each window, the signed binary functional connectivity is estimated using the methods outlined in refs. 38 and 40. The true model order and the same significance level are used for all methods (see *SI Appendix*, section 5, for more details). Fig. 3*E* (last two rows) shows the connectivity maps obtained by refs. 38 and 40. On a qualitative level, Fig. 3*E* reveals the favorable performance of our proposed framework in terms of both identification and tracking of the GC influences. The method of ref. 40 results in both high false positive and false negative errors and fails to track the GC dynamics due to highly variable parameter estimates. Similarly, the method of ref. 38 shows poor false positive rejection and tracking performance.
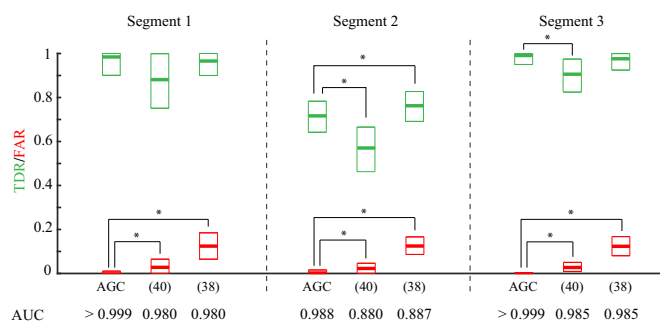
**Fig. 4.** Performance comparison of AGC inference with the methods of refs. 40 and 38 in terms of TDR (green) and FAR (red) for the three segments of the simulation period. Boxes represent the mean and 90% confidence intervals. Asterisks indicate significant difference with effect size of $r \geq 0.8$ (Wilcoxon signed-rank test, $P < 0.001$). AUC values corresponding to the ROC performance of the three algorithms are reported at the bottom (see *SI Appendix*, section 7 for more details).

To quantify the foregoing performance comparison, we repeated the simulation for $R = 500$ realizations of spike trains randomly generated based on the network dynamics in Fig. 3*A*. Fig. 4 represents the performance results in terms of true detection rate (TDR) and false alarm rate (FAR), which are shown in green and red, respectively. Boxes represent the mean and 90% confidence intervals. Based on the Wilcoxon signed-rank test with $P < 0.001$, our method has a significantly lower FAR compared with both ref. 38 (effect sizes of $r = 1$ for all segments) and ref. 40 ($r = 0.8, 0.86$ and $0.98$ for the three segments, respectively). Our achieved TDRs are also significantly higher than those of ref. 40 ($r = 0.73, 0.996$ and $0.94$ for the three segments, respectively) and are only outperformed by ref. 38 in the middle segment ($r = 0.19, 0.86$ and $0.27$ for the three segments, respectively). By varying the significance levels for all of the algorithms, we also evaluated their receiver operating characteristic (ROC) performance, whose area under the curve (AUC) values for the three segments are indicated in Fig. 4. Further details on the comparison procedure as well as assessing the robustness of our method to the choice of parameters are given in *SI Appendix*, section 7. It is noteworthy that our method is the only one with consistently low FAR ($< 1\%$), while maintaining high TDR. Finally, both methods in refs. 38 and 40 output binary connectivity maps, as opposed to AGC, which provides normalized continuous-valued test strengths of the detected GC links. In the spirit of easing reproducibility, we have archived a MATLAB implementation that fully generates Fig. 3 on GitHub (https://github.com/Arsha89/AGC_Analysis).

**Application to Real Data: Spontaneous Activity in the Mouse Auditory Cortex.** In this section, we apply our proposed method to experimentally recorded neuronal population data from the mouse auditory cortex. We imaged the spontaneous activity in the auditory cortex of an awake mouse with in vivo two-photon calcium imaging (see *SI Appendix*, section 13 for details of the experimental procedures). Within an imaged field of view, the activity of $N_{cells} = 219$ neurons is recorded at a sampling rate of $f_s \approx 30$ Hz for a total duration of $T \approx 22$ min. Spike trains are inferred from the fluorescence traces using the constrained-foopsi technique (57). For GC inference, a subset of $C = 20$ neurons exhibiting high spiking activity were selected, as many of the neurons in the ensemble are relatively silent. The FDR is controlled at a rate of $\alpha = 0.005$ for testing the $|\mathcal{C}| = 380$ possible GC links. Fig. 5 *A* and *B* show the time course of the noncentrality estimates and $J$-statistics for four selected candidate GC links, respectively. These representative GC links consist of a persis-

tent link $(6 \mapsto 4)$ (blue), two transient links $(1 \mapsto 18)$ (red) and $(2 \mapsto 12)$ (green), and an insignificant link $(8 \mapsto 9)$ (magenta). Fig. 5 *C* and *D* show four snapshots of the AGC map estimates, respectively in the matrix form and as a network overlaid on the slice, at time-stamps $\{8.33, 11.66, 16.66, 22.22\}$ min. Other than the three color-coded significant links, the rest of the detected G-causal links are indicated by black arrows.

The detected G-causal maps are considerably sparse (maximum $\sim 16$ out of 380 possible links), with a few persistent GC links and a multitude of transient links emerging and vanishing over time (Movies S1 and S2). The sparsity of the AGC maps is consistent with sparse activity in auditory cortex (58). A careful inspection of the spatial pattern of the AGC links reveals that the detected links correspond to distances in the range of $[150, 200]$ μm. These distances are consistent with in vitro measurements of the spatial patterns of intralaminar connectivity within the mouse auditory cortex (59), showing a significant peak in the connection probability within the mean radial range of $[120, 200]$ μm. These results indicate that the proposed AGC method is able to detect underlying connectivity patterns among neurons.

**Application to Real Data: Ferret Cortical Activity During Active Behavior.** Studies of the PFC have revealed its association with high-level executive functions such as decision making and attention (60–62). In particular, recent findings suggest that PFC is engaged in cognitive control of auditory behavior (62), through a top-down feedback to sensory cortical areas, resulting in enhancement of goal-directed behavior. It is conjectured in ref. 63 that the top-down feedback from PFC triggers adaptive changes in the receptive field properties of A1 neurons during active attentive behavior, to facilitate the processing of task-specific stimulus features. This conjecture has been examined in the context of visual processing, where top-down influences exerted on the visual cortical pathways have been shown to alter the functional properties of cortical neurons (64, 65).
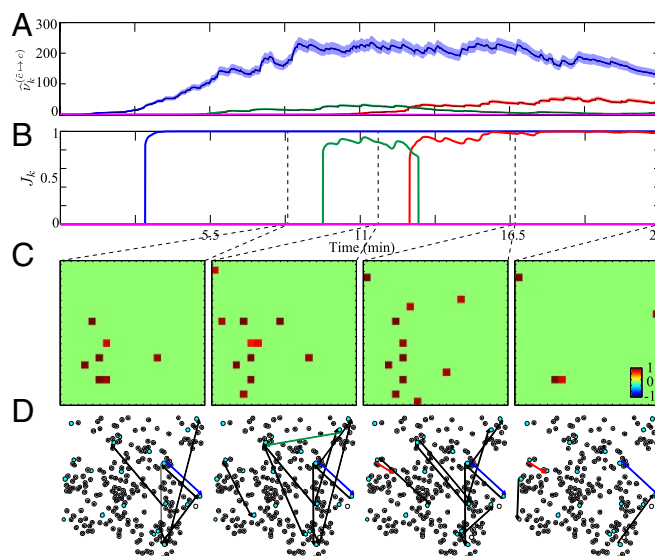


**Fig. 5.** Adaptive G-causal interactions among ensemble of neurons in mouse auditory cortex under spontaneous activity. The time course of estimated GC changes for four selected GC links obtained through (*A*) noncentrality parameter $\hat{\nu}_k$ and (*B*) $J$-statistics $J_k$. (*C*) AGC map estimates $\hat{\Phi}_k$ at four selected points in time, marked by the dashed vertical lines in the top panel. (*D*) network maps overlaid on the slice, showing cells with black circles and the selected cells highlighted in cyan. The detected GC links are depicted in black directed arrows and colored for the selected links.

To examine this conjecture at a single-unit level, we apply our proposed AGC inference method to single-unit spiking activities from an ensemble of neurons simultaneously recorded from two cortical regions of A1 and PFC in ferrets during a series of passive listening and active auditory task conditions. In this application, we sought to reveal the significant task-specific changes in the G-causal interactions within or between PFC and A1 regions at the single-unit level during active behavior. We used the spike data recordings from a large set of experiments (more than 35) conducted on three ferrets for GC inference analysis (data from the Neural Systems Laboratory, Institute for Systems Research, University of Maryland, College Park, MD). During each trial in an auditory discrimination task, the ferrets were presented with a random sequence of broadband noise-like acoustic stimuli known as temporally orthogonal ripple combinations (TORCs) along with randomized presentations of the target tone. Ferrets were trained to attend to the spectrotemporal features of the presented sounds and discriminate the tonal target from the background reference stimuli (see ref. 63 for details of the experimental procedures). Due to their broadband noise-like features, the TORCs and the corresponding neural responses admit efficient estimation of the spectrotemporal tuning of the primary auditory neurons via sparse regression (43, 66).

Fig. 6 shows our results on a selected experiment in which a total number of $C = 9$ single units were detected through spike sorting (five units in A1 and four units in PFC). The selected experiment consists of three main blocks: passive listening pre-task, active task, and passive listening post-task, composed of $R = 4, 4$, and 6 repetitions, respectively. Within each repetition, a complete set of 30 randomly permuted TORCs were presented along with a randomized presentation of the target tone at $f = 2.5$ kHz. Fig. 6A shows the activity of all of the units during the first repetition of each block, separated by vertical dashed lines. Fig. 6B shows the time courses of the inferred $J$-statistics, where each row represents the significant incoming GC links from all of the other units. Each unit and its significant outgoing GC links are color-coded uniquely as labeled on the right side of each panel. For brevity, the significant GC links that show a degree of persistence during at least one block of the experiment

are plotted. Fig. 6C depicts the representative network maps of the GC links among the nine units during the three main blocks, where each significant GC link from panel Fig. 6B is indicated by a directional link. Finally, Fig. 6D exhibits snapshots of the STRFs of all of the five A1 units, taken at the end point of each block. The red arrow marks the tonal target.

Three major task-specific dynamic effects can be inferred from Fig. 6: (i) a significant bottom-up GC link from the target-tuned A1 unit during active behavior, (ii) a persistent task-relevant top-down GC link, and (iii) task-relevant plasticity and rapid tuning changes within A1. First, unit 4 in A1 shows strong frequency selectivity to the target around $f = 2.5$ kHz during the whole experiment (vertical dashed lines, Fig. 6D). Moreover, its STRF dynamics reveal a plastic shift of the target-tuned facilitative regions to shorter latencies following the active attentive behavior (upward arrow, middle panel; see also Movie S3). Strikingly, a bottom-up GC link from the very same strongly target-tuned unit to PFC (red link, $4 \mapsto 8$) emerges during the active task (second row, Fig. 6B), temporally preceding any top-down significant GC link.

The second effect appears as a strong top-down GC link (green link, $8 \mapsto 3$) which builds up during the active auditory behavior and even persists during a few repetitions of the post-active condition (fifth row, Fig. 6B). The onset of this top-down GC link coincides with a dramatic and rapid change in the STRF of the A1 unit 3, which was initially tuned to $\sim 8$ kHz (downward arrow, Fig. 6D, Left) but eventually gets suppressed at this nontarget frequency (Fig. 6D, Middle) by getting G-causally influenced by the PFC unit 8. This effect reveals the relationship between the top-down network dynamics and the changes in the tuning of the A1 units. We examine the dynamics of the parameters of the foregoing bottom-up and top-down links in detail in *SI Appendix, section 10*, for further clarification. The third effect concerns the emergence and strengthening of frequency selectivity in some of the A1 units (e.g., units 1 and 2, Fig. 6D, Right) to the target tone, which alludes to a salient synaptic reinforcement effect within A1 during and after the active task.

In addition to these interregion GC links, multiple instances of GC links within A1 (e.g., $5 \mapsto 4$) or within PFC (e.g., $9 \mapsto 8$) emerge or vanish during the active block, which accounts
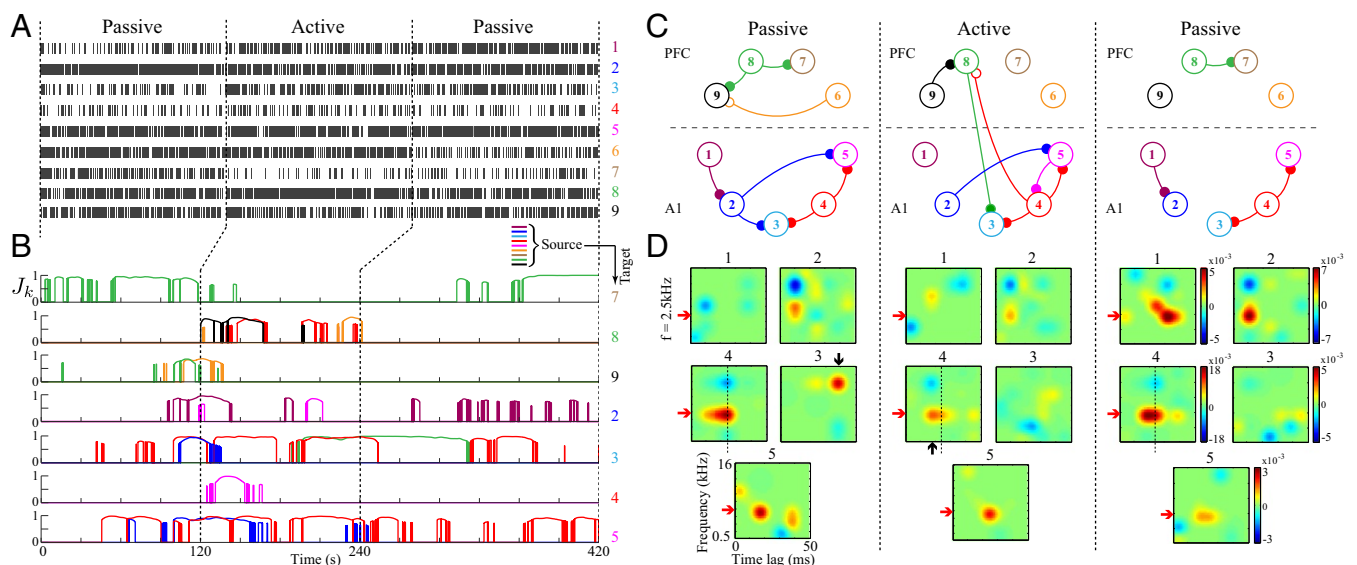


**Fig. 6.** Dynamic inference of G-causal influences between single-units in ferret PFC and A1 during auditory task. (A) Spike trains corresponding to the first repetition of each block. (B) Time course of significant GC changes through $J$-statistics for selected single-units (FDR controlled at $\alpha = 0.1$). (C) Detected patterns of network AGC maps during three main blocks of experiment. (D) Spectrotemporal receptive field (STRF) snapshots of A1 units at the end points of three blocks of experiment.

for the task-specific network-level changes within the cortical regions that are involved in active listening. A salient instance of this phenomenon can be observed in the dynamics of unit 8, whose GC links within PFC significantly change during the active behavior: As it gets G-causally linked to the lower-level A1 region, its GC links to the other PFC units fade away (rows 1, 3 and 5, Fig. 6*B*). It is noteworthy that the fluctuating instances of the *J*-statistics (e.g., Fig. 6*B*, fourth row, third segment) are due to the FDR control procedure, and there is no evidence to believe that they have a neurophysiological basis. To reduce these fluctuations, one can choose a higher FDR rate. If these effects persist at high FDR rates, further inspection of the cross-history coefficients is needed to assess their possible neurophysiological basis (see *SI Appendix*, section 10 for further discussion).

To validate our results in the absence of ground truth, we assess their reliability using surrogate data obtained by random shuffling and network subsampling in *SI Appendix*, section 11 and verify the robustness of the inferred task-dependent functional network dynamics against the aforementioned adversarial perturbations. In conclusion, our methodology enabled the extraction of the top-down and bottom-up network-level dynamics that were previously conjectured in ref. 63 to be involved in active attentive behavior, at the neuronal scale with high spatiotemporal resolution. In *SI Appendix*, section 12 we present our analysis of another experiment, which further corroborates our findings.

## Discussion and Concluding Remarks

**Summary and Extensions of Our Contributions.** Most widely adopted time series analysis techniques for quantifying functional causal relations among the nodes in a network assume static functional structures or otherwise enforce dynamics using sliding windows. While they have proven successful in analyzing stationary Gaussian time series, when applied to spike recordings from neuronal ensembles undergoing rapid task-dependent dynamics they hinder a precise statistical characterization of the sparse dynamic neuronal functional networks underlying adaptive behavior.

To address these shortcomings, we developed a dynamic inference paradigm for extracting functional neuronal network dynamics in the sense of Granger, by integrating techniques from adaptive filtering, compressed sensing, point process theory, and high-dimensional statistics. We proposed a measure of time-varying GC, namely AGC, and demonstrated its utility through theoretical analysis, algorithm development, and application to synthetic and real data. Our analysis of the mouse auditory cortical data revealed unique features of the functional neuronal network structures underlying spontaneous activity at unprecedented spatial resolution. Application of our techniques to simultaneous recordings from the ferret auditory and prefrontal cortical areas suggested evidence for the role of rapid top-down and bottom-up functional dynamics across these areas involved in robust attentive behavior.

The plug-and-play nature of the algorithms used in our framework enables it to be generalized for application to various other domains beyond neuroscience, such as the analysis of social networks or gene regulatory networks. As an example, the GLM models can be generalized to account for $m$-ary data, the forgetting factor mechanism for inducing adaptivity can be extended to state-space models governing the coefficient dynamics, and the FDR correction can be replaced by more recent techniques such as knockoff filters (67). To ease reproducibility and aid the adoption of our method, we have archived a MATLAB implementation on GitHub (https://github.com/Arsha89/AGC_Analysis).

**Limitations of Our Approach.** In closing, it is worth discussing two potential limitations of our proposed paradigm.

***Confounding effects due to network subsampling.*** A common criticism of statistical causality measures, such as the GC, directed information, or transfer entropy, is susceptibility to latent confounding causal effects arising from network subsampling. In practice, these methods are typically applied to a small subnetwork of the circuits involved in neuronal processing. Given that each neuron may receive thousands of synaptic inputs, lack of access to a large number of latent confounding inputs can affect the validity of the causal inference results obtained by these methods.

We have evaluated the robustness of our method against such confounding effects using comprehensive numerical studies in *SI Appendix*, section 9. These studies involve scenarios with deterministic and stochastic latent common inputs as well as confounding effects due to network subsampling and suggest that our techniques indeed exhibit a degree of immunity to such confounding effects. We argue that this performance is due to explicit modeling of the dynamics of the Granger causal effects in the GLM framework, invoking the sparsity hypothesis, and using sharp statistical inference procedures (see *SI Appendix*, section 9 for further discussion).

***Biological interpretation.*** The functional network characterization provided by our framework must not be readily interpreted as direct or synaptic connections that result in causal effects. Our analysis results in a sparse number of GC interactions between neurons that can appear and vanish over time in a task-specific fashion. While it is possible that these connections reflect synaptic contacts between neurons, as changes in synaptic strengths can be induced rapidly within minutes (68), the observed GC dynamics could also be due to other underlying mechanisms such as desynchronization of inputs, altered shunting, or dendritic filtering. Thus, these plasticity effects remain to be tested with ground truth experiments. An alternative and inclusive view is that these links reflect a measure of information transferred from one neuron to another.

The relatively rapid switching of these links, however, must be interpreted with caution: While some of the rapid fluctuations are due to the use of the FDR control procedure (as discussed in the *Applications*), sudden emergence or disappearance of a link does not necessarily imply sudden changes in the causal structure or information transfer in the network. A sudden disappearance of a steady link most likely reflects the fact that given the amount of currently available data, there is not enough evidence to maintain the existence of the link at the group level with the desired statistical confidence; similarly, a sudden emergence of a link most likely implies that enough evidence has just been accumulated to justify its presence with statistical confidence. The gradual effects of these interactions are indeed reflected in the dynamics of the noncentrality parameters estimated by our methods.

As demonstrated by the applications of our inference procedures, our framework provides a robust characterization of the dynamic statistical dependencies in the network in the sense of Granger at high temporal resolution. This characterization can be readily used at a phenomenological level to describe the dynamic network-level functional correlates of behavior, as demonstrated by our real data applications. More importantly, this characterization can serve as a guideline in forming hypotheses for further testing of the direct causal effects using experimental procedures such as lesion studies, microstimulation, or optogenetics in animal models.

1. Olshausen BA, Field DJ (1996) Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381:607–609.
2. Olshausen BA, Field DJ (2004) Sparse coding of sensory inputs. *Curr Opin Neurobiol* 14:481–487.
3. Sporns O, Zwi JD (2004) The small world of the cerebral cortex. *Neuroinformatics* 2:145–162.
4. Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB (2005) Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol* 3:e68.
5. Rehn M, Sommer FT (2007) A network that uses few active neurons to code visual input predicts the diverse shapes of cortical receptive fields. *J Comput Neurosci* 22:135–146.
6. Druckmann S, Hu T, Chklovskii DB (2012) A mechanistic model of early sensory processing based on subtracting sparse representations. *Advances in Neural Information Processing Systems 25, eds Bartlett P, Pereira F, Burges CJC, Bottou L, Weinberger KZ (Curran Associates, Red Hook, NY)*, pp 1988–1996.
7. Ganguli S, Sompolinsky H (2012) Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis. *Annu Rev Neurosci* 35:485–508.
8. Babadi B, Sompolinsky H (2014) Sparseness and expansion in sensory representations. *Neuron* 83:1213–1226.
9. Greicius MD, Krasnow B, Reiss AL, Menon V (2003) Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proc Natl Acad Sci USA* 100:253–258.
10. Damoiseaux J, et al. (2006) Consistent resting-state networks across healthy subjects. *Proc Natl Acad Sci USA* 103:13848–13853.
11. Hagmann P, et al. (2008) Mapping the structural core of human cerebral cortex. *PLoS Biol* 6:e159.
12. Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes: II. Simultaneous spike trains. *Biophys J* 7:419–440.
13. Gerstein GL, Perkel DH (1969) Simultaneously recorded trains of action potentials: Analysis and functional interpretation. *Science* 164:828–830.
14. Brody CD (1999) Correlations without synchrony. *Neural Comput* 11:1537–1551.
15. Aertsen AM, Gerstein GL (1985) Evaluation of neuronal connectivity: Sensitivity of cross-correlation. *Brain Res* 340:341–354.
16. Bernasconi C, König P (1999) On the directionality of cortical interactions studied by structural analysis of electrophysiological recordings. *Biol Cybern* 81:199–210.
17. Kamiński M, Ding M, Truccolo WA, Bressler SL (2001) Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biol Cybern* 85:145–157.
18. Goebel R, Roebroeck A, Kim DS, Formisano E (2003) Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magn Reson Imaging* 21:1251–1261.
19. Brovelli A, et al. (2004) Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality. *Proc Natl Acad Sci USA* 101:9849–9854.
20. Geweke JF (1984) Measures of conditional linear dependence and feedback between time series. *J Am Stat Assoc* 79:907–915.
21. Geweke J (1982) Measurement of linear dependence and feedback between multiple time series. *J Am Stat Assoc* 77:304–313.
22. Kaminski M, Blinowska KJ (1991) A new method of the description of the information flow in the brain structures. *Biol cybernetics* 65:203–210.
23. Baccalá LA, Sameshima K (2001) Partial directed coherence: A new concept in neural structure determination. *Biol Cybern* 84:463–474.
24. Sommerlade L, et al. (2012) Inference of Granger causal time-dependent influences in noisy multivariate time series. *J Neurosci Methods* 203:173–185.
25. Milde T, et al. (2010) A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate ar models and its application in analysis of laser-evoked brain potentials. *Neuroimage* 50:960–969.
26. Havlicek M, Jan J, Brazdil M, Calhoun VD (2010) Dynamic granger causality based on Kalman filter for evaluation of functional network connectivity in fMRI data. *Neuroimage* 53:65–77.
27. Möller E, Schack B, Arnold M, Witte H (2001) Instantaneous multivariate EEG coherence analysis by means of adaptive high-dimensional autoregressive models. *J Neurosci Methods* 105:143–158.
28. Hesse W, Möller E, Arnold M, Schack B (2003) The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. *J Neurosci Methods* 124:27–44.
29. Astolfi L, et al. (2008) Tracking the time-varying cortical connectivity patterns by adaptive multivariate estimators. *IEEE Trans Biomed Eng* 55:902–913.
30. Sato JR, et al. (2006) A method to produce evolving functional connectivity maps during the course of an fMRI experiment using wavelet-based time-varying Granger causality. *Neuroimage* 31:187–196.
31. Valdés-Sosa PA, et al. (2005) Estimating brain functional connectivity with sparse multivariate autoregression. *Philos Trans R Soc Lond B Biol Sci* 360:969–981.
32. Stevenson IH, et al. (2009) Bayesian inference of functional connectivity and network structure from spikes. *IEEE Trans Neural Syst Rehabil Eng* 17:203–213.
33. Zhou Z, et al. (2009) Detecting directional influence in fMRI connectivity analysis using PCA based granger causality. *Brain Res* 1289:22–29.
34. Dhamala M, Rangarajan G, Ding M (2008) Analyzing information flow in brain networks with nonparametric Granger causality. *Neuroimage* 41:354–362.
35. Sameshima K, Baccalá LA (1999) Using partial directed coherence to describe neuronal ensemble interactions. *J Neurosci Methods* 94:93–103.
36. Krumin M, Shoham S (2010) Multivariate autoregressive modeling and Granger causality analysis of multiple spike trains. *Comput intelligence Neurosci* 2010:10.
37. Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN (2005) A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J Neurophysiol* 93:1074–1089.
38. Okatan M, Wilson MA, Brown EN (2005) Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity. *Neural Comput* 17:1927–1961.
39. Nedungadi AG, Rangarajan G, Jain N, Ding M (2009) Analyzing multiple spike trains with nonparametric Granger causality. *J Comput Neurosci* 27:55–64.
40. Kim S, Putrino D, Ghosh S, Brown EN (2011) A Granger causality measure for point process models of ensemble neural spiking activity. *PLoS Comput Biol* 7:e1001110.
41. Quinn CJ, Coleman TP, Kiyavash N, Hatsopoulos NG (2011) Estimating the directed information to infer causal relationships in ensemble neural spike train recordings. *J Comput Neurosci* 30:17–44.
42. Kim S, Quinn CJ, Kiyavash N, Coleman TP (2014) Dynamic and succinct statistical analysis of neuroscience data. *Proc IEEE* 102:683–698.
43. Sheikhattar A, Fritz JB, Shamma SA, Babadi B (2016) Recursive sparse point process regression with application to spectrotemporal receptive field plasticity analysis. *IEEE Trans Signal Process* 64:2026–2039.
44. Smith A, Brown EN (2003) Estimating a state-space model from point process observations. *Neural Comput* 15:965–991.
45. Paninski L (2004) Maximum likelihood estimation of cascade point-process neural encoding models. *Netw Comput Neural Syst* 15:243–262.
46. Paninski L, Pillow J, Lewi J (2007) Statistical models for neural encoding, decoding, and optimal stimulus design. *Prog Brain Res* 165:493–507.
47. Haykin SS (1996) *Adaptive Filter Theory* (Prentice Hall, Upper Saddle River, NJ).
48. Brown EN, Nguyen DP, Frank LM, Wilson MA, Solo V (2001) An analysis of neural receptive field plasticity by point process adaptive filtering. *Proc Natl Acad Sci USA* 98:12261–12266.
49. Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci* 6:1216–1223.
50. Wilks SS (1938) The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann Math Stat* 9:60–62.
51. Davidson RR, Lever WE (1970) The limiting distribution of the likelihood ratio statistic under a class of local alternatives. *Sankhyā: Indian J Stat Ser A* 32:209–224.
52. Peers H (1971) Likelihood ratio and associated test criteria. *Biometrika* 58:577–587.
53. Van de Geer S, Bühlmann P, Ritov Y, Dezeure R (2014) On asymptotically optimal confidence regions and tests for high-dimensional models. *Ann Stat* 42:1166–1202.
54. Javanmard A, Montanari A (2014) Confidence intervals and hypothesis testing for high-dimensional regression. *The J Machine Learn Res* 15:2869–2909.
55. Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188.
56. Francis et al. (2018) Small networks encode decision-making in primary auditory cortex. *Neuron* 97:885–897.
57. Pnevmatikakis EA, et al. (2016) Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* 89:285–299.
58. Hromádka T, Zador AM (2009) Representations in auditory cortex. *Curr Opin Neurobiol* 19:430–433.
59. Watkins PV, Kao JP, Kanold PO (2014) Spatial pattern of intra-laminar connectivity in supragranular mouse auditory cortex. *Front Neural circuits* 8:15.
60. Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202.
61. Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574.
62. Buschman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315:1860–1862.
63. Fritz JB, David SV, Radtke-Schuller S, Yin P, Shamma SA (2010) Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nat Neurosci* 13:1011–1019.
64. Gilbert CD, Li W (2013) Top-down influences on visual processing. *Nat Rev Neurosci* 14:350–363.
65. Piëch V, Li W, Reeke GN, Gilbert CD (2013) Network model of top-down influences on local gain and contextual interactions in visual cortex. *Proc Natl Acad Sci USA* 110:E4108–E4117.
66. Klein DJ, Simon JZ, Depireux DA, Shamma SA (2006) Stimulus-invariant processing and spectrotemporal reverse correlation in primary auditory cortex. *J Comput Neurosci* 20:111–136.
67. Barber RF, Candès EJ (2015) Controlling the false discovery rate via knockoffs. *Ann Stat* 43:2055–2085.
68. Cooke SF, Bear MF (2010) Visual experience induces long-term potentiation in the primary visual cortex. *J Neurosci* 30:16304–16313.