Inferring and validating mechanistic models of neural microcircuits based on spike-train data

Josef Ladenbauer[1*], Sam McKenzie[2], Daniel Fine English[3], Olivier Hagens[4] and Srdjan Ostojic[1†]

[1] Laboratoire de Neurosciences Cognitives et Computationnelles, INSERM U960, École Normale Supérieure, PSL Research University, 29 rue d'Ulm, 75005 Paris, France
[2] Neuroscience Institute, New York University, 20 Cooper Square, New York, NY 10003, USA
[3] School of Neuroscience, Virginia Tech, 970 Washington St. SW, Blacksburg, VA 24061, USA
[4] Laboratory of Neural Microcircuitry, Brain Mind Institute, School of Life Sciences, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
[*] josef.ladenbauer@gmail.com [†] srdjan.ostojic@ens.fr

Abstract

The interpretation of neuronal spike train recordings often relies on abstract statistical models that allow for principled parameter estimation and model selection but provide only limited insights into underlying microcircuits. In contrast, mechanistic models are useful to interpret microcircuit dynamics, but are rarely quantitatively matched to experimental data due to methodological challenges. Here we present analytical methods to efficiently fit spiking circuit models to single-trial spike trains. Using derived likelihood functions, we statistically infer the mean and variance of hidden inputs, neuronal adaptation properties and connectivity for coupled integrate-and-fire neurons. Comprehensive evaluations on synthetic data, validations using ground truth in-vitro and in-vivo recordings, and comparisons with existing techniques demonstrate that parameter estimation is very accurate and efficient, even for highly subsampled networks. Our methods bridge statistical, data-driven and theoretical, model-based neurosciences at the level of spiking circuits, for the purpose of a quantitative, mechanistic interpretation of recorded neuronal population activity.

Introduction

In recent years neuronal spike train data have been collected at an increasing pace, with the ultimate aim of unraveling how neural circuitry implements computations that underlie behavior. Often these data are acquired from extracellular electrophysiological recordings in vivo without knowledge of neuronal input and connections between neurons. To interpret such data, the recorded spike trains are frequently analyzed by fitting abstract parametric models that describe statistical dependencies in the data. A typical example consists in fitting generalized linear models (GLMs) to characterize the mapping between measured sensory input and neuronal spiking activity [1–5]. These approaches are very useful for quantifying the structure in the data, and benefit from statistically principled parameter estimation and model selection methods. However, their interpretive power is limited as the underlying models typically do not incorporate prior biophysical constraints, and therefore have a limited capacity for identifying circuit mechanisms.

Mechanistic models of coupled neurons on the other hand involve biophysically interpretable variables and parameters, and have proven essential for analyzing the dynamics of neural circuits. At the top level of complexity for this purpose are detailed models of the Hodgkin–Huxley type [6,7], which can include a plethora of properties such as morphology and patterns of ionic currents. Comparisons between models of this type and recorded spike trains have revealed that multiple combinations of biophysical parameters give rise to the same firing patterns [8,9]. This observation motivates the use of models at an intermediate level of complexity, and in particular integrate-and-fire (I&F) neurons, which implement in a simplified manner the key biophysical constraints with a reduced number of effective parameters and can be equipped with various mechanisms such as spike initiation [10–12], adaptive excitability [13,14] or distinct compartments [15,16] to generate diverse spiking behaviors [17,18] and model multiple neuron types [19,20]. I&F models can reproduce and predict

neuronal activity with a remarkable degree of accuracy [11, 21, 22], essentially matching the performance of biophysically detailed models with many parameters [17, 23]; thus, they have become state-of-the-art models for describing neural activity in in-vivo-like conditions [11, 19, 20, 24]. In particular, they have been applied in a multitude of studies on local circuits [25–31], network dynamics [32–36], learning/computing in networks [37–40], as well as in neuromorphic hardware systems [37, 41–43].

I&F neurons can be fit in straightforward ways to membrane voltage recordings with knowledge of the neuronal input, typically from in-vitro preparations [11, 17, 19, 20, 24, 44, 45]. Having only access to the spike times as in a typical in-vivo setting however poses a substantial challenge for the estimation of parameters. Estimation methods that rely on numerical simulations to maximize a likelihood or minimize a cost function [46–48] strongly suffer from the presence of intrinsic variability in this case.

To date, model selection methods based on extracellular recordings are thus much more advanced and principled for statistical/phenomenological models than for mechanistic circuit models. To bridge this methodological gap, here we present analytical tools to efficiently fit I&F circuits to observed spike times from a single trial. By maximizing analytically computed likelihood functions, we infer the statistics of hidden inputs, input perturbations, neuronal adaptation properties and synaptic coupling strengths, and evaluate our approach extensively using synthetic data. Importantly, we validate our inference methods for all of these features using ground truth in-vitro and in-vivo data from whole-cell [49, 50] and combined juxtacellular-extracellular recordings [51]. Systematic comparisons with existing model-based and model-free methods on synthetic data and electrophysiological recordings [49–52] reveal a number of advantages, in particular for the challenging task of estimating synaptic couplings from highly subsampled networks.

Our work demonstrates that neuronal and network properties can be accurately and efficiently estimated from spike trains using biophysically constrained I&F models, and indicates the potential of this approach for the inference of microcircuit models from extracellular recordings. Numerical implementations for our methods are provided under a free license as an open source project.

Results

Our results are structured as follows. We outline our methods in section 1 and evaluate them in the subsequent sections using both synthetic data and ground truth recordings: for statistical inference of fluctuating background input in section 2, input perturbations in section 3, synaptic coupling in section 4, and neuronal adaptation in section 5. Comparisons with related approaches are included in sections 2-5. Section 6 summarizes information on implementation and computational demands.

1. Maximum likelihood estimation for integrate-and-fire neurons

Maximum likelihood estimation is a principled method for fitting statistical models to observations. Given observed data D and a model that depends on a vector of parameters $\boldsymbol{\theta}$, the estimated value of the parameter vector is determined by maximizing the likelihood that the observations are generated by the model, $p(\mathrm{D}|\boldsymbol{\theta})$, with respect to $\boldsymbol{\theta}$. This method features several attractive properties, among them: (1) the distribution of maximum likelihood estimates is asymptotically Gaussian with mean given by the true value of $\boldsymbol{\theta}$; (2) the variances of the parameter estimates achieve a theoretical lower bound, the Cramer-Rao bound, as the sample size increases [53]. It should further be noted that maximizing the likelihood $p(\mathrm{D}|\boldsymbol{\theta})$ within plausible limits for the parameter values is equivalent to maximizing the posterior probability density for the parameters given the data, $p(\boldsymbol{\theta}|\mathrm{D})$, without prior knowledge about the parameters except for the limits (i.e., assuming a uniform prior distribution of $\boldsymbol{\theta}$).

Let us first focus on single neurons. The data we have are neuronal spike times, which we collect in the ordered set $\mathrm{D} := \{t_1, \ldots, t_K\}$. We consider neuron models of the integrate-and-fire (I&F) type, which describe the membrane voltage dynamics by a differential equation together with a reset condition that simplifies the complex, but rather stereotyped, dynamics of action potentials (spikes). Here we focus on the classical leaky I&F model [54] but also consider a refined variant that includes a nonlinear description of the spike-generating sodium current at spike initiation and is known as exponential I&F model [10]. An extended

I&F model that accounts for neuronal spike rate adaptation [14, 17, 55] is included in Results section 5. Each model neuron receives fluctuating inputs described by a Gaussian white noise process with (time-varying) mean $\mu(t)$ and standard deviation σ (for details on the models see Methods section 1).

We are interested in the likelihood $p(\mathrm{D}|\boldsymbol{\theta})$ of observing the spike train D from the model with parameter vector $\boldsymbol{\theta}$. As spike emission in I&F models is a renewal process (except in presence of adaptation, see below) this likelihood can be factorized as

$$p(\mathrm{D}|\boldsymbol{\theta}) = \prod_{k=1}^{K-1} p(t_{k+1}|t_k, \mu[t_k, t_{k+1}], \boldsymbol{\theta}), \tag{1}$$

where $\mu[t_k, t_{k+1}] := \{\mu(t) \,|\, t \in [t_k, t_{k+1}]\}$ denotes the mean input time series across the time interval $[t_k, t_{k+1}]$. In words, each factor in Eq. (1) is the probability density value of a spike time conditioned on knowledge about the previous spike time, the parameters contained in $\boldsymbol{\theta}$ and the mean input time series across the inter-spike interval (ISI). Below we refer to these factors as conditioned spike time likelihoods. We assume that $\mu[t_k, t_{k+1}]$ can be determined using available knowledge, which includes the parameters in $\boldsymbol{\theta}$ as well as the observed spike times. Note that we indicate the parameter μ separately from $\boldsymbol{\theta}$ due to its exceptional property of variation over time.

For robust and rapid parameter estimation using established optimization techniques we need to compute $p(\mathrm{D}|\boldsymbol{\theta})$ as accurately *and* efficiently as possible. Typical simulation-based techniques are not well suited because they can only achieve a noisy approximation of the likelihood that depends on the realization of the input fluctuations and is difficult to maximize. This poses a methodological challenge which can be overcome using analytical tools that have been developed for I&F neurons in the context of the *forward* problem of calculating model output for given parameters [56–60]. These tools led us to the following methods that we explored for the *inverse* problem of parameter estimation:

- Method **1** calculates the factors of the likelihood (Eq. (1), right hand side) by solving a Fokker-Planck partial differential equation using suitable numerical solution schemes (for details see Methods section 2). In model scenarios where the mean input is expected to vary only little in $[t_k, t_{k+1}]$ and contains perturbations with weak effects, for example, synaptic inputs that cause relatively small postsynaptic potentials, we can write $\mu(t) = \mu_0^k + J\mu_1(t)$ with small $|J|$, where μ_0^k may vary between ISIs. This allows us to employ the first order approximation

$$p(t_{k+1}|t_k, \mu[t_k, t_{k+1}], \boldsymbol{\theta}) \approx p_0(t_{k+1}|t_k, \boldsymbol{\theta}) + J\, p_1(t_{k+1}|t_k, \mu_1[t_k, t_{k+1}], \boldsymbol{\theta}), \tag{2}$$

  where $\boldsymbol{\theta}$ contains parameters that remain constant within ISIs, including μ_0^k; J is indicated separately for improved clarity. p_0 denotes the conditioned spike time likelihood in absence of input perturbations and p_1 the first order correction. We indicate the use of this approximation explicitly by method **1a**.

- Method **2** involves an approximation of the spike train by an inhomogeneous Poisson point process. The spike rate $r(t)$ of that process is effectively described by a simple differential equation derived from the I&F model and depends on the mean input up to time t as well as the other parameters in $\boldsymbol{\theta}$ (for details see Methods section 3). In this case the factors in Eq. (1), right, are expressed as

$$p(t_{k+1}|t_k, \mu[t_k, t_{k+1}], \boldsymbol{\theta}) \approx r(t_{k+1}|\mu[t_1, t_{k+1}], \boldsymbol{\theta}) \exp\left(-\int_{t_k}^{t_{k+1}} r(\tau|\mu[t_1, \tau], \boldsymbol{\theta})\, d\tau\right). \tag{3}$$

For each of these methods the likelihood (Eq. (1)) is then maximized using well-established algorithms (see Methods section 4). Notably, for the leaky I&F model this likelihood is mathematically guaranteed to be free from local maxima [61] so that any optimization algorithm will converge to its global maximum, which ensures reliable and tractable fitting.

The most accurate and advantageous method depends on the specific setting, as illustrated for different scenarios in the following sections. Several scenarios further allow to benchmark our methods against a related approach from Pillow and colleagues, which also uses the Fokker-Planck equation [61, 62]. We compared the different methods in terms of estimation accuracy and computation time.

2. Inference of background inputs

We first consider the spontaneous activity of an isolated recorded neuron. This situation is modeled with an I&F neuron receiving a stationary noisy background input with constant mean μ and standard deviation σ. For this scenario method **1** is more accurate than method **2** while similarly efficient, and therefore best suited. We would like to point out that not all model parameters need to be estimated. The membrane voltage can be scaled such that the parameters of interest for estimation are limited to μ and σ, together with the membrane time constant τ_{m} in case of the leaky I&F model (for details see Methods section 1).

2.1 Evaluation using synthetic data

An example of the simulated ground truth data, which consists of the membrane voltage time series including spike times, is shown in Fig. 1a together with ISI and membrane voltage histograms. Note that for estimation we only use the spike times. By maximizing the likelihood the true parameter values are well recovered (Fig. 1b) and we obtain an accurate estimate of the ISI density. In addition, we also obtain an accurate estimate for the unobserved membrane voltage density, which can be calculated using a slight modification of method **1** in a straightforward way once the parameter values are determined. Interestingly, the fits are accurate regardless of whether the membrane time constant τ_{m} is set to the true value (Fig. 1a,b), estimated or set to a wrong value within a biologically plausible range (Figs. 1c and S1a).

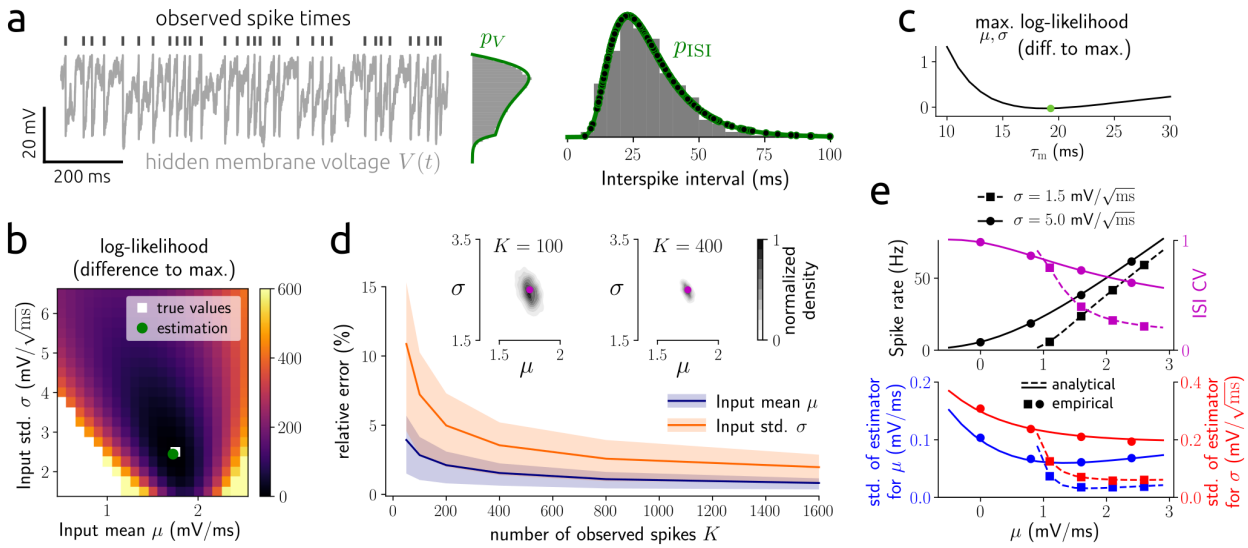

Figure 1. Estimation results for background input statistics using synthetic data. a: example membrane voltage time series with indicated spike times from a leaky I&F neuron (Eqs. (5)–(7)) together with membrane voltage and ISI histograms, and associated probability densities p_V, p_{ISI} calculated using the Fokker-Planck equation (see Methods sections 2 and 3). Parameter estimates are shown in **b**. Dots denote the values of the conditioned spike time likelihoods, i.e., the factors in Eq. (1), which are points of p_{ISI} here. Method **1** was used for parameter estimation. **b**: log-likelihood subtracted from the maximal value as a function of the input mean μ and standard deviation σ, based on 400 spikes from the example in **a**, with true and estimated parameter values indicated. **c**: maximal log-likelihood across μ and σ as a function of fixed τ_{m}, subtracted from the maximal value across τ_{m} which is attained at the indicated value. **d**: mean and central 50 % (i.e., 25th-75th percentile) of relative errors between estimated and true parameter values for μ and σ as a function of number of spikes K. Insets: empirical density of parameter estimates with true values indicated for $K = 100$ and $K = 400$. **e**, top: spike rate and ISI coefficient of variation (CV) calculated analytically using the Fokker-Planck equation (lines; see Methods sections 2 and 3) and empirically using numerical simulations (dots) as a function of μ for different values of σ; bottom: standard deviation of estimates for μ and σ according to the theoretical bound given by the Cramer-Rao inequality (lines; see Methods section 5) and empirical values from simulations (dots) for $K = 400$. In **a**,**b**,**d**,**e** τ_{m} was set to the true value.

We next evaluated the estimation accuracy for different numbers of observed spikes (Fig. 1d). As little as

50 spikes already lead to a good solution with a maximum average relative error of about 10 %. Naturally, the estimation accuracy increases with the number of observed spikes. Moreover, the variance of the parameter estimates decreases as the number of spikes increases (see insets in Fig. 1d).

To further quantify how well the different parameters can be estimated from a spike train of a given length, we analytically computed the Cramer-Rao bound (see Methods section 5). This bound limits the variance of any unbiased estimator from below and is approached by the variance of a maximum likelihood estimator. For τ_{m} fixed and a reasonable range of values for μ and σ we consistently find that the variance of the estimates of both input parameters decreases with decreasing ISI variability (Fig. 1e) and the variance for μ is smaller than that for σ (Fig. 1d,e). If τ_{m} is included in the estimation, the variance of its estimates is by far the largest and that for σ the smallest (in relation to the range of biologically plausible values, Fig. S1b,c). Together with the results from Figs. 1c and S1a this motivates to keep τ_{m} fixed.

Comparisons of our method with the approach from Pillow et al. 2005 [62] show clear improvements on estimation accuracy (Fig. S1c) and computation time (Fig. S1d: reduction by two orders of magnitude). Finally, we tested our method on the exponential I&F model, which involves an additional nonlinearity and includes a refractory period. For this model we also obtain accurate estimates (Fig. S1e).

2.2 Validation using in-vitro ground truth data and in-vivo recordings

We validated our inference method using somatic whole-cell recordings of cortical pyramidal cells (PYRs) [49] and fast-spiking interneurons (INTs) exposed to injected fluctuating currents. A range of stimulus statistics, in terms of different values for the mean μ_I and standard deviation σ_I of these noise currents, was applied and each cell responded to multiple different stimuli (examples are shown in Fig. 2a; for details see Methods section 9.1). We estimated the input parameters μ and σ of an I&F neuron from the observed spike train for each stimulus by maximizing the spike train likelihood. Note that since we considered only spikes, and not the membrane potential, we did not estimate the input resistance and rest potential; therefore, the input parameters were defined up to arbitrary offset and scale factors.

Model fitting yielded an accurate reproduction of the ISI distributions (Fig. 2a). Importantly, the estimated input statistics captured well the true stimulus statistics (Fig. 2b,c). In particular, estimated and true mean input as well as estimated and true input standard deviations were strongly correlated for all cells (Fig. 2b,c). The correlation coefficients between estimated and ground truth values for INTs are larger than those for PYRs, as reflected by the concave shape of the estimated μ values as a function of μ_I. This shape indicates a saturation mechanism that is not included in the I&F model. Indeed, it can in part be explained by the intrinsic adaptation property of PYRs (see Results section 5). Furthermore, correlation coefficients are slightly increased for longer stimuli (15 s compared to 5 s duration) due to improved estimation accuracy for longer spike trains (Fig. S2). Comparison with a Poisson point process showed that the I&F model is the preferred one across all cells and stimuli according to the Akaike information criterion (AIC), which takes into account both goodness of fit and complexity of a model (Fig. 2d).

Noise injections in vitro mimic in a simplified way the background inputs that lead to spontaneous neural activity in vivo, and certain dynamical aspects may not be well captured. Therefore, we additionally considered spike-train data obtained from extracellular multi-channel recordings in primary auditory cortex of awake ferrets during silence [52] (for details see Methods section 9.2). Examples of baseline data in terms of ISI histograms together with estimation results are shown in Fig. 2e, and estimated parameter values of the background input for all cells are visualized in Fig. 2f. The I&F model with fluctuating background input captures well a range of ISI histograms that characterize the baseline spiking activity of cortical neurons. Also here the I&F model appears to be the preferred one compared to a Poisson process for almost all cells according to the AIC (Fig. 2g).

3. Inference of input perturbations

We next focus on the effects of synaptic or stimulus-driven inputs at known times. Specifically, we consider $\mu(t) = \mu_0 + J\mu_1(t)$, where μ_0 denotes the background mean input and $\mu_1(t)$ reflects the superposition of input pulses triggered at known times. For this scenario we apply and compare methods **1a** and **2**.

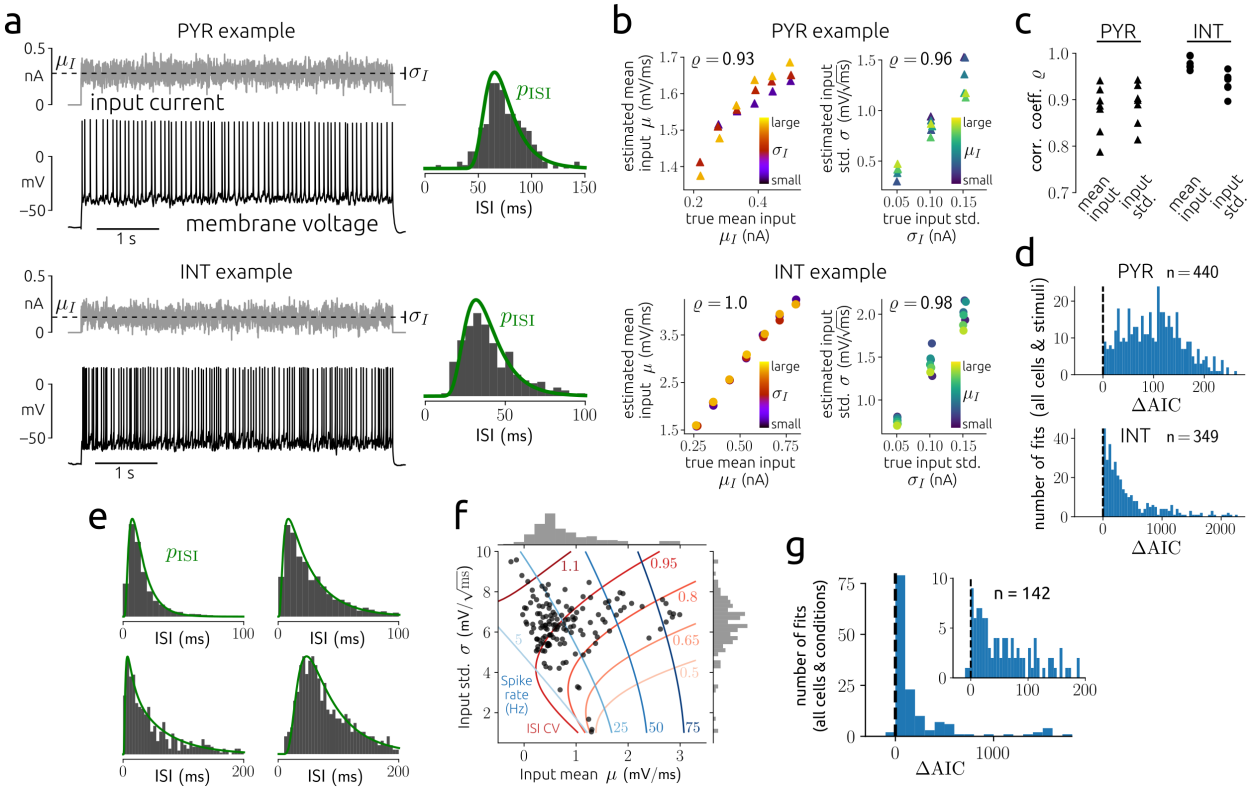

Figure 2. Estimation results for background input statistics using in-vitro and in-vivo data. **a**: examples of recorded membrane voltage in response to fluctuating input with indicated mean μ_I and standard deviation σ_I (vertical bar marks $\mu_I \pm \sigma_I$) together with ISI histogram and density p_{ISI} that corresponds to the fitted leaky I&F model for a PYR (top) and an INT (bottom). Method **1** was used for parameter estimation. **b**: estimated input parameters (μ, σ) versus empirical input statistics (μ_I, σ_I) with Pearson correlation coefficient ϱ indicated for an example PYR and INT each. **c**: ϱ for input mean and standard deviation for all seven PYRs and six INTs. **d**: histograms of AIC difference (ΔAIC) between the Poisson and I&F models across all stimuli for all PYRs (top) and all INTs (bottom). **e**: examples of ISI histograms from in-vivo recordings and densities p_{ISI} from fitted I&F neurons. **f**: estimates of parameters for the background input together with contour lines of equal spike rate and ISI CV calculated from p_{ISI}. **g**: histogram of AIC difference between the Poisson and I&F models across all cells and conditions.

3.1 Evaluation using synthetic data

For the evaluation we model $\mu_1(t)$ by the superposition of alpha functions with time constant τ, triggered at known event times with irregular inter-event intervals. We estimate the perturbation strength J as well as τ which determines the temporal extent over which the perturbation acts. Estimation accuracy for a range of perturbation strengths is shown in Fig. 3a,b. Note that the input perturbations are relatively weak, producing mild deflections of the hidden membrane voltage which are difficult to recognize visually in the membrane voltage time series in the presence of noisy background input (Fig. 3a). Both methods perform comparably well for weak input perturbations. As $|J|$ increases the estimation accuracy of method **2** increases, whereas that of method **1a** decreases (Fig. 3b) because it is based on a weak coupling approximation.

We further assessed the sensitivity of our estimation methods for the detection of weak input perturbations, and considered a model-free method based on cross-correlograms (CCGs) between spike trains and perturbation times (Fig. 3c,d) for comparison. Briefly, detection sensitivity was measured by the fraction of significant estimates of J and CCG extrema for positive lags, respectively, compared to the estimated values and CCGs without perturbations from large numbers of realizations (for details see Methods section 6). The model-free approach estimates the probability that the input and the spike train are coupled, but does not

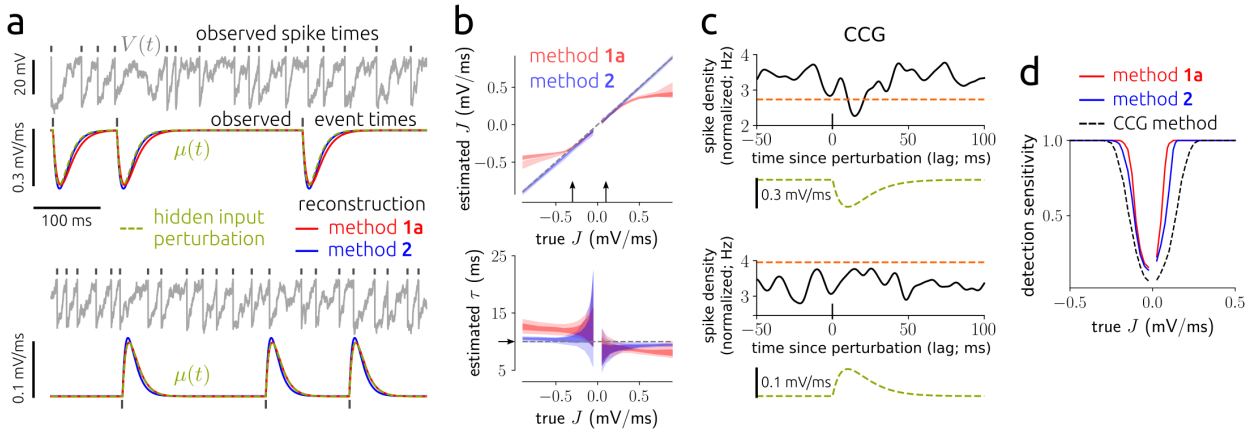

Figure 3. Estimation results for input perturbations using synthetic data. a: example membrane voltage and mean input time series (true and reconstructed ones using the estimated parameter values) for weak excitatory (lower panel) and somewhat stronger inhibitory (upper panel) perturbations. Methods **1a** and **2** were used for parameter estimation. **b**: central 50 % (i.e., 25th-75th percentile) of estimates for J and τ as a function of true J for two spike train lengths (100 s, light shade; 500 s, dark shade). Arrows indicate the parameter choices in **a**. **c**: normalized CCGs (i.e., spike density curves aligned to perturbation onset times) with significance threshold (orange dashed line) for the indicated mean input perturbations, corresponding to the examples in **a**. The CCGs are normalized such that their integral over all lags equals one. **d**: detection sensitivity as a function of J computed from the estimation methods (solid lines) or using CCGs (dashed line) based on spike trains of length 100 s (cf. **c**). For details see Methods section 6.

provide additional information on the shape of that coupling. Both model-based estimation methods are more sensitive in detecting weak perturbations than the model-free approach, with method **1a** expectedly performing best (Fig. 3d).

For additional benchmarks we considered the approach from [62] (Fig. S3). We compared estimation accuracy for the parameters of perturbations (Fig. S3a), background input and the membrane time constant (Fig. S3b), detection sensitivity (Fig. S3c) as well as computation time (Fig. S3d). Both of our methods are clearly more accurate in terms of parameter estimation (Fig. S3a,b) and thus, reconstruction (Fig. S3e). As a result, detection sensitivity is improved (Fig. S3c), while computation time is strongly reduced (Fig. S3d).

Using our model we then examined how the statistics of the background input affect the ability to detect weak input perturbations. For this purpose we analytically computed the change in expected ISI caused by a weak, brief input pulse (Fig. S4). That change depends on the time within the ISI at which the perturbation occurs (Fig. S4a). By comparing the maximal relative change of expected ISI across parameter values for the background input we found that a reduction of spiking variability within biologically plausible limits (ISI CV $\gtrsim 0.3$) typically increases detectability according to this measure (Fig. S4b).

3.2 Validation using in-vitro ground truth data

We validated our inference approach using whole-cell recordings of cortical pyramidal cells that received injections of a fluctuating current and an additional artificial excitatory postsynaptic current (aEPSC) [50]. The fluctuating current was calibrated for each cell to produce ~ 5 spikes/s and membrane voltage fluctuations of maximal ~ 10 mV amplitude. aEPSC bumps with 1 ms rise time and 10 ms decay time were triggered at simulated presynaptic spike times with rate 5 Hz, and their strength varied in different realizations (for details see Methods section 9.3).

We applied methods **1a** and **2** to detect these artificial synaptic inputs based on the presynaptic and postsynaptic spike times only, and compared the detection performance with that of a CCG-based method. Specifically, input perturbations were described by delayed pulses for method **1a** (which yields improved computational efficiency) and alpha functions for method **2** (cf. Results section 3.1). We quantified how

much data is necessary to detect synaptic input of a given strength. For this purpose we computed the log-likelihood ratio between the I&F model with ($J \neq 0$) and without ($J = 0$) coupling in a cross-validated setting, similarly to the approach that has been applied to this dataset previously using a Poisson point process GLM [50]. For the model-free CCG method we compared the CCG peak for positive lags to the peaks from surrogate data generated by distorting the presynaptic spike times with a temporal jitter in a large number of realizations. Detection time is then defined respectively as the length of data at which the model with coupling provides a better fit than the one without according to the likelihood on test data [50] (Fig. 4a), and at which the CCG peak crosses the significance threshold (Fig. 4b; for details see Methods section 9.3).

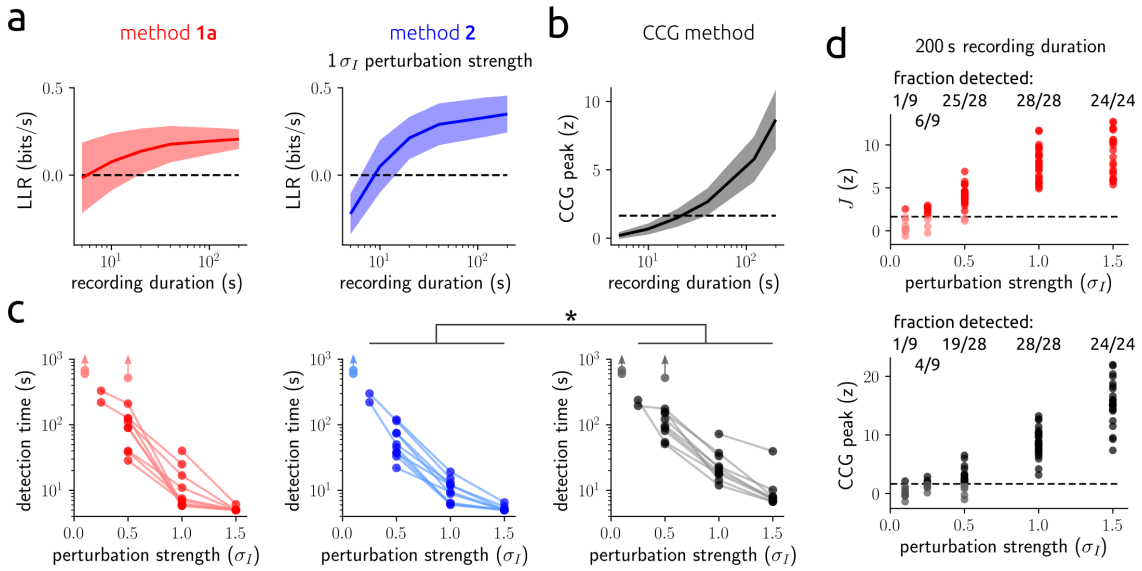

Figure 4. Estimation results for artificial synaptic input perturbations using in-vitro data. **a**: log-likelihood ratio (LLR) of the I&F model with/without coupling on test data (10-fold cross-validation) as a function of recording duration for fixed aEPSC strength (in unit of σ_I, the standard deviation of the fluctuating input current) using methods **1a** (left) and **2** (right). Shaded areas in **a** and **b** denote mean ± standard deviation across all cells. LLR > 0 indicates successful detection. **b**: CCG peak z-score as a function of recording duration for aEPSC strength 1 σ_I. Dashed line indicates the significance threshold (95th percentile of the standard normal distribution). **c**: detection time for each cell as a function of aEPSC strength (in unit of σ_I) for method **1a** (left), **2** (center), and the CCG method (right). Lines connect detection results for the same cell, light dots denote misses. In those cases, the detection time, if it exists, must be longer than the recording length, as indicated by arrows. * significant reduction of detection time ($p = 0.0276$, two-tailed paired t-test, sample size $n = 32$). **d**: estimated perturbation strength z-score using method **1a** (top) and CCG peak z-score (bottom), both as a function of aEPSC strength for 200 s recording length. Recordings were split into 200 s segments with up to 40 s overlap. Dashed lines indicate the significance threshold (cf. **b**). Dark dots mark successful detections.

All three methods were able to successfully detect most aEPSC perturbations, and method **2** required the least data for detection (Fig. 4c; overall reduction in detection time, also compared to the GLM used previously: Figs. 2F,G in [50] are directly comparable to Figs. 4a,c). We further considered an alternative detection criterion for the I&F model, based on surrogate data as used for the CCG method. This approach thus enables a more direct comparison. Comparing the detection performance for fixed recording duration across all perturbation strengths and available cells shows that method **1a** is more sensitive than the CCG method (success rate 0.86 versus 0.78, Fig. 4d). We conclude that both methods, **1a** with delayed pulses and **2** with an alpha function kernel, are well suited to detect weak aEPSCs in this in-vitro dataset, and their sensitivity is increased compared to a model-free CCG method.

4. Inference of synaptic coupling

In the previous section we showed that we can successfully estimate the perturbations in the spiking of an individual neuron that may be elicited by inputs from another neuron. We now turn to estimating synaptic couplings in a network. We consider the situation in which we have observed spike trains of N neurons. We fit these data using a network model in which each neuron i receives independent fluctuating background input with neuron-specific mean $\mu_i(t)$ and variance σ_i^2, and neurons are coupled through delayed current pulses which cause postsynaptic potentials of size $J_{i,j}$ with time delay $d_{i,j}$, for $i, j \in \{1, \ldots, N\}$. The mean background input may vary over time to reflect large amplitude variations in the external drive (and thus, spiking activity) that are slow compared to the fast input fluctuations captured by the Gaussian white noise process (cf. Results section 4.2 below). Our aim is therefore to estimate the coupling strengths in addition to the statistics of background inputs caused by unobserved neurons.

We collect the observed spike times of all N neurons in the set D and separate the matrix of coupling strengths $\mathbf{J}$ from all other parameters in $\boldsymbol{\theta}$ for improved clarity. Since $\mu_i(t)$ is assumed to vary slowly we approximate it by one value across each ISI. The overall mean input for neuron i across the k-th ISI can therefore be expressed as $\mu_i^k + \sum_{j=1}^N J_{i,j}\mu_j^1(t)$, where $J_{i,j}\mu_j^1(t)$ describes the synaptic input for neuron i elicited by neuron j taking the delay $d_{i,j}$ into account. The likelihood $p(\mathrm{D}|\boldsymbol{\theta}, \mathbf{J})$ can be factorized into conditioned spike time likelihoods, where each factor is determined by the parameters in $\boldsymbol{\theta}$ together with a specific subset of all coupling strengths and knowledge of the spike times that we have observed. Assuming reasonably weak coupling strengths, each of these factors can be approximated by the sum of the conditioned spike time likelihood in absence of input perturbations and a first order correction due to synaptic coupling (cf. Eq. (2)) to obtain

$$p(\mathrm{D}|\boldsymbol{\theta}, \mathbf{J}) \approx \prod_{i=1}^{N} \prod_{k=1}^{K_i-1} p_0(t_i^{k+1}|t_i^k, \boldsymbol{\theta}) + \sum_{j=1}^{N} J_{i,j}\, p_1(t_i^{k+1}|t_i^k, \mu_j^1[t_i^k, t_i^{k+1}], \boldsymbol{\theta}), \tag{4}$$

where t_i^k denotes the k-th of K_i observed spike times and $\boldsymbol{\theta}$ contains all parameters, including μ_i^k (except for $\mathbf{J}$). Note that the mean input perturbations μ_j^1 depend on the spike times of neuron j taking the delay $d_{i,j}$ into account. The approximation (4) implies that an individual synaptic connection on one neuron has a negligible effect for the estimation of a connection on another neuron, which is justified by the assumption of weak coupling. This allows for the application of method **1a**, by which the likelihood can be calculated in an efficient way (for details see Methods section 7).

4.1 Evaluation and validation using synthetic data

We first evaluated our method on simulated data for small ($N = 10$) as well as larger ($N = 50$) fully observed networks of neurons with constant background input statistics. The number of parameters inferred per network, which include mean and standard deviation of background inputs, coupling strengths and delays, excluding self-couplings, amounts to $2N^2$. The estimated parameter values show a remarkable degree of accuracy (Fig. 5a,b left and Fig. S5a).

In a more realistic scenario, the N recorded neurons belong to a larger network that is subsampled through the measurement process. The unobserved neurons therefore contribute additional, hidden inputs. In the fitted model, the effect of these unobserved neurons on neuron i is absorbed in the estimated parameters (μ_i, σ_i) of the background noise. Specifically, the total external input to neuron i, originating from a large number M_i of unobserved neurons whose spike trains are represented by independent Poisson processes with rates r_j, can be approximated for reasonably small coupling strengths with a background noise of mean $\mu_i = \sum_{j=N+1}^{N+M_i} J_{i,j} r_j$ and variance $\sigma_i^2 = \sum_{j=N+1}^{N+M_i} J_{i,j}^2 r_j$. This is the classical diffusion approximation [63]. Because of shared connections from unobserved neurons, the inputs received by the different observed neurons are in general correlated, with correlation strengths that depend on the degree of overlap between the unobserved presynaptic populations. Although the model we fit to the observed data assumes uncorrelated background inputs, the effects of input correlations on the neuronal ISI distributions are approximated in the estimation of μ_i and σ_i.

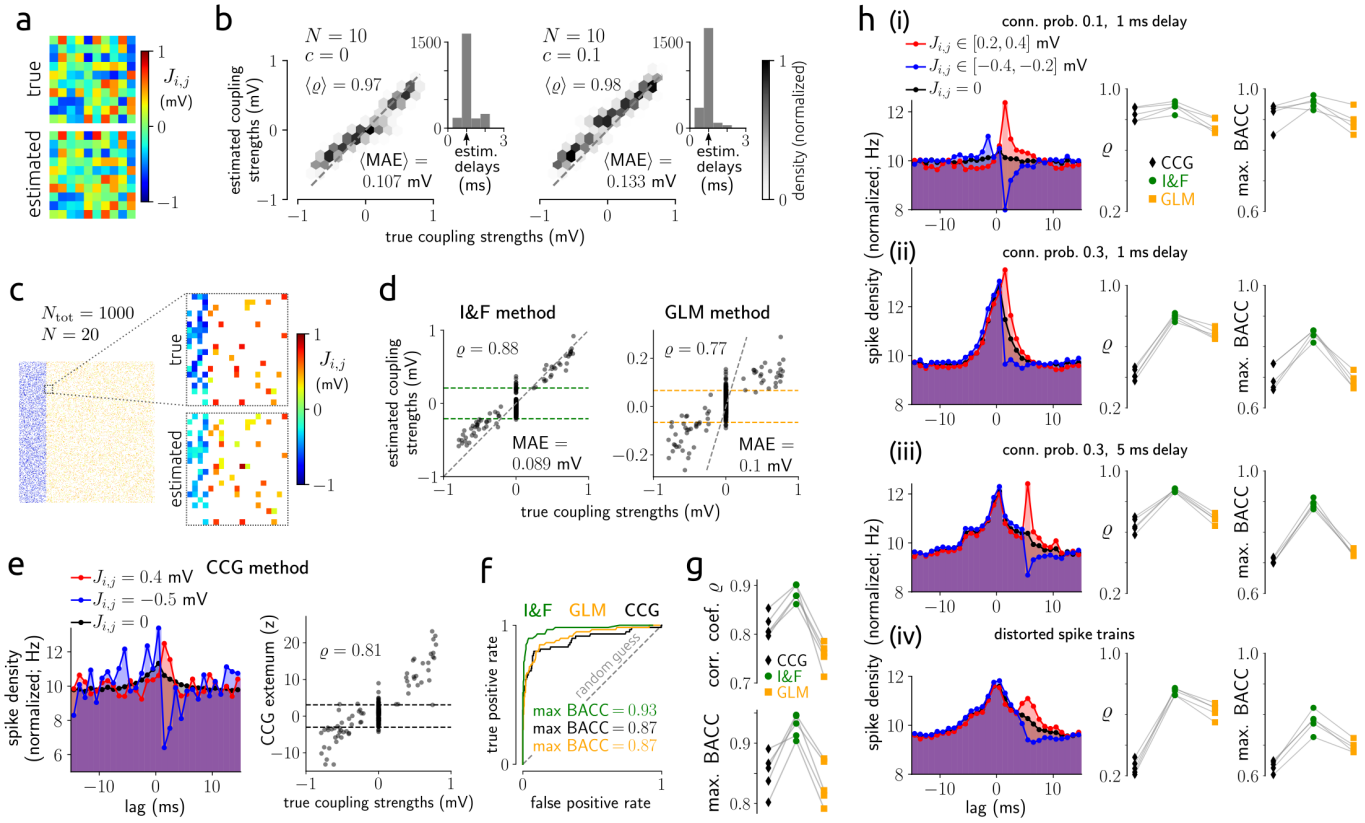

Figure 5. Estimation results for synaptic coupling strengths using synthetic data. **a**: example coupling matrix using 5 min long spike trains from $N = 10$ neurons and estimation method **1a**. **b**: density of estimated vs. true coupling strengths from 25 networks of $N = 10$ neurons for uncorrelated ($c = 0$, left, as in **a**) and correlated ($c = 0.1$, right) external input fluctuations. Average Pearson correlation coefficient $\langle\varrho\rangle$ and mean absolute error $\langle$MAE$\rangle$ between true and estimated coupling strengths are indicated. Insets: histogram of estimated delays across all networks, arrow marks the true (global) delay. **c**: true and estimated coupling strengths of $N = 20$ neurons from a randomly coupled network of $N_{\text{tot}} = 1000$ (800 excitatory, 200 inhibitory) neurons with connection probabilities of 0.1 for excitatory and 0.4 for inhibitory synapses, heterogeneous synaptic strengths and statistics of external inputs; $c = 0$, delay 1 ms. **d**–**f** show inference results from the I&F method compared to GLM- and CCG-based methods for the network in **c**. **d**: estimated vs. true coupling strengths for the I&F method (left, as used in **c**) and the GLM method (right). Dashed grey lines mark identity. A discrimination threshold for the presence of connections was applied (indicated by green and orange dashed lines; for details see Methods section 7). **e**, left: spike train CCGs for an excitatory connection (red), an inhibitory connection (blue) and averaged over all uncoupled pairs (black); right: CCG extremum z-score for positive lags vs. true coupling strengths, with discrimination threshold indicated. **f**: receiver operating characteristic (ROC) curves for the detection of synapses, i.e., true positive rate vs. false positive rate as the discrimination threshold varies, for the three methods. Indicated maximal balanced accuracy (BACC) values correspond to the thresholds visualized in **d** and **e**, right. Dashed line represents random guessing. **g**: Pearson correlation coefficient ϱ (top, cf. **d** and **e**, right) and max. BACC (cf. **f**) for the three methods using 5 different networks as in **c**. Results from the same network are connected by grey lines. **h**: effects of connection probability and delay for subsampled networks of 500 excitatory and 500 inhibitory neurons. Left column: spike train CCGs averaged over subsets of excitatory (red) and inhibitory (blue) connections as indicated, and over all uncoupled pairs (black). Center and right columns: ϱ and max. BACC for the three methods (as in **g**). Connection probability and delay values are indicated in **h**(i)–**h**(iii); **h**(iv): setting as in **h**(iii), but spike trains were perturbed by a temporal jitter (random values in $[-1, 1]$ ms). Other parameter values as in **c**. Results in **h**(i)–**h**(iv) are from 5 networks each. In **c**–**h** 10 min long recordings were used.

To assess the influence of correlations due to unobserved common inputs, we first fitted our model to data generated from a network with correlated background inputs. The estimation accuracy of synaptic strengths is still good in case of weak correlations of the external input fluctuations (correlation coefficient $c \leq 0.1$ for each pair of observed neurons). Empirical values for such noise correlations, measured in experimental studies by the correlation coefficient of pair-wise spike counts on a short timescale, are typically very small [36,64,65].

We next tested our method on data generated by explicitly sub-sampling networks of randomly connected neurons, and compared the results with those from two classical approaches: a model-free CCG method (cf. Results section 3.2) and a method based on a Poisson point process GLM that is well constrained for the synthetic data [66] (for details see Methods section 7). Note that GLMs are very flexible models with typically many parameters that need to be optimized. Consequently, these models can potentially yield good fits, on the other hand, overfitting is a serious issue [5] typically tackled by strong regularization [3,67,68]. Therefore, we chose a GLM that is tailored to capture the spiking dynamics of the (ground truth) I&F network model with a minimal number of parameters.

First, we considered classical networks of 800 excitatory and 200 inhibitory neurons [33] with connection probabilities of 0.1 for excitatory connections, 0.4 for inhibitory connections, and a plausible range of coupling strengths. Inference results from partial observations ($N = 20$) show that the I&F method outperforms the other two approaches on accuracy of both estimated coupling strengths and detection of connections (Fig. 5c-g, for different recording lengths see Fig. S5b).

To gain more insight into the challenges caused by shared input from unobserved neurons and how they affect the different methods, we then varied connection probability and delay. We considered networks with equal numbers of excitatory and inhibitory neurons to ensure equal conditions for the estimation of the two connection types (Figs. 5h and S5c-f). For relatively sparse, subsampled networks (connection probability 0.1, $N = 20$ observed neurons out of 1000) all three methods perform well, and the I&F method shows only a slight improvement in terms of correlation between true and estimated coupling strengths and detection accuracy (Fig. 5h(i), for detailed inference results see Fig. S5c). The inference task becomes increasingly difficult as connectivity increases (connection probability 0.3, see Figs. 5h(ii) and S5d). In this case the correlations between spike trains of uncoupled pairs of neurons are clearly stronger on average, particularly at small time lags, which renders CCGs unsuitable to distinguish the effects of numerous synaptic connections from uncoupled pairs. As a consequence, the number of false positives and misclassified inhibitory connections increases. Hence the accuracy of the CCG method is substantially reduced. This also causes strongly impaired accuracy of the GLM method, especially with respect to detection of synapses, whereas our I&F method appears to be much less affected. Increased synaptic delays in such networks lead to improved inference results for all three methods (Figs. 5h(iii) and S5e). Intuitively, this shifts the effects of connections in the CCG to larger lag values and thus away from the peak at zero lag caused by correlated inputs (Fig. 5h(iii) left). Nevertheless, also in this case the I&F method remains the most accurate one. We further considered small spike train distortions (using a temporal jitter) to mimic strong measurement noise, which naturally caused an overall reduction in accuracy but most strongly affected the CCG method (Figs. 5h(iv) and S5f).

Finally, we would like to note that the connectivity is not the only determinant of noise correlations in random networks; an increase of coupling strengths also caused increased spike train correlations for uncoupled pairs (Fig. S6). For rather sparse networks (connection probability 0.1) the benefit of stronger couplings outweighs that disadvantage for inference (Fig. S6a). However, for slightly increased connectivity (connection probability 0.2) the noise correlations are strongly amplified, leading to an increase in mainly false positives and misclassified inhibitory connections (Fig. S6b, in comparison to Figs. 5h(ii) and S5d).

In sum, our approach yields accurate inference results for subsampled networks as long as the correlations between the hidden inputs, due to shared connections from unobserved neurons, are not too large. In particular, it outperforms classical CCG-based and GLM-based methods. Notably, the I&F method can further be accommodated for variations of the background input statistics, and thus, time-varying spike rates, caused by unobserved neurons or network interaction.

4.2 Validation using in-vivo ground truth data

We validated our inference of synaptic coupling using simultaneous extracellular recordings and juxtacellular stimulations of hippocampal neuronal ensembles in awake mice [51]. Following the approach developed in [51], we estimated connectivity by applying the I&F method to spontaneous, extracellularly recorded spiking activity, and assessed the accuracy of our estimates by comparison with ground truth data. Ground truth connectivity was obtained by evoking spikes in single PYRs juxtacellularly using short current pulses, while recording extracellular spike trains of local INTs (for an example see Fig. 6a). Ground truth values for the presence and absence of synaptic connections were derived from spike train CCGs using the evoked presynaptic spikes, taking into account co-modulation caused by common network drive (for details see Methods section 9.4).

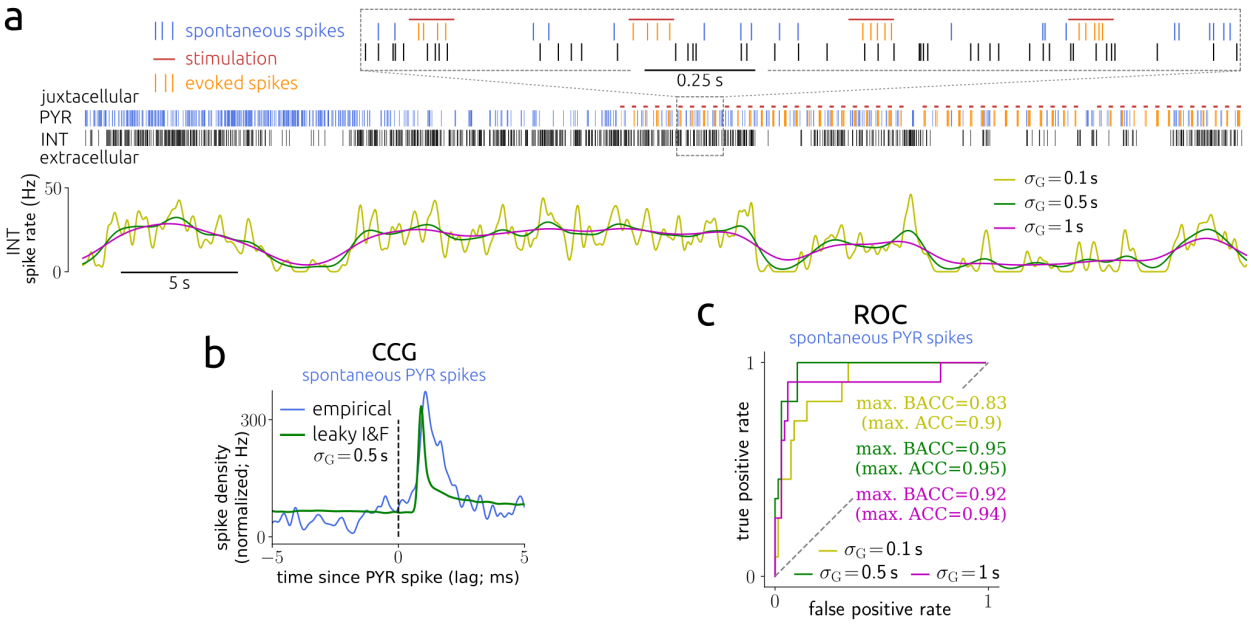

Figure 6. Estimation results for synaptic couplings using in-vivo data. **a**: example spike trains of a juxtacellularly recorded PYR and an extracellularly recorded INT, with stimulation pulses indicated (top panel). Instantaneous spike rate of the INT computed via kernel density estimation using a Gaussian kernel with width σ_G as indicated (bottom panel). **b**: CCG between spontaneous PYR and INT spike times and as computed from an I&F neuron fitted to the INT spike train with estimated coupling strength and delay, corresponding to the example in **a**. Note that synaptic coupling was modeled using delayed, transient pulses. Method **1a** was used for parameter estimation. **c**: ROC curves across 78 PYR-INT pairs using the spontaneous PYR spikes (up to 1000 per pair) for three different values of σ_G. Maximal (balanced) accuracy values are indicated.

An important aspect of these data is that spontaneous activity appeared to be highly non-stationary, so that the spike trains of the recorded neurons were typically co-modulated. To infer synaptic couplings from the spontaneous spike trains with our model-based approach, we accounted for network co-modulation in two ways: (i) through small temporal perturbations of the PYR spike times, used to compute coupling strength z-scores; (ii) through estimated variations of the background mean input for the (potentially postsynaptic) INTs, that is, μ_i^k in Eq. (4) varied between ISIs. These variations were inferred from the instantaneous spike rate, which typically varied at multiple timescales over the duration of the recordings that lasted up to ∼2 h (Fig. 6a). We, therefore, estimated the variations of mean input at three different timescales separately and inferred synaptic couplings for each of these (see Methods section 9.4).

Although spontaneous activity was highly non-stationary, our inference of the connectivity appeared to be very accurate. Comparisons with ground truth estimates demonstrated accuracy of up to 0.95 (for the intermediate timescale variant; Fig. 6c). Moreover, reducing the number of spikes used for inference

did not lead to an appreciable decrease of reproduction accuracy (Fig. S7a). Nevertheless, using instead a model-free CCG method (cf. Results sections 3.2 and 4.1) on the spontaneous spike trains yielded comparable detection accuracy (Fig. S7b, for an example CCG see Fig. 6b). This fact and the observation that the timescale affects detection performance only weakly (cf. Fig. 6c) may be explained by the large signal-to-noise ratio in this dataset (Fig. S7c), as the focus in [51] was on strong connections. We would also like to remark that a GLM-based method was previously applied and compared to the CCG-based method on these data [51], resulting in similar detection accuracy at strongly increased computational demands (200-400 s GPU computing time for a GLM fit, depending on the recording length, versus 20 ms CPU computing time for the CCG analysis).

5. Inference of neuronal adaptation

We next extend the model neurons to account for spike rate adaptation – a property of many types of neurons, including pyramidal cells [69–71]. It can be observed by a gradual change in spiking activity following an immediate response upon an abrupt change of input strength, as shown in Fig. 7a,d. This behavior is typically mediated by a calcium-activated, slowly decaying transmembrane potassium current, which rapidly accumulates when the neuron spikes repeatedly [71, 72]. In the extended I&F neuron model [14, 55] this adaptation current is represented by an additional variable w that is incremented at spike times by a value Δw, exponentially decays with slow time constant τ_w in between spikes, and subtracts from the mean input, acting as a negative feedback on the membrane voltage (Fig. 7a, see Methods section 8).

In contrast to classical I&F neurons, in the generalized model with adaptation, spiking is not a renewal process: given a spike time t_k the probability of the next spike depends on all previous spike times. That dependence is however indirect, as it is mediated through the effective mean input $\mu(t) - w(t)$ across the ISI $[t_k, t_{k+1}]$. This effective mean input can be explicitly expressed using the parameter values in $\boldsymbol{\theta}$ together with the observed spike times, and then inserted in Eqs. (1) for estimation. Here, method **1** is best suited and can be applied efficiently by exploiting the fact that w varies within ISIs in a rather stereotyped way (for details see Methods section 8). Methods **1a** and **2** are less well suited for this scenario because the adaptation variable can accumulate to substantial values, thereby opposing the assumption of weak variations of the mean input; moreover, the spike train of an adapting neuron deviates strongly from a Poisson process.

5.1 Evaluation using synthetic data

We first evaluated the inference method using simulated ground truth data for constant statistics of the background inputs (cf. Results section 2). An example of the membrane voltage and adaptation current time series is depicted in Fig. 7b. The true values for the adaptation parameters (the strength Δw and time constant τ_w) are well recovered and we obtain an accurate estimate of the adaptation current as it evolves over time. Here, too, the estimation accuracy depends on the number of observed spikes (Fig. 7c) and relative errors are on average less than 10 % for 500 spikes.

Comparisons of our method with the approach from [62] on simultaneous inference of adaptation and background input parameters show clear improvements in terms of estimation accuracy (Fig. S8a,b,d) and computation time (Fig. S8c).

5.2 Validation using in-vitro recordings

To validate our inference method for adaptation parameters, we used the recordings of neurons stimulated by noise currents that we examined in Results section 2.2. Several cells, predominantly PYRs, exhibited clear spike rate adaptation (for an example see Fig. 7d). Accordingly, the adaptive I&F model yielded a clearly improved fit compared to the nonadaptive model for all but one PYRs as shown by the AIC (Fig. 7e; for details see Methods section 9.1). On the other hand, for all except one INTs the nonadaptive model turned out to be the preferred one, which is consistent with the observation that INTs generally exhibit little spike rate adaptation compared to PYRs [73].

We examined the mean input estimates from the adaptive model in comparison to the nonadaptive model for the neurons where the adaptive model was preferred (Fig. 7e,f). For all of those cells, including adaptation

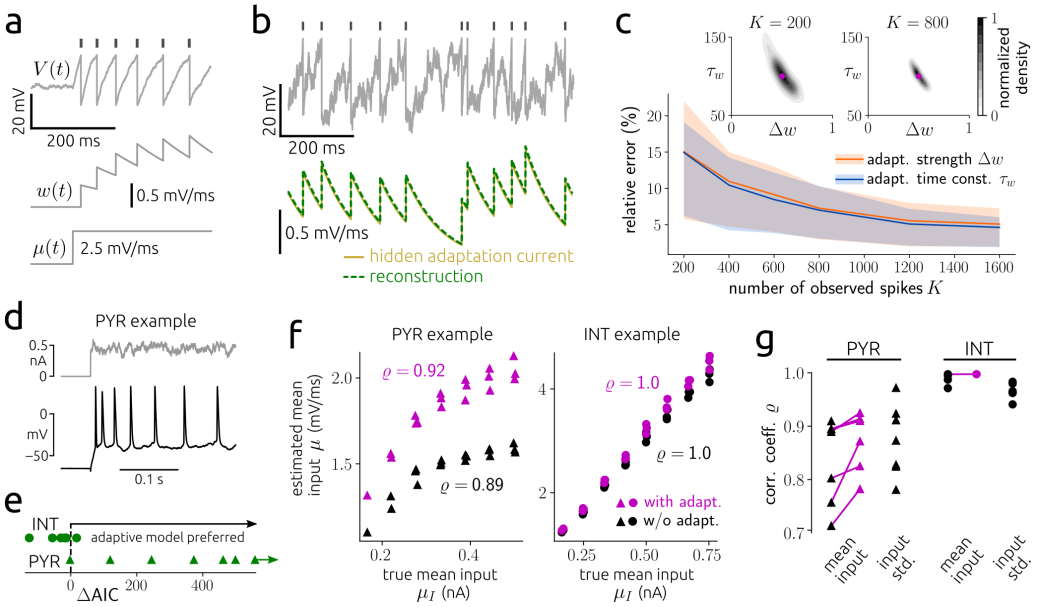

Figure 7. Estimation results for neuronal adaptation using synthetic and in-vitro data. a: membrane voltage with indicated spike times (top) and adaptation variable (center) of an adaptive leaky I&F neuron in response to a step of mean input (bottom, for small input noise intensity σ). **b**: example membrane voltage and true adaptation current time series as well as the reconstruction using estimated adaptation parameters Δw, τ_w (based on 1000 spikes) and the observed spike times. The adaptation current is considered as normalized by the (unknown) membrane capacitance and therefore in units of mV/ms. Method **1** was used for estimation. **c**: mean and central 50 % (i.e., 25th-75th percentile) of relative errors between estimated and true values for Δw and τ_w as a function of number of spikes K. Insets: empirical density of estimated parameter values with true values indicated for $K = 200$ and $K = 800$. **d**: example recorded membrane voltage in response to an injected noisy step current showing spike rate adaptation. **e**: AIC difference (ΔAIC) between the nonadaptive and adaptive leaky I&F models for all seven PYRs and six INTs. **f**: estimated mean input μ as a function of empirical mean input μ_I for the adaptive and nonadaptive models (magenta and black symbols, respectively) for two example cells. ϱ denotes Pearson correlation coefficient. **g**: ϱ for input mean and standard deviation for the two models and all neurons. Estimation results in **e**-**g** were obtained using 15 s long stimuli (corresponding to 3 repetitions of 5 s stimuli with equal statistics).

increased the correlation coefficient between estimated and empirical mean input. The remaining room for improvement of this correlation for PYRs indicates that there are likely multiple adaptation mechanisms (with different timescales) at work [73]. Note that the intrinsic adaptation current effectively subtracts from the mean input (but does not affect the input standard deviation). Indeed, the presence of adaptation in the model insignificantly affects the correlation coefficient between estimated and empirical input standard deviation (Fig. S9). This can be explained by the fact that the adaptation variable varies slowly compared to the membrane voltage and typical mean ISI (estimated $\tau_w > 4\tau_{\text{m}}$ on average); therefore, it affects the ISI mean more strongly than its variance which is predominantly influenced by the fast input fluctuations (parameter σ).

These results provide an indirect validation of our method and demonstrate its potential to infer intrinsic neuronal adaptation properties based on observed spike trains alone.

6. Implementation and computational complexity

We have implemented our methods for parameter estimation (**1**, **1a** and **2**) using the Python programming language and applying the libraries Scipy [74] for optimization and Numba [75] for low-level machine acceleration. The code for representative estimation examples from sections 2–5 is available at GitHub: https://github.com/neuromethods/inference-for-integrate-and-fire-models. Computation times for example

inference problems are summarized in table S1.

Discussion

We presented efficient, statistically principled methods to fit I&F circuit models to single-trial spike trains, and we evaluated and validated them extensively using synthetic, in-vitro and in-vivo ground truth data. Our approach allows to accurately infer hidden neuronal input statistics and adaptation currents as well as coupling strengths for I&F networks. We demonstrated that (1) the mean and variance of neuronal inputs are well recovered even for relatively short spike trains; (2) for a sufficient, experimentally plausible number of spikes, weak input perturbations triggered at known times are detected with high sensitivity, (3) coupling strengths are faithfully estimated even for subsampled networks, and (4) neuronal adaptation strength and timescale are accurately inferred. By applying our methods to suitable electrophysiological datasets, we could successfully infer the statistics of in-vivo-like fluctuating inputs, detect input perturbations masked by background noise, reveal intrinsic adaptation mechanisms, and reconstruct in-vivo synaptic connectivity.

Related approaches

Previously several likelihood-based methods related to ours have been proposed, considering uncoupled I&F neurons with [61, 62, 76] or without adaptation [77, 78] for constant [77] or time-varying [61, 62, 76, 78] input statistics. All of these methods employ the Fokker-Planck equation and numerically calculate the spike train likelihood with high precision [77] or using approximations [61, 62, 76, 78]. We benchmarked our methods against that from [61, 62], which was applicable to the estimation of single neuron parameters and for which an efficient implementation is available (cf. SI section S3). Our methods clearly outperformed the previous one in terms of estimation accuracy as well as computation time, owing to optimized numerical discretization/interpolation schemes (method **1**) and effective approximations (method **2**). Notably, our methods extend these previous approaches, centered on inference for single leaky I&F models, to the estimation of synaptic coupling in *networks* of *generalized* I&F neurons.

In the absence of likelihoods, methods for parameter fitting typically involve numerical simulations and distance measures that are defined on possibly multiple features of interest [46–48, 79, 80]. Evolutionary algorithms [46, 47, 79], brute-force search [48] or more principled Bayesian techniques [80] are then used to minimize the distances between observed and model-derived features. While these likelihood-free, simulation-based methods can be applied to more complex models they exhibit disadvantages: the distance measures usually depend on additional parameters and their evaluation depends on the particular realization of noise or randomness considered in the model. Optimization can therefore be exceedingly time-consuming. Furthermore, the principled, likelihood-based tools for model comparison (such as AIC and log-likelihood ratio) are not applicable in this case.

Estimation of synaptic coupling

Here we directly estimated synaptic coupling strengths for leaky I&F networks with fluctuating external inputs from observed spike trains. Our method is conceptually similar to those presented in [81, 82], but does not rely on an approximation of the spike train likelihood that assumes vanishing [81] or small [82] amplitudes of input fluctuations. Strongly fluctuating inputs are typically required to produce in-vivo-like spiking statistics.

We validated our approach first using simulations of network models. By applying our method to a subset of randomly selected neurons, we were able to accurately uncover the true connections in spite of co-modulation from unobserved units in the network. Inference results are accurate as long as coupling in the network is reasonably sparse (connection probability $\lesssim 0.3$) and weak (magnitude of postsynaptic potentials $\lesssim 1$ mV) so that co-modulation effects due to common presynaptic inputs are not too strong. Notably, our approach outperformed a straightforward, model-free method based on CCGs as well as an approach based on a phenomenological, point process GLM [66] throughout these validations.

This does not imply that GLM-based methods are generally less accurate in inferring synaptic connectivity: point process GLMs are flexible models which can be designed and optimized to fit the observed spike trains well [2, 3, 67, 83]. However, that approach is prone to overfitting unless strong constraints or regularization are enforced [3, 5, 66–68]. An advantage of our approach in this respect is that the basic mechanistic principles included in I&F models provide a natural regularization and reduce the number of model parameters, which strongly reduces the risk of overfitting. Note that our method **2** is essentially based on mapping I&F models to simplified, constrained GLM-like models [57, 58, 84].

In a second validation step we tested our approach on in-vivo, extracellularly recorded data [51]. Comparison with ground truth estimates obtained using juxta-cellular stimulations showed that our methods accurately identified the existing connections, although the network activity was highly non-stationary and generated strong co-modulations. Note that in that case, we did not have access to ground truth estimates of coupling strengths, and could therefore only validate the detection of connections.

Alternative approaches to infer connectivity from spike trains, other than those addressed above, have employed models of sparsely and linearly interacting point processes [85], or have been designed in a model-free manner [51, 86, 87], for example, using CCGs [51, 86] similarly to our comparisons. A general challenge in subsampled networks arises from pairwise spike train correlations at small time lags generated by shared connections from unobserved neurons, regardless of whether a direct connection is present. These spurious correlations impair our ability to distinguish the effects of synaptic connections from those caused by correlated inputs, especially when coupling delays are small. One of the benefits of our approach is that it includes an explicit, principled mechanism to account for the effects of unobserved neurons, which are absorbed in the estimated statistics of the fluctuating background inputs. Correlated fast input fluctuations are not directly modeled, their effects are compensated for in the estimation of the background input parameters, whereas shared input dynamics on a longer timescale are explicitly captured by slow variations of the mean input for each neuron. This facilitates the isolation of pairwise synaptic interactions from common drive. Furthermore, we corrected for a potential bias in coupling strength estimates, due to effects of unobserved neurons that are not captured otherwise, by our inference method.

Several related studies have focused on a theoretical link between network structure and correlated spiking activity recorded from a large number of neurons, without attempting to explicitly estimate synaptic connections (see [88–94] and [95] for review). Of major relevance in this regard is the extent to which effective interactions among observed neurons are reshaped by coupling to unobserved neurons [83, 96]. Current methods to estimate coupling strengths from observed spike trains may be further advanced using these theoretical insights.

Possible methodological extensions

Here we described synaptic interaction by a simple form of delayed current pulses that cause (connection-specific) postsynaptic potentials. In principle, a conductance-based model of synaptic coupling could also be considered. For this case the presented methods may be adjusted accordingly using results from [97], but it would generally complicate the calculations in methods **1** and **2**.

Throughout this work we assumed that the mean input trajectory across an ISI can be determined using available knowledge (that is, the model parameters and observed spike times). In Results section 4.2 we extracted the variations of the mean input from estimates of the instantaneous neuronal spike rate at different timescales. A useful extension may be to consider a separate stochastic process that governs the evolution of the mean input, allowing to extract the most appropriate timescale from the data [98], which in turn could benefit the estimation of synaptic couplings using our approach.

Another interesting direction for future work is to extend our approach for various forms of synaptic plasticity, such as short-term depression/facilitation and spike-timing-dependent plasticity, and to infer the corresponding parameters from observed spike trains [99–101]. In term, short-term plasticity is particular apparent in in-vivo recordings [51], and including it in our models would be important for estimating coupling strengths.

Towards a quantitative inference of neural circuits from spiking activity

I&F neurons are a popular tool for interpreting spiking activity in terms of simple circuit models (see, e.g., [25–31]). Such approaches typically start by hypothesizing a structure for the underlying circuit based on available physiological information, and then examine the behavior of the resulting model as a function of the critical biophysical parameters. The model is then validated by qualitatively comparing the model output with experimental data. Specifically, the model activity is required to resemble key features of the experimental data in an extended region of the parameter space. If that is not the case, the model is rejected and a different one is sought.

An important benefit of this approach is that it provides a mechanistic interpretation and understanding of recorded activity in terms of biological parameters in a neural circuit. A major limitation is, however, that it typically relies on a qualitative comparison with the data to select or reject models. The methods presented here open the door to a more quantitative, data-driven approach, in which this class of spiking circuit models can be evaluated and compared based on their fitting performance (cf. Figs. 2d,g, 4a,c and 7e for such comparisons) as is routinely the case for more abstract statistical models (see, e.g., [4]).

Methods

1. Integrate-and-fire neuron models

We consider typical I&F models subject to fluctuating inputs. The dynamics of the membrane voltage V are governed by

$$\frac{dV}{dt} = f(V) + \mu(t) + \sigma\xi(t) \tag{5}$$

$$\text{if } V(t) \geq V_\text{s} \text{ then } V(t) \leftarrow V_\text{r}, \tag{6}$$

where μ is the mean input, σ the standard deviation of the input, ξ a (unit) Gaussian white noise process, i.e., $\langle \xi(t)\xi(t+\tau)\rangle = \delta(\tau)$ with expectation $\langle\cdot\rangle$, V_s is the threshold (or spike) voltage and V_r the reset voltage. For the leaky I&F model the function f is given by

$$f(V) := -\frac{V}{\tau_\text{m}} \tag{7}$$

where τ_m denotes the membrane time constant. It should be noted that for the methods used in this paper f can be any arbitrary real-valued function. For example, in the exponential I&F model, used for Fig. S1e, f is a nonlinear function that includes an exponential term (for details see SI section S1). The parameter values are $V_\text{s} = 30$ mV, $V_\text{r} = 0$ mV, $\tau_\text{m} = 20$ ms, $\mu = 1.75$ mV/ms, $\sigma = 2.5$ mV/$\sqrt{\text{ms}}$ if not stated otherwise in figures or captions.

1.1 Voltage scaling

It is not meaningful to estimate all model parameters: a change of V_s or V_r in the leaky I&F model can be completely compensated in terms of spiking dynamics by appropriate changes of $\mu(t)$ and σ. This can be seen using the change of variables $\tilde{V} := (V - V_\text{r})/(V_\text{s} - V_\text{r})$. Consequently, we may restrict the estimation to $\mu(t)$, σ, τ_m and set the remaining parameters to reasonable values.

2. Method 1: conditioned spike time likelihood

It is useful to express the factors in Eq. (1), the conditioned spike time likelihoods, in terms of the ISI probability density p_ISI,

$$p(t_{k+1}|t_k, \mu[t_k, t_{k+1}], \boldsymbol{\theta}) = p_\text{ISI}(s_k|\mu_\text{ISI}[0, s_k], \boldsymbol{\theta}), \tag{8}$$

where $s_k := t_{k+1} - t_k$ is the length of the k-th ISI and μ_ISI is the mean input across that ISI given by $\mu_\text{ISI}[0, s_k] = \mu[t_k, t_{k+1}]$. The technical advantage of this change of variables becomes most obvious for constant

μ. In this case the density function p_{ISI} needs to be computed only once in order to evaluate the spike train likelihood, Eq. (1), for a given parametrization μ, $\boldsymbol{\theta}$.

Given a spike at $t = t_0$ the probability density of the next spike time is equal to the ISI probability density $p_{\mathrm{ISI}}(s)$ where $s := t - t_0 \geq 0$ denotes the time since the last spike. This quantity can be approximated by numerical simulation in an intuitive way: starting with initial condition $V(t_0) = V_{\mathrm{r}}$ one follows the neuronal dynamics given by Eq. (5) in each of n realizations of the noise process until the membrane voltage crosses the value V_{s} and records that spike time t_i in the i-th realization. The set of times $\{t_i\}$ can then be used to compute p_{ISI}, where the approximation error decreases as n increases. We can calculate p_{ISI} analytically in the limit $n \to \infty$ by solving the Fokker-Planck partial differential equation (PDE) [102, 103] that governs the dynamics of the membrane voltage probability density $p_V(V, s)$,

$$\frac{\partial p_V}{\partial s} + \frac{\partial q_V}{\partial V} = 0 \tag{9}$$

$$q_V := [f(V) + \mu_{\mathrm{ISI}}(s)]\, p_V - \frac{\sigma^2}{2}\frac{\partial p_V}{\partial V} \tag{10}$$

with mean input $\mu_{\mathrm{ISI}}(s) = \mu(t)$, subject to the initial and boundary conditions

$$p_V(V, 0) = \delta(V - V_{\mathrm{r}}) \tag{11}$$

$$p_V(V_{\mathrm{s}}, s) = 0 \tag{12}$$

$$\lim_{V \to -\infty} q_V(V, s) = 0. \tag{13}$$

The ISI probability density is then given by the probability flux at V_{s},

$$p_{\mathrm{ISI}}(s|\mu_{\mathrm{ISI}}[0, s], \boldsymbol{\theta}) = q_V(V_{\mathrm{s}}, s). \tag{14}$$

In the field of probability theory p_{ISI} is also known as first passage time density.

In method **1a** we consider the first order approximation (2) for weak perturbations of the mean input, $\mu(t) = \mu_0^k + J\mu_1(t)$ with small $|J|$ during the k-th ISI. In this case the k-th spike time likelihood is expressed as

$$p_{\mathrm{ISI}}(s_k|\mu_{\mathrm{ISI}}[0, s_k], \boldsymbol{\theta}) \approx p^0_{\mathrm{ISI}}(s_k|\boldsymbol{\theta}) + J\, p^1_{\mathrm{ISI}}(s_k|\mu^1_{\mathrm{ISI}}[0, s_k], \boldsymbol{\theta}) \tag{15}$$

$$\mu_{\mathrm{ISI}}[0, s_k] = \mu^0_{\mathrm{ISI}} + J\mu^1_{\mathrm{ISI}}[0, s_k], \tag{16}$$

where $\mu^1_{\mathrm{ISI}}[0, s_k] = \mu_1[t_k, t_{k+1}]$ and $\boldsymbol{\theta}$ contains parameters that remain constant within ISIs, including $\mu^0_{\mathrm{ISI}} = \mu_0^k$.

Numerical solution schemes to compute p_{ISI} (method **1**) or p^0_{ISI} and p^1_{ISI} (method **1a**) in accurate and efficient ways are provided in SI section S2. It should be noted that these functions do not need to be computed for each observed ISI separately; instead, we pre-calculate them for a reasonable set of trajectories $\mu_{\mathrm{ISI}}[0, s_{\max}]$, where $s_{\max}$ is the largest observed ISI, and use interpolation for each evaluation of a spike time likelihood (cf. Eq. (15)).

3. Method 2: derived spike rate model

Method **2** requires the (instantaneous) spike rate $r(t)$ of the model neuron described by Eqs. (5) and (6), which can be calculated by solving a Fokker-Planck system similar to Eqs. (9)–(13):

$$\frac{\partial p_V}{\partial t} + \frac{\partial q_V}{\partial V} = 0, \qquad q_V := [f(V) + \mu(t)]\, p_V - \frac{\sigma^2}{2}\frac{\partial p_V}{\partial V}, \qquad r(t) = q_V(V_{\mathrm{s}}, t), \tag{17}$$

subject to the conditions

$$p_V(V_{\mathrm{s}}, t) = 0 \qquad \lim_{V \to -\infty} q_V(V, t) = 0 \tag{18}$$

$$\lim_{V \searrow V_{\mathrm{r}}} q_V(V, t) - \lim_{V \nearrow V_{\mathrm{r}}} q_V(V, t) = q_V(V_{\mathrm{s}}, t), \tag{19}$$

where Eq. (19) accounts for the reset condition (6). The steady-state solution of this system (for constant mean input) can be conveniently calculated [56]. Obtaining the time-varying solution of Eqs. (17)–(19) is computationally more demanding and can be achieved, e.g., using a finite volume method as described in SI section S2 (see [57]).

As an efficient alternative, reduced models have been developed to approximate the spike rate dynamics of this Fokker-Planck system by a low-dimensional ordinary differential equation (ODE) that can be solved much faster [57–59, 104]. Here we employ a simple yet accurate reduced model from [57] (the LNexp model, based on [58]) adapted for leaky I&F neurons with constant input variance σ^2. This model is derived via a linear-nonlinear cascade ansatz, where the mean input is first linearly filtered and then passed though a nonlinear function to yield the spike rate. Both components are determined from the Fokker-Planck system and can be conveniently calculated without having to solve Eqs. (17)–(19) forward in time: the linear temporal filter is obtained from the first order spike rate response to small amplitude modulations of the mean input and the nonlinearity is obtained from the steady-state solution [57, 58]. The filter is approximated by an exponential function and adapted to the input in order to allow for large deviations of μ. This yields a one-dimensional ODE for the filter application,

$$\frac{d\mu_{\mathrm{f}}}{dt} = \frac{\mu(t) - \mu_{\mathrm{f}}}{\tau_\mu(\mu_{\mathrm{f}}|\boldsymbol{\theta})}, \tag{20}$$

where μ_{f} is the filtered mean input and τ_μ is the (state dependent) time constant. The spike rate is given by the steady-state spike rate of the Fokker-Planck system evaluated at $\mu = \mu_{\mathrm{f}}$,

$$r(t) = r_\infty(\mu_{\mathrm{f}}|\boldsymbol{\theta}). \tag{21}$$

In order to efficiently simulate this model we pre-calculate τ_μ and r_∞ for a reasonable range of mean input values and use look-up tables during time integration. Note that this model is based on the derivation in [58] with filter approximation scheme proposed in [57] which leads to improved accuracy of spike rate reproduction for the sensitive low input regime [57]. For a given mean input time series $\mu[t_0, t]$ we calculate $r(t|\mu[t_0, t], \boldsymbol{\theta})$ using the initial condition $\mu_{\mathrm{f}}(t_0) = \mu(t_0)$.

4. Likelihood maximization

We maximized the logarithm of the likelihood (log-likelihood),

$$\mathrm{argmax}_{\boldsymbol{\theta}}\ \log p(\mathrm{D}|\boldsymbol{\theta}) = \mathrm{argmax}_{\boldsymbol{\theta}} \sum_{k=1}^{K-1} \log p(t_{k+1}|t_k, \mu[t_k, t_{k+1}], \boldsymbol{\theta}) \tag{22}$$

for individual neurons, using Eq. (1), and similarly for networks using the logarithm of Eq. (4). Optimization was performed using a simplex algorithm [105] as implemented in the Scipy package for Python. It should be noted that our method is not restricted to this algorithm; alternative, gradient-based optimization techniques, for example, may likely lead to reduced estimation times.

5. Calculation of the Cramer-Rao bound

We computed the Cramer-Rao bound for the variance of parameter estimates in Results section 2. This bound is approached by the variance of a maximum likelihood estimator as the number of realizations increases. Let θ denote the vector of parameters for estimation (contained in $\boldsymbol{\theta}$), e.g., $\theta = (\mu, \sigma, \tau_{\mathrm{m}})^{\mathrm{T}}$. In case of a single (non-adapting) model neuron with constant input moments the Cramer-Rao bound for the variance of estimates of θ_i from spike trains with K spikes is then given by $[\mathcal{I}(\theta)]^{-1}_{i,i}/(K-1)$, where $\mathcal{I}(\theta)$ is the Fisher information matrix per ISI defined by

$$\mathcal{I}(\theta)_{i,j} := -\int_0^\infty p_{\mathrm{ISI}}(s|\mu[0, s], \boldsymbol{\theta})\, \frac{\partial^2}{\partial\theta_i\, \partial\theta_j} \log p_{\mathrm{ISI}}(s|\mu[0, s], \boldsymbol{\theta})\, ds. \tag{23}$$

6. Modeling input perturbations

In Results section 3 we consider input perturbations of the form $\mu(t) = \mu_0 + J\mu_1(t)$, where $\mu_1(t)$ is described by the superposition of alpha functions with time constant τ, triggered at times $\tilde{t}_1, \ldots, \tilde{t}_L$,

$$\mu_1(t) = \sum_{l=1}^{L} H(t - \tilde{t}_l) \frac{t - \tilde{t}_l}{\tau} \exp\left(1 - \frac{t - \tilde{t}_l}{\tau}\right) \tag{24}$$

with Heaviside step function H. The alpha functions are normalized such that their maximum value is 1 when considered in isolation. As an alternative we also considered delayed delta pulses instead of alpha kernels,

$$\mu_1(t) = \sum_{l=1}^{L} \delta(t - \tilde{t}_l - d), \tag{25}$$

where d denotes the time delay. The perturbation onset (trigger) times in Results section 3.1 were generated by randomly sampling successive separation intervals $\tilde{t}_{l+1} - \tilde{t}_l$ from a Gaussian distribution with 200 ms mean and 50 ms standard deviation. In Results section 3.1 we only considered alpha kernels, in Results section 3.2 we used both alpha functions and delta pulses.

6.1 Detection sensitivity

In Results section 3.1 we quantified the sensitivity to detect weak input perturbations using our estimation methods (**1a** and **2**) in comparison with a detection method based on the generated data only. For a given parametrization N_r spike trains were simulated using different realizations of neuronal input noise and perturbation onset times. Detection sensitivity was quantified by comparing the inferred perturbation strengths from data generated with ($J \neq 0$) and without ($J = 0$) perturbations.

For the I&F-based methods it was calculated by the fraction of $N_\mathrm{r} = 50$ estimates of J for true $J > 0$ ($J < 0$) that exceeded the 95$^\mathrm{th}$ percentile (fell below the 5$^\mathrm{th}$ percentile) of estimates without perturbation. The model-free reference method was based on CCGs between the spike trains and perturbation times (in other words, spike density curves aligned to perturbation onset times). For each realization one such curve was calculated by the differences between spike times and the perturbation onset times using a Gaussian kernel with 3 ms standard deviation. Detection sensitivity was assessed by the fraction of $N_\mathrm{r} = 300$ CCGs for which a significant peak (for $J > 0$) or trough (for $J < 0$) appeared in the interval [0, 100 ms]. Significance was achieved for true $J > 0$ ($J < 0$) if the curve maximum (minimum) exceeded the 95$^\mathrm{th}$ percentile (fell below the 5$^\mathrm{th}$ percentile) of maxima (minima) in that interval without perturbation.

7. Network model

In Results section 4 we consider networks of N_tot coupled leaky I&F neurons from which the spike trains of $N \leq N_\mathrm{tot}$ neurons have been observed. These networks are given by

$$\frac{dV_i}{dt} = -\frac{V_i}{\tau_\mathrm{m}} + \mu_i(t) + \sum_{j=1}^{N_\mathrm{tot}} J_{i,j}\mu_j^1(t) + \sigma_i \eta_i(t) \tag{26}$$

$$\mu_i^1(t) = \sum_{k=1}^{K_i} \delta(t - t_i^k - d_{i,j}) \tag{27}$$

$$\text{if } V_i(t) \geq V_\mathrm{s} \text{ then } V_i(t) \leftarrow V_\mathrm{r} \tag{28}$$

for $i \in \{1, \ldots, N_\mathrm{tot}\}$, where $J_{i,j}$ denotes the coupling strength between presynaptic neuron j and postsynaptic neuron i, t_i^k is the k-th of K_i spike times of neuron i, and $d_{i,j}$ is the delay. η_i describes the fluctuations of external input received from unobserved neurons,

$$\eta_i(t) = \sqrt{1-c}\,\xi_i(t) + \sqrt{c}\,\xi_c(t), \tag{29}$$

where ξ_i, ξ_c are independent unit Gaussian white noise processes, i.e., $\langle \xi_i(t)\xi_j(t+\tau)\rangle = \delta_{ij}\delta(\tau)$, $i,j \in \{1,\ldots,N_{\text{tot}}, c\}$, and c is the input correlation coefficient. We considered uncorrelated or weakly correlated external input fluctuations, i.e., $c = 0$ or $c = 0.1$. Note that the input variation for neuron i caused by (observed or unobserved) neuron j across the interval $[t_i^k, t_i^{k+1}]$, denoted by $J_{i,j}\mu_j^1[t_i^k, t_i^{k+1}]$ (as used in Eq. (4)), is determined by the spike times of neuron j that occur in the interval $[t_i^k - d_{i,j}, t_i^{k+1} - d_{i,j}]$. For simulated data we chose identical delays across the network, but this is not a restriction of our inference method (see section 7.1 below). Coupling strengths were uniformly sampled in $[-0.75, 0.75]$ mV (Figs. 5a,b and S5a), otherwise excitatory/inhibitory connections were randomly generated with specified probabilities and coupling strengths were then uniformly sampled with mean ± 0.5 mV (Figs. 5c-h and S5b-f) or mean ± 1 mV (Fig. S6), respectively. Autapses were excluded, i.e., $J_{i,i} = 0$. Network simulations for Figs. 5c-h, S5b-f and S6 were performed using the Python-based Brian2 simulator [106].

7.1 I&F method

Our method fits an I&F network, described by Eqs. (26)–(29) with $c = 0$ for the N observed neurons (i.e., $i \in \{1,\ldots,N\}$) to the spike train data by maximizing the likelihood (4). In the fitted model the effects of unobserved neurons are absorbed by the parameters $\mu_i(t)$ and σ_i, where $\mu_i(t)$ is discretized in an event-based way: the background mean input for neuron i during its k-th ISI is represented by a constant value μ_i^k (cf. Eq. (4)). In Results section 4.1 $\mu_i(t)$ was assumed to be constant over time, whereas in Results section 4.2 it varied between ISIs; for details on the estimation of these variations see Methods section 9.4 below.

The logarithm of the spike train likelihood (4) (cf. Methods section 4) was optimized in the following way, justified by the assumption of weak coupling. First, the parameters of the background input, μ_i^k and σ_i, were estimated for each neuron in isolation (all $J_{i,j} = 0$). Consequently, effects of observed and unobserved neurons are reflected by these estimates. Then, the coupling strength $J_{i,j}$ and delay $d_{i,j}$ were estimated given μ_i^k and σ_i for each i,j-pair. These two steps were performed in parallel over (postsynaptic) neurons. We estimated couplings in a pair-wise manner to save computation time, assuming that the transient effects of an individual synaptic connection on the spiking probability of a neuron are negligible for the estimation of other synaptic connections. This is justified for weak coupling and network activity where synchronous spikes across the network occur sparsely.

We then corrected for a potential systematic bias in the estimated coupling strengths of a network in Results section 4.1 as follows. We perturbed all presynaptic spike times by a temporal jitter (random values uniformly sampled in the interval $[-10, 10]$ ms) to mask any of the transient effects caused by synaptic connections, and re-estimated the coupling strengths for multiple such realizations for each postsynaptic neuron i given μ_i^k and σ_i. The averaged bias that was estimated across the network from this procedure was then subtracted from the original estimates.

7.2 GLM method

For comparison we used a method that fits a point process GLM to the data. From the class of GLM-based approaches [2, 3, 67] we chose one that is well suited for reconstruction from spike train data generated by an I&F network as specified in Methods section 7 and for which an efficient implementation is available [66]. In the GLM network model the incoming spike trains, after incurring transmission delays, are filtered by a leaky integrator with a time constant and a (constant) baseline activity parameter for each neuron. The resulting membrane potential is passed through an exponential link function, which transforms it into the time-varying rate of a Poisson point process that generates the output spike train of the neuron. The spike train is also fed back as an input to the neuron itself to model refractory, post-spike properties. The coupling terms in the GLM and I&F networks are equivalent: both models use delayed delta pulses.

Using maximum likelihood estimation this GLM method inferred $N^2 + 3N$ parameters per network with N neurons: the coupling strengths (including self-feedback) as well as time constant, baseline parameter and delay, one for each neuron. Note that for the simulated data only one (global) delay value was used for all connections in a network. Hence, the GLM method estimated fewer parameters compared to the I&F method (which inferred $2N^2$ parameters). For details on the elaborate inference technique, which includes regularized

optimization and cross-validation, and an available Python implementation using the library Cython for accelerated program execution we refer to [66].

7.3 CCG method

In the model-free, CCG-based method a connection strength was estimated by the z-score of the extremum in the spike train CCG across positive lags for each pair of neurons. Note that the lag denotes the time since a presynaptic spike. Z-scores were obtained using estimates from surrogate data generated by perturbing the presynaptic spike times by a temporal jitter (random values between -10 ms and 10 ms, cf. Methods section 7.1) in a large number of realizations.

7.4 Detection of couplings

For each of the three methods above we assessed detection performance in the following way. A discrimination threshold J_{thresh} for the presence or absence of connections was applied to estimated coupling strengths $\hat{J}_{i,j}$. Accordingly, the presence of a connection (i,j-pair) was assured by the condition $|\hat{J}_{i,j}| > J_{\text{thresh}}$. The true positive rate (sensitivity) was given by the number TP of connections for which the estimation satisfied $|\hat{J}_{i,j}| > J_{\text{thresh}}$ and a true connection was present ($|J_{i,j}| > 0$), divided by the number P of true connections. The true negative rate (specificity) was given by the number TN of connections for which $|\hat{J}_{i,j}| \leq J_{\text{thresh}}$ and a true connection was absent ($J_{i,j} = 0$), divided by the number N of absent connections. Receiver operating characteristic (ROC) curves were generated from sensitivity and specificity as a function of J_{thresh}. Accuracy (ACC) and balanced accuracy (BACC) are defined as $\text{ACC} = (\text{TP} + \text{TN})/(\text{P} + \text{N})$ and $\text{BACC} = (\text{TP}/\text{P} + \text{TN}/\text{N})/2$, respectively.

8. Modeling spike rate adaptation

In Results section 5 we consider an extended I&F neuron model that includes an additional adaptation (current) variable w that is incremented at spike times, slowly decays, and counteracts the input to the neuron [14,55]: Eqs. (5) and (6) where the mean input $\mu(t)$ is replaced by an effective mean input $\mu(t) - w(t)$, with

$$\frac{dw}{dt} = -\frac{w}{\tau_w} \tag{30}$$

$$\text{if } V(t) \geq V_{\text{s}} \text{ then } w(t) \leftarrow w(t) + \Delta w. \tag{31}$$

Here τ_w is the adaptation time constant and Δw denotes the spike-triggered increment.

For known spike times, contained in set D, the effective mean input can be written as $\mu(t) - \Delta w\, \mu_1(t|\text{D}, \tau_w)$, where μ_1 between spike times t_k and t_{k+1} is explicitly expressed by

$$\mu_1(t|\text{D}, \tau_w) = \sum_{i=1}^{k} H(t - t_i) \exp\left(-\frac{t - t_i}{\tau_w}\right), \tag{32}$$

$t \in [t_k, t_{k+1}]$, with Heaviside step function H, assuming the adaptation current just prior to the first spike is zero, $w(t_1^-) = 0$. This means, for given parameters μ, Δw, τ_w the effective mean input time series is determined by the observed spike train (up to t_k).

Note, that in general the mean input perturbations caused by adaptation vary from spike to spike, $\mu_1(t_k|\text{D}, \tau_w) \neq \mu_1(t_l|\text{D}, \tau_w)$ for $t_k \neq t_l \in \text{D}$. To efficiently evaluate the likelihood $p(\text{D}|\boldsymbol{\theta})$ using Eqs. (1) and (8) we calculate $p_{\text{ISI}}(s|\mu_{\text{ISI}}[0, s], \boldsymbol{\theta})$ with $\mu_{\text{ISI}}(s) = \mu(s) - w_0 \exp(-s/\tau_w)$, $s \geq 0$ for a reasonable range of values for w_0 and interpolate to obtain $p_{\text{ISI}}(s_k|\mu_{\text{ISI}}[0, s_k], \boldsymbol{\theta})$ with $\mu_{\text{ISI}}[0, s_k] = \mu[t_k, t_{k+1}] - \Delta w\, \mu_1[t_k, t_{k+1}]$ using Eq. (32).

9. Fitting I&F neurons to spike trains from electrophysiological recordings

9.1 Estimation of neuronal input statistics using in-vitro ground truth data

We used somatic whole-cell current clamp recordings from primary somatosensory cortex in acute brain slices (for details see [49]). Layer 5 PYRs were recorded in wild-type mice [49], fast-spiking layer 5 INTs were selected among the fluorescing cells of a GAD67-GFP transgenic line [107]. Only cells with an access resistance ≤25 MΩ (PYR: 18.3±1.5 MΩ, n = 7; INT: 19.5±4.0 MΩ, n = 6) and a drift in the resting membrane potential ≤7.5 mV (PYR: 3.2±3.0 mV, n = 7; INT: 3.1±3.7 mV, n = 6) throughout the recording were retained for further analysis. Seven PYRs and six INTs were stimulated with a fluctuating current $I(t)$ generated according to an Ornstein-Uhlenbeck process,

$$\frac{dI}{dt} = \frac{\mu_I - I}{\tau_I} + \sqrt{\frac{2}{\tau_I}}\sigma_I \xi(t), \tag{33}$$

where τ_I denotes the correlation time, μ_I and σ_I are the mean and standard deviation of the stationary normal distribution, i.e., $\lim_{t\to\infty} I(t) \sim \mathcal{N}(\mu_I, \sigma_I^2)$, and ξ is a unit Gaussian white noise process. Somatic current injections lasted 5 s and were separated by inter-stimulus intervals of at least 25 s. Different values for μ_I and σ_I were used and each combination was repeated three times. The correlation time was set to 3 ms. Spike times were defined by the time at which the membrane voltage crossed 0 mV from below, which was consistent with a large depolarization rate $dV/dt > 10$ mV/ms [49]. An absolute refractory period of 3 ms was assumed.

For each neuron we fitted a leaky I&F model with and without adaptation (cf. Methods sections 1 and 8). Note that the injected current $I(t)$ can be well approximated by a Gaussian white noise process as considered in our model because of the small correlation time τ_I. In Results section 2 we estimated the input parameters μ and σ for nonadaptive model neurons from each 5 s-long spike train recording as well as from each combined 3×5 s-long recording (using the three repetitions with identical stimulus parameters which effectively yielded 15 s long stimuli). To exclude onset transients (i.e., increased spike rate upon stimulus onset) we used the central 90% of ISIs for each stimulus, ensuring that ISIs lasted > 5 ms. For comparison we considered a Poisson process with constant rate. In Results section 5 we additionally estimated the adaptation parameters Δw and τ_w per neuron across all available stimuli in the combined 15 s stimulus setting. Here we used all ISIs (including the short ones at stimulus onset) in order to unmask adaptation effects. Parameter estimation was accomplished using method **1**. To compare the quality of the models and avoid over-fitting we used the AIC [53,108], given by $2N_\theta - 2\max_\theta \log p(\mathrm{D}|\boldsymbol{\theta})$, where N_θ denotes the number of estimated parameters (θ is a subvector of $\boldsymbol{\theta}$) for a particular model. For the adaptive I&F model $N_\theta = 4$, for the nonadaptive I&F model $N_\theta = 2$, and for the Poisson model $N_\theta = 1$. The preferred model from a set of candidate models is the one with the smallest AIC value.

9.2 Estimation of neuronal input statistics using data from awake animals

We used single unit spike trains from extracellular recordings of two adult female ferrets in an awake, spontaneous state. The animals were listening to acoustic stimuli separated by periods of silence lasting 0.4 s which we used for model fitting. Neural activity from primary auditory cortex was recorded using a 24 channel-electrode and spikes were sorted using an automatic clustering algorithm followed by a manual adjustment of the clusters (for details see [52]). Spike trains with > 50 ISIs during silence periods were considered for fitting. 71 single units passed that threshold in each of two behavioral conditions (passive listening vs. engaged in a discrimination task). Model neurons were fit in either behavioral condition separately, resulting in 142 sets of estimated parameters. We employed the leaky I&F model (Eqs. (5), (6)) with constant background input mean μ. For robust estimation we used the central 95% of ISIs, ensuring that ISIs lasted > 2.5 ms. For comparison we considered a Poisson process with constant rate and compared the quality of the models using the AIC.

9.3 Estimation of input perturbations using in-vitro ground truth data

We used whole-cell recordings of pyramidal neurons in slices of rat visual cortex. Ten neurons were stimulated with an input current that consisted of transient bumps reflecting an artificial excitatory postsynaptic current (aEPSC) immersed in background noise (for details see [50]: experiment 1). The background noise was generated as in Methods section 9.1 with $\tau_I = 5$ ms, μ_I tuned to maintain ~ 5 spikes/s and σ_I adjusted to produce membrane voltage fluctuations with $\sim$ 15-20 mV peak-to-peak amplitude. aEPSC traces were generated by convolving a simulated presynaptic spike train with a synaptic kernel described by a difference of exponentials (rise time 1 ms, decay time 10 ms). Presynaptic spikes were generated by a gamma renewal process (shape 2, scale 2.5) with 5 spikes/s on average; aEPSC amplitudes triggered by a single spike ranged from 0.1 σ_I to 1.5 σ_I. Current was injected in segments of 46 s length with at least 10 repetitions per aEPSC strength (except for one cell). The first and last 3 s of each segment were discarded from the analysis, as in [50].

Using only the presynaptic and postsynaptic spike times, we fitted an I&F neuron where input perturbations were described using delta pulses or alpha functions (cf. Methods section 6). For the former model we applied method **1a**, for the latter method **2**. Similarly to [50] we defined detection time as the minimal total length of spike train data for which the model with input perturbations ($J \neq 0$) yields a larger likelihood on test data compared to the respective model without input perturbations ($J = 0$); this was indicated by a positive log-likelihood ratio on test data from 5-fold cross-validation. In addition, we assessed detection performance on a fixed amount of data that consisted of five consecutive segments (i.e., 200 s recording duration), where adjacent 5-segment blocks shared one segment. Coupling strength z-scores were computed using estimates from surrogate data generated by perturbing the presynaptic spike times by a temporal jitter (cf. Methods section 7) in a large number of realizations.

9.4 Estimation of synaptic connections using in-vivo ground truth data

We used combined juxtacellular-extracellular recordings of neuronal ensembles from the hippocampal CA1 region in awake mice (for details see [51]). Neurons were separated into PYRs and INTs according to their spike waveform and spiking statistics. Spikes were evoked in single PYRs by short current pulses (50-100 ms) applied at intervals of variable length using juxtacellular electrodes while recording extracellular spikes of local INTs. PYR spikes which occurred during a stimulus were considered as evoked, and those which occurred at all other times were considered as spontaneous. All spikes that occurred during sharp-wave ripple events were discarded from the analyses and we only considered INTs that fired at least 3 spikes/s on average. A total of 78 PYR-INT pairs were included for estimation of synaptic couplings.

For each INT we fitted a leaky I&F neuron receiving background input and (potential) synaptic input from the recorded PYR such that each presynaptic spike causes a delayed (with delay d) postsynaptic potential of size J (cf. Eqs. (26)–(29) with $c = 0$, where we omit the indices i, j here for simplicity).

To account for changes in background input statistics over the recording duration, which lasted up to $\sim$2 h, and to reflect low-frequency network co-modulation induced by common network drive, the background mean input was allowed to vary over time. The parameters to be estimated are thus $\mu(t)$, J, d and σ. Estimation consisted of three steps. First, we inferred the statistics $\mu(t)$ and σ of background inputs (described by a non-stationary stochastic process) for $J = 0$ in the following way. We computed the empirical instantaneous spike rate $r(t)$ of the INT from the observed spike train via kernel density estimation using a Gaussian kernel with width $\sigma_{\mathrm{G}} \in \{0.1, 0.5, 1\}$ s. The estimated empirical spike rate varies over time much slower than the timescale at which changes of mean input translate to changes of spike rate in the I&F model. This justifies the approximation $r(t) \approx r_\infty(\mu(t)|\boldsymbol{\theta})$ (cf. Methods section 3), which allowed us to efficiently evaluate the spike train likelihood for fixed σ by applying method **1** with mean input assumed constant within each ISI, given by $\mu(t) = r_\infty^{-1}(r(t)|\boldsymbol{\theta})$ (at the center between consecutive spike times). The likelihood was then maximized with respect to σ. Given the parameters for the background inputs (one value of μ per ISI and one for σ) we next maximized the likelihood of the full model in the second step with respect to J and d using method **1a**. In the third step we assessed the significance of synaptic coupling estimates using surrogate data, similarly as in the previous section. We perturbed the presynaptic spike times by a small temporal jitter (random values between -5 and +5 ms) and re-estimated J and d. This was repeated 100 times and

z-scores were computed from the estimated coupling strengths. Notably, since spike times are shifted by only small values, effects due to network co-modulation which occurs on a slower timescale are preserved in the surrogate data. In this way we obtained a coupling strength z-score for each PYR-INT pair and for each of the three values of σ_{G}.

We validated our results against ground truth connection labels obtained from juxtacellular evoked activity using a model-free method based on spike train CCGs [51] (for details see SI section S4). Based on these labels we computed ROC curves as well as ACC and BACC (cf. Methods section 7.4) using a classification (z-score) threshold value J^z_{thresh}. Accordingly, the presence of an estimated connection was assured by the condition $\hat{J}^z > J^z_{\mathrm{thresh}}$, where $\hat{J}^z$ denotes the connection strength (z-score) estimate for a given PYR-INT pair. Note, that the ground truth labels indicate positive excitatory connections (positives) and absent connections (negatives).

To test the validity of our approach we estimated connectivity using only the first evoked PYR spikes of each stimulation pulse, which are maximally decoupled from network co-modulation, and compared the results with the ground truth labels. This assessment yielded excellent agreement, with ACC and BACC values of up to 0.97 and 0.95, respectively (for $\sigma_{\mathrm{G}} = 0.1$ s).

Code availability

Python code for our methods and estimation examples are available under a free license at https://github.com/neuromethods/inference-for-integrate-and-fire-models.

Data availability

No experimental data was collected for this study. The study involved available datasets from previous experimental studies with ethical approval granted. These datasets are available either online http://dx.doi.org/10.6084/m9.figshare.1144467 or from the authors on reasonable request.

Acknowledgments

We thank Dimitra Maoutsa for her support on the GLM implementation, and the anonymous reviewers for their constructive comments and helpful suggestions. This work was supported by Deutsche Forschungsgemeinschaft in the framework of Collaborative Research Center 910, the Programme Emergences of the City of Paris, ANR project MORSE (ANR-16-CE37-0016), and the program "Ecoles Universitaires de Recherche" launched by the French Government and implemented by the ANR, with the reference ANR-17-EURE-0017. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

J.L. and S.O. designed the study. J.L. led the project, developed and implemented the methods, performed the evaluations, validations and benchmark tests, and wrote the manuscript. S.M., D.E. and O.H. provided electrophysiological data. S.O. supervised the study and edited the manuscript. S.M. and O.H. provided feedback on the manuscript.

Competing interests

The authors declare no competing financial interests.

References

1. Chichilnisky EJ. A simple white noise analysis of neuronal light responses. Network. 2001;12:199–213.

2. Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects. J Neurophysiol. 2005;93:1074–1089.

3. Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, et al. Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature. 2008;454:995–999.

4. Latimer KW, Yates JL, Meister MLR, Huk AC, Pillow JW. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. Science. 2015;349:184–187.

5. Aljadeff J, Lansdell BJ, Fairhall AL, Kleinfeld D. Analysis of Neuronal Spike Trains, Deconstructed. Neuron. 2016;91:221–259.

6. Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol. 1952;117:500–544.

7. Gouwens NW, Berg J, Feng D, Sorensen SA, Zeng H, Hawrylycz MJ, et al. Systematic generation of biophysically detailed models for diverse cortical neuron types. Nat Commun. 2018;9:710.

8. Prinz AA, Billimoria CP, Marder E. Alternative to Hand-Tuning Conductance-Based Models: Construction and Analysis of Databases of Model Neurons. J Neurophysiol. 2003;90:3998–4015.

9. Marder E, Taylor AL. Multiple models to capture the variability in biological neurons and networks. Nat Neurosci. 2011;14:133–138.

10. Fourcaud-Trocmé N, Hansel D, van Vreeswijk C, Brunel N. How spike generation mechanisms determine the neuronal response to fluctuating inputs. J Neurosci. 2003;23:11628–11640.

11. Badel L, Lefort S, Brette R, Petersen CCH, Gerstner W, Richardson MJE. Dynamic I-V curves are reliable predictors of naturalistic pyramidal-neuron voltage traces. J Neurophysiol. 2008;99:656–666.

12. Platkiewicz J, Brette R. A threshold equation for action potential initiation. PLOS Comput Biol. 2010;6:e1000850.

13. Richardson MJE, Brunel N, Hakim V. From subthreshold to firing-rate resonance. J Neurophysiol. 2003;89:2538–2554.

14. Ladenbauer J, Augustin M, Obermayer K. How adaptation currents change threshold, gain and variability of neuronal spiking. J Neurophysiol. 2014;111:939–953.

15. Ostojic S, Szapiro G, Schwartz E, Barbour B, Brunel N, Hakim V. Neuronal Morphology Generates High-Frequency Firing Resonance. J Neurosci. 2015;35:7056–7068.

16. Ladenbauer J, Obermayer K. Weak electric fields promote resonance in neuronal spiking activity: analytical results from two-compartment cell and network models. bioRxiv preprint. 2018; p. 1–24.

17. Brette R, Gerstner W. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. J Neurophysiol. 2005;94:3637–3642.

18. Naud R, Marcille N, Clopath C, Gerstner W. Firing patterns in the adaptive exponential integrate-and-fire model. Biol Cybern. 2008;99:335–347.

19. Harrison PM, Badel L, Wall MJ, Richardson MJE. Experimentally Verified Parameter Sets for Modelling Heterogeneous Neocortical Pyramidal-Cell Populations. PLOS Comput Biol. 2015;11:e1004165.

20. Teeter C, Iyer R, Menon V, Gouwens N, Feng D, Berg J, et al. Generalized leaky integrate-and-fire models classify multiple neuron types. Nat Commun. 2018;9:709.

21. Jolivet R, Schürmann F, Berger TK, Naud R, Gerstner W, Roth A. The quantitative single-neuron modeling competition. Biol Cybern. 2008;99:417–426.

22. Gerstner W, Naud R. How Good Are Neuron Models? Science. 2009;326:379–380.

23. Pospischil M, Piwkowska Z, Bal T, Destexhe A. Comparison of different neuron models to conductance-based post-stimulus time histograms obtained in cortical pyramidal cells using dynamic-clamp in vitro. Biol Cybern. 2011;105:167–180.

24. Pozzorini C, Mensi S, Hagens O, Naud R, Koch C, Gerstner W. Automated High-Throughput Characterization of Single Neurons by Means of Simplified Spiking Models. PLOS Comput Biol. 2015;11:e1004275.

25. de Solages C, Szapiro G, Brunel N, Hakim V, Isope P, Buisseret P, et al. High-Frequency Organization and Synchrony of Activity in the Purkinje Cell Layer of the Cerebellum. Neuron. 2008;58:775–788.

26. Giridhar S, Doiron B, Urban NN. Timescale-dependent shaping of correlation by olfactory bulb lateral inhibition. Proc Natl Acad Sci USA. 2011;108:5843–5848.

27. Litwin-Kumar A, Chacron MJ, Doiron B. The Spatial Structure of Stimuli Shapes the Timescale of Correlations in Population Spiking Activity. PLOS Comput Biol. 2012;8:e1002667.

28. Potjans TC, Diesmann M. The cell-type specific cortical microcircuit: Relating structure and activity in a full-scale spiking network model. Cereb Cortex. 2014;24:785–806.

29. Bendor D. The role of inhibition in a computational model of an auditory cortical neuron during the encoding of temporal information. PLOS Comput Biol. 2015;11:e1004197.

30. Blot A, de Solages C, Ostojic S, Szapiro G, Hakim V, Léna C. Time-invariant feed-forward inhibition of Purkinje cells in the cerebellar cortex in vivo. J Physiol. 2016;10:2729–2749.

31. Kanashiro T, Ocker GK, Cohen MR, Doiron B. Attentional modulation of neuronal variability in circuit models of cortex. eLife. 2017;6:1–38.

32. Brunel N, Hakim V. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. Neural Comput. 1999;11:1621–1671.

33. Brunel N. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. J Comput Neurosci. 2000;8:183–208.

34. Izhikevich EM, Edelman G. Large-scale model of mammalian thalamocortical systems. Proc Natl Acad Sci USA. 2008;105:3593–3598.

35. Litwin-Kumar A, Doiron B. Slow dynamics and high variability in balanced cortical networks with clustered connections. Nat Neurosci. 2012;15:1498–1505.

36. Doiron B, Litwin-Kumar A, Rosenbaum R, Ocker GK, Josić K. The mechanics of state-dependent neural correlations. Nat Neurosci. 2016;19:383–393.

37. Schmuker M, Pfeil T, Nawrot MP. A neuromorphic network for generic multivariate data classification. Proc Natl Acad Sci USA. 2014;111:2081–2086.

38. Gütig R. Spiking neurons can discover predictive features by aggregate-label learning. Science. 2016;351:aab4113.

39. Gilra A, Gerstner W. Predicting non-linear dynamics by stable local learning in a recurrent spiking neural network. eLife. 2017;6:e28295.

40. Bellec G, Salaj D, Subramoney A, Legenstein R, Maass W. Long short-term memory and learning-to-learn in networks of spiking neurons. In: Advances in Neural Information Processing Systems; 2018. p. 1–14.

41. Neftci E, Binas J, Rutishauser U, Chicca E, Indiveri G, Douglas RJ. Synthesizing cognition in neuromorphic electronic systems. Proc Natl Acad Sci USA. 2013;110:E3468–E3476.

42. Nawrocki RA, Voyles RM, Shaheen SE. A Mini Review of Neuromorphic Architectures and Implementations. IEEE Trans Electron Devices. 2016;63:3819–3829.

43. Davies M, Srinivasa N, Lin TH, Chinya G, Cao Y, Choday SH, et al. Loihi: A Neuromorphic Manycore Processor with On-Chip Learning. IEEE Micro. 2018;38:82–99.

44. Gerstner W, Kistler WM, Naud R, Paninski L. Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition. Cambridge, UK: Cambridge University Press; 2014.

45. Hawrylycz M, Anastassiou C, Arkhipov A, Berg J, Buice M, Cain N, et al. Inferring cortical function in the mouse visual system through large-scale systems neuroscience. Proc Natl Acad Sci USA. 2016;113:7337–7344.

46. Druckmann S, Banitt Y, Gidon A, Schürmann F, Markram H, Segev I. A novel multiple objective optimization framework for constraining conductance-based neuron models by experimental data. Front Neurosci. 2007;1:7–18.

47. Rossant C, Goodman DFM, Fontaine B, Platkiewicz J, Magnusson AK, Brette R. Fitting neuron models to spike trains. Front Neurosci. 2011;5:1–8.

48. Stringer C, Pachitariu M, Steinmetz NA, Okun M, Bartho P, Harris KD, et al. Inhibitory control of correlated intrinsic variability in cortical networks. eLife. 2016;5:1–33.

49. Mensi S, Hagens O, Gerstner W, Pozzorini C. Enhanced Sensitivity to Rapid Input Fluctuations by Nonlinear Threshold Dynamics in Neocortical Pyramidal Neurons. PLOS Comput Biol. 2016;12:e1004761.

50. Volgushev M, Ilin V, Stevenson IH. Identifying and Tracking Simulated Synaptic Inputs from Neuronal Firing: Insights from In Vitro Experiments. PLOS Comput Biol. 2015;11:e1004167.

51. English DF, McKenzie S, Evans T, Kim K, Yoon E, Buzsáki G. Pyramidal Cell-Interneuron Circuit Architecture and Dynamics in Hippocampal Networks. Neuron. 2017;96:505–520.

52. Bagur S, Averseng M, Elgueda D, David S, Fritz J, Yin P, et al. Go/No-Go task engagement enhances population representation of target stimuli in primary auditory cortex. Nat Commun. 2018;9:2529.

53. Millar RB. Maximum Likelihood Estimation and Inference. Wiley; 2011.

54. Brunel N, Van Rossum MCW. Lapicque's 1907 paper: From frogs to integrate-and-fire. Biol Cybern. 2007;97:337–339.

55. Gigante G, Mattia M, Giudice P, Del Giudice P. Diverse population-bursting modes of adapting spiking neurons. Phys Rev Lett. 2007;98:148101.

56. Richardson MJE. Spike-train spectra and network response functions for non-linear integrate-and-fire neurons. Biol Cybern. 2008;99:381–92.

57. Augustin M, Ladenbauer J, Baumann F, Obermayer K. Low-dimensional spike rate models derived from networks of adaptive integrate-and-fire neurons: comparison and implementation. PLOS Comput Biol. 2017;13:e1005545.

58. Ostojic S, Brunel N. From spiking neuron models to linear-nonlinear models. PLOS Comput Biol. 2011;7:e1001056.

59. Mattia M, Del Giudice P. Population dynamics of interacting spiking neurons. Phys Rev E. 2002;66:051917.

60. Burkitt AN. A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. Biol Cybern. 2006;95:1–19.

61. Paninski L, Pillow JW, Simoncelli EP. Maximum Likelihood Estimation of a Stochastic Integrate-and-Fire Neural Encoding Model. Neural Comput. 2004;16:2533–2561.

62. Pillow JW, Paninski L, Uzzell VJ, Simoncelli EP, Chichilnisky EJ. Prediction and Decoding of Retinal Ganglion Cell Responses with a Probabilistic Spiking Model. J Neurosci. 2005;25:11003–11013.

63. Tuckwell HC. Introduction to Theoretical Neurobiology. Cambridge, UK: Cambridge University Press; 1988.

64. Ecker AS, Berens P, Keliris GA, Bethge M, Logothetis NK, Tolias AS. Decorrelated neuronal firing in coritcal micorcircuits. Science. 2010;327:584–587.

65. Cohen MR, Kohn A. Measuring and interpreting neuronal correlations. Nat Neurosci. 2011;14:811–819.

66. Zaytsev YV, Morrison A, Deger M. Reconstruction of recurrent synaptic connectivity of thousands of neurons from simulated spiking activity. J Comput Neurosci. 2015;39:77–103.

67. Stevenson IH, London BM, Oby ER, Sachs NA, Reimer J, Englitz B, et al. Functional Connectivity and Tuning Curves in Populations of Simultaneously Recorded Neurons. PLOS Comput Biol. 2012;8:e1002775.

68. Gerhard F, Deger M, Truccolo W. On the stability and dynamics of stochastic spiking neuron models: Nonlinear Hawkes process and point process GLMs. PLOS Comput Biol. 2017;13:e1005390.

69. Madison DV, Nicoll RA. Control of the repetitive discharge of rat CA1 pyramidal neurones in vitro. J Physiol. 1984;354:319–331.

70. Schwindt PC, Spain WJ, Foehring RC, Chubb MC, Crill WE. Slow conductances in neurons from cat sensorimotor cortex in vitro and their role in slow excitability changes. J Neurophysiol. 1988;59:450–467.

71. Stocker M. Ca(2+)-activated K+ channels: molecular determinants and function of the SK family. Nat Rev Neurosci. 2004;5:758–770.

72. Schwindt PC, Spain WJ, Crill WE. Calcium-dependent potassium currents in neurons from cat sensorimotor cortex. J Neurophysiol. 1992;67:216–226.

73. La Camera G, Rauch A, Thurbon D, Lüscher HR, Senn W, Fusi S. Multiple time scales of temporal response in pyramidal and fast spiking cortical neurons. J Neurophysiol. 2006;96:3448–3464.

74. Oliphant TE. Python for Scientific Computing. Comput Sci Eng. 2007;9:10–20.

75. Lam SK, Pitrou A, Seibert S. Numba: A LLVM-based python JIT compiler. In: Proc. LLVM Compil. Infrastruct. HPC; 2015. p. 1–6.

76. Dong Y, Mihalas S, Russell A, Etienne-Cummings R, Niebur E. Estimating Parameters of Generalized Integrate-and-Fire Neurons from the Maximum Likelihood of Spike Trains. Neural Comput. 2011;23:2833–2867.

77. Mullowney P, Iyengar S. Parameter estimation for a leaky integrate-and-fire neuronal model from ISI data. J Comput Neurosci. 2008;24:179–194.

78. Kim H, Shinomoto S. Estimating nonstationary input signals from a single neuronal spike train. Phys Rev E. 2012;86:1–12.

79. Carlson KD, Nageswaran JM, Dutt N, Krichmar JL. An efficient automated parameter tuning framework for spiking neural networks. Front Neurosci. 2014;8:1–15.

80. Lueckmann JM, Goncalves PJ, Bassetto G, Öcal K, Nonnenmacher M, Macke JH. Flexible statistical inference for mechanistic models of neural dynamics. In: Advances in Neural Information Processing Systems; 2017. p. 1–18.

81. Cocco S, Leibler S, Monasson R. Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. Proc Natl Acad Sci USA. 2009;106:14058–14062.

82. Monasson R, Cocco S. Fast inference of interactions in assemblies of stochastic integrate-and-fire neurons from spike recordings. J Comput Neurosci. 2011;31:199–227.

83. Vidne M, Ahmadian Y, Shlens J, Pillow JW, Kulkarni J, Litke AM, et al. Modeling the impact of common noise inputs on the network activity of retinal ganglion cells. J Comput Neurosci. 2012;33:97–121.

84. Mensi S, Naud R, Gerstner W. From Stochastic Nonlinear Integrate-and-Fire to Generalized Linear Models. In: Advances in Neural Information Processing Systems; 2011. p. 1–9.

85. Pernice V, Rotter S. Reconstruction of sparse connectivity in neural networks from spike train covariances. J Stat Mech. 2013;3.

86. Pastore VP, Massobrio P, Godjoski A, Martinoia S. Identification of excitatory-inhibitory links and network topology in large-scale neuronal assemblies from multi-electrode recordings. PLOS Comput Biol. 2018;14:e1006381.

87. Casadiego J, Maoutsa D, Timme M. Inferring Network Connectivity from Event Timing Patterns. Phys Rev Lett. 2018;121:054101.

88. Ostojic S, Brunel N, Hakim V. How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. J Neurosci. 2009;29:10234–53.

89. Pernice V, Staude B, Cardanobile S, Rotter S. How structure determines correlations in neuronal networks. PLOS Comput Biol. 2011;7:e1002059.

90. Trousdale J, Hu Y, Shea-Brown E, Josić K. Impact of network structure and cellular response on spike time correlations. PLOS Comput Biol. 2012;8:e1002408.

91. Tetzlaff T, Helias M, Einevoll GT, Diesmann M. Decorrelation of Neural-Network Activity by Inhibitory Feedback. PLOS Comput Biol. 2012;8:e1002596.

92. Rosenbaum R, Smith MA, Kohn A, Rubin JE, Doiron B. The spatial structure of correlated neuronal variability. Nat Neurosci. 2017;20:107–114.

93. Ocker GK, Josić K, Shea-Brown E, Buice MA. Linking structure and activity in nonlinear spiking networks. PLOS Comput Biol. 2017;13:e1005583.

94. Huang C, Ruff DA, Pyle R, Rosenbaum R, Cohen MR, Doiron B. Circuit Models of Low-Dimensional Shared Variability in Cortical Networks. Neuron. 2019;101:1–12.

95. Ocker GK, Hu Y, Buice MA, Doiron B, Josić K, Rosenbaum R, et al. From the statistics of connectivity to the statistics of spike times in neuronal networks. Curr Opin Neurobiol. 2017;46:109–119.

96. Brinkman BAW, Rieke F, Shea-Brown E, Buice MA. Predicting how and when hidden neurons skew measured synaptic interactions. PLOS Comput Biol. 2018;14:e1006490.

97. Richardson M. Effects of synaptic conductance on the voltage distribution and firing rate of spiking neurons. Phys Rev E. 2004;69:051918.

98. Donner C, Opper M, Ladenbauer J. Inferring the dynamics of neural populations from single-trial spike trains using mechanistic models. Cosyne Abstr, Lisbon, PT; bioRxiv preprint. 2019; p. 1–22.

99. Ghanbari A, Malyshev A, Volgushev M, Stevenson IH. Estimating short-term synaptic plasticity from pre- and postsynaptic spiking. PLOS Comput Biol. 2017;13:e1005738.

100. Stevenson I, Koerding K. Inferring spike-timing-dependent plasticity from spike train data. In: Advances in Neural Information Processing Systems; 2011. p. 1–9.

101. Linderman SW, Stock CH, Adams RP. A Framework for Studying Synaptic Placticity with Neural Spike Train Data. In: Advances in Neural Information Processing Systems; 2014. p. 1–9.

102. Risken H. The Fokker-Planck Equation: Methods of Solutions and Applications. Berlin: Springer; 1996.

103. Ostojic S. Inter-spike interval distributions of spiking neurons driven by fluctuating inputs. J Neurophysiol. 2011;106:361–373.

104. Schaffer ES, Ostojic S, Abbott LF. A complex-valued firing-rate model that approximates the dynamics of spiking networks. PLOS Comput Biol. 2013;9:e1003301.

105. Nelder JA, Mead R. A Simplex Method for Function Minimization. Comput J. 1965;7:308–313.

106. Stimberg M, Goodman DFM, Benichoux V, Brette R. Equation-oriented specification of neural models for simulations. Front Neuroinform. 2014;8:1–14.

107. Tamamaki N, Yanagawa Y, Tomioka R, Miyazaki JI, Obata K, Kaneko T. Green Fluorescent Protein Expression and Colocalization with Calretinin, Parvalbumin, and Somatostatin in the GAD67-GFP Knock-In Mouse. J Comp Neurol. 2003;467:60–79.

108. Akaike H. A New Look at the Statistical Model Identification. IEEE Trans Autom Control. 1974;19:716–723.

109. Gosse L. Computing qualitatively correct approximations of balance laws. vol. 2. Springer; 2013.

110. LeVeque RJ. Finite volume methods for hyperbolic problems. Cambridge University Press; 2002.

Supplementary Information

S1. Exponential integrate-and-fire neuron model

The exponential I&F model is specified by Eqs. (5) and (6) with function f defined by

$$f(V) := \frac{\Delta_\mathrm{T}}{\tau_\mathrm{m}} \exp\left(\frac{V - V_\mathrm{T}}{\Delta_\mathrm{T}}\right) - \frac{V}{\tau_\mathrm{m}}, \tag{34}$$

where Δ_T is the threshold slope factor and V_T denotes the effective threshold voltage. The exponential term in Eq. (34) effectively approximates the rapidly increasing Na^+ current at spike initiation [10] and yields an improved fit to intracellular measurements of current-voltage relationships (typically obtained in-vitro) compared to the leaky I&F model [11, 19]. The value of V_s does not play an important role as long as it is sufficiently larger than the effective threshold V_T. In fact, when V increases beyond V_T, it diverges to infinity in finite time due to the exponential term, which defines a spike. In practice, however, the spike is said to occur when V reaches V_s. The exponential I&F model used here is further equipped with an absolute refractory period of length T_ref, during which $V(t)$ is clamped at V_r.

A change of V_T or Δ_T in the exponential I&F model can be completely compensated in terms of spiking dynamics by appropriate changes of μ, σ, V_r and V_s. This can be seen using the change of variables $\tilde{V} := (V - V_\mathrm{T})/\Delta_\mathrm{T}$. Consequently, the relevant parameters for inference from spike train data are μ, σ, τ_m, V_r and T_ref.

S2. Numerical solution for ISI density p_ISI

Below we present two numerical solution schemes: the first one uses a finite volume discretization and is generally applicable for arbitrary variations of the input parameters, the second one uses the Fourier transform and is suitable for small amplitude variations of the parameters. Depending on the model scenario the latter scheme may be computationally more efficient than the former. Python code for both numerical schemes is available at: https://github.com/neuromethods/inference-for-integrate-and-fire-models

S2.1 Solution based on finite volume discretization

This method employs a recently developed finite volume numerical scheme with implicit time discretization and Scharfetter-Gummel flux approximation, adapted from [57] for the first passage time problem. It provides an accurate solution of Eqs. (9)–(13) for arbitrary variations of the mean input. In brief, we first discretize the domain $[V_\mathrm{lb}, V_\mathrm{s}]$ into N_V equidistant grid cells $[V_{m-\frac{1}{2}}, V_{m+\frac{1}{2}}]$ with centers V_m, $m \in \{1, \ldots, N_V\}$, $V_1 < V_2 < \cdots < V_{N_V}$, where $V_{\frac{1}{2}} = V_\mathrm{lb}$ and $V_{N_V+\frac{1}{2}} = V_\mathrm{s}$ are the out-most cell borders. Within each cell the numerical approximation of $p_V(V, s)$ is assumed to be constant and corresponds to the average value denoted by $p_V(V_m, s)$. Integrating Eq. (9) over the volume of cell m, and applying the divergence theorem, yields

$$\frac{\partial}{\partial s} p_V(V_m, s) = \frac{q_V(V_{m-\frac{1}{2}}, s) - q_V(V_{m+\frac{1}{2}}, s)}{\Delta V}, \tag{35}$$

where $\Delta V = V_2 - V_1$. To solve Eq. (35) forward in time (represented by the ISI variable s) the fluxes at the borders of each cell are approximated using the first order Scharfetter-Gummel flux [109],

$$q_V(V_{m+\frac{1}{2}}, s) = v_{m+\frac{1}{2}}(s) \frac{p_V(V_m, s) - p_V(V_{m+1}, s) \exp\left(-v_{m+\frac{1}{2}}(s)\Delta V/D\right)}{1 - \exp\left(-v_{m+\frac{1}{2}}(s)\Delta V/D\right)}, \tag{36}$$

where $v_{m+\frac{1}{2}}(s) = f(V_{m+\frac{1}{2}}) + \mu_\mathrm{ISI}(s)$ and $D = \frac{1}{2}\sigma^2$ denote the drift and diffusion coefficients, respectively.

For the time discretization we rewrite Eq. (35) (with Eq. (36)) in vectorized form and approximate the involved time derivative as first order backward difference to ensure numerical stability. This yields in each

time step of length Δs a linear system for the vector $\mathbf{p}^{n+1}$ of probability density values at s_{n+1}, given the values $\mathbf{p}^n$ at the previous time step s_n, with vector elements $p_m^n = p_V(V_m, s_n)$,

$$(\mathbf{I} - \frac{\Delta s}{\Delta V}\,\mathbf{G}^n)\mathbf{p}^{n+1} = \mathbf{p}^n, \tag{37}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{G}^n \in \mathbb{R}^{N_V \times N_V}$ is a tridiagonal matrix that contains the discretization of the membrane voltage (cf. Eqs. (35), (36)), including the absorbing and reflecting boundary conditions (Eqs. (12) and (13)).

The ISI density, Eq. (14), in this representation is obtained by evaluating the flux, Eq. (36), at the spike voltage V_s, taking into account the absorbing boundary condition, Eq. (12), and introducing an auxiliary ghost cell [110] with center V_{N_V+1}, which yields

$$p_\mathrm{ISI}(s_{n+1}|\mu_\mathrm{ISI}[0, s_n], \boldsymbol{\theta}) = q_V(V_{N_V+\frac{1}{2}}, s_{n+1}) = v_{N_V+\frac{1}{2}}(s_n)\,\frac{1+\exp(-v_{N_V+\frac{1}{2}}(s_n)\Delta V/D)}{1-\exp(-v_{N_V+\frac{1}{2}}(s_n)\Delta V/D)}\,p_{N_V}^{n+1}. \tag{38}$$

For additional details we refer to [57] (incl. supplemental material therein). Note that this method also allows for the variation of other parameters, in addition to the mean input, within the ISI. Naturally, the finite volume scheme can be used to compute the first order approximation (15) for small amplitude variations of the mean input in a straightforward way: using $J = 0$ for p_ISI^0, and a small value $J > 0$ to calculate p_ISI, to obtain $p_\mathrm{ISI}^1 \approx (p_\mathrm{ISI} - p_\mathrm{ISI}^0)/J$. An alternative method to compute p_ISI^0 and p_ISI^1, which may be computationally more efficient in certain cases, is presented in the next section.

S2.2 Solution for weak input perturbations based on Fourier transform

The method described below is inspired by [56] and provides a solution of Eqs. (9)–(13) for constant mean input or small amplitude variations of it through a perturbative approach. It extends the technique proposed in [56] to the case of time-varying mean input within the ISI, and can be adjusted for small amplitude variations of other parameters in a straightforward way.

We consider $\mu_\mathrm{ISI}(s) = \mu_\mathrm{ISI}^0 + J\mu_\mathrm{ISI}^1(s)$ with small $|J|$ according to Eq. (16). The solution of the first passage time system to first order in J can be expressed as $p_V(V, s) \approx p_V^0(V, s) + Jp_V^1(V, s)$. We insert this in the Fokker-Planck system, Eqs. (9)–(13), and neglect terms of order > 1 in J. The resulting system is then Fourier-transformed over time using separation of variables, $p_V(V, s) = x(V)y(s)$, such that $\hat{p}_V(V;\omega) = \mathcal{F}[p_V(V, s)] = x(V)\hat{y}(\omega)$, with Fourier transform and the inverse transform defined by

$$\hat{y}(\omega) := \mathcal{F}[y(s)] = \int_{-\infty}^{\infty} y(s)e^{-\mathrm{i}\omega s}ds, \qquad y(s) = \mathcal{F}^{-1}[\hat{y}(\omega)] = \int_{-\infty}^{\infty} \frac{\hat{y}(\omega)}{2\pi}e^{\mathrm{i}\omega s}d\omega, \tag{39}$$

where ω denotes angular frequency. This yields the following two (coupled) systems, one for the steady-state solution $\hat{p}_V^0(V;\omega)$,

$$\frac{d\hat{q}_V^0}{dV} = -\mathrm{i}\omega\hat{p}_V^0 \tag{40}$$

$$\frac{d\hat{p}_V^0}{dV} = \frac{2}{\sigma^2}\Big([f(V) + \mu_\mathrm{ISI}^0]\,\hat{p}_V^0 - \hat{q}_V^0\Big) \tag{41}$$

subject to

$$\hat{p}_V^0(V_\mathrm{s};\omega) = 0, \qquad \lim_{V\to-\infty}\hat{q}_V^0(V;\omega) = 0 \tag{42}$$

$$\lim_{V\searrow V_\mathrm{r}}\hat{q}_V^0(V;\omega) - \lim_{V\nearrow V_\mathrm{r}}\hat{q}_V^0(V;\omega) = 1 \tag{43}$$

and one for $\hat{p}_V^1(V;\omega)$,

$$\frac{d\hat{q}_V^1}{dV} = -\mathrm{i}\omega\hat{p}_V^1 \tag{44}$$

$$\frac{d\hat{p}_V^1}{dV} = \frac{2}{\sigma^2}\Big([f(V) + \mu_\mathrm{ISI}^0]\,\hat{p}_V^1 + \hat{h}(V;\omega) - \hat{q}_V^1\Big) \tag{45}$$

with function $\hat{h}(V;\omega) = \mathcal{F}[\mu^1_{\mathrm{ISI}}(s)p^0_V(V,s)]$, subject to

$$\hat{p}^1_V(V_{\mathrm{s}};\omega) = 0, \qquad \lim_{V\to-\infty} \hat{q}^1_V(V;\omega) = 0. \tag{46}$$

The ISI density (in the frequency domain) is then given by

$$\hat{p}_{\mathrm{ISI}}(\omega) \approx \hat{p}^0_{\mathrm{ISI}}(\omega) + J\hat{p}^1_{\mathrm{ISI}}(\omega) = \hat{q}^0_V(V_{\mathrm{s}};\omega) + J\hat{q}^1_V(V_{\mathrm{s}};\omega). \tag{47}$$

We solve the linear ordinary differential equations (40)–(46) for each fixed frequency ω by splitting the solutions,

$$\hat{p}^0_V(V;\omega) = \hat{p}^0_{\mathrm{ISI}}(\omega)\hat{p}^0_\alpha(V;\omega) + \hat{p}^0_\beta(V;\omega), \qquad \hat{q}^0_V(V;\omega) = \hat{p}^0_{\mathrm{ISI}}(\omega)\hat{q}^0_\alpha(V;\omega) + \hat{q}^0_\beta(V;\omega) \tag{48}$$

and analogously for $\hat{p}^1_V(V;\omega)$, $\hat{q}^1_V(V;\omega)$. The components $(\hat{p}^0_\alpha, \hat{q}^0_\alpha)$ and $(\hat{p}^0_\beta, \hat{q}^0_\beta)$ respectively solve Eqs. (40)–(41) with $\hat{p}^0_\alpha = 0$, $\hat{q}^0_\alpha = 1$ and $\hat{p}^0_\beta = \hat{q}^0_\beta = 0$ at $V = V_{\mathrm{s}}$ and $\lim_{V\searrow V_{\mathrm{r}}} \hat{q}^0_\beta - \lim_{V\nearrow V_{\mathrm{r}}} \hat{q}^0_\beta = 1$. The components $(\hat{p}^1_\alpha, \hat{q}^1_\alpha)$ and $(\hat{p}^1_\beta, \hat{q}^1_\beta)$ solve Eqs. (44)–(45) with $\hat{p}^1_\alpha = 0$, $\hat{q}^1_\alpha = 1$ and $\hat{p}^1_\beta = \hat{q}^1_\beta = 0$ at $V = V_{\mathrm{s}}$. These solution components can be conveniently computed by backward integration from V_{s} to a sufficiently small lower bound $V_{\mathrm{lb}} < V_{\mathrm{r}}$. We then obtain $\hat{p}^0_{\mathrm{ISI}}$ and $\hat{p}^1_{\mathrm{ISI}}$ by satisfying the reflecting boundary conditions (in (42) and (46)):

$$\hat{p}^0_{\mathrm{ISI}}(\omega) = -\frac{\hat{q}^0_\beta(V_{\mathrm{lb}};\omega)}{\hat{q}^0_\alpha(V_{\mathrm{lb}};\omega)}, \qquad \hat{p}^1_{\mathrm{ISI}}(\omega) = -\frac{\hat{q}^1_\beta(V_{\mathrm{lb}};\omega)}{\hat{q}^1_\alpha(V_{\mathrm{lb}};\omega)}. \tag{49}$$

The ISI density in the time domain is finally calculated by the inverse transform, $p_{\mathrm{ISI}}(s|\mu_{\mathrm{ISI}}[0,s],\boldsymbol{\theta}) = \mathcal{F}^{-1}[\hat{p}_{\mathrm{ISI}}(\omega)]$ using Eqs. (47) and (49).

In practice, we first solve the steady-state system (40)–(43), evaluate the function $\hat{h}$ which appears in Eq. (45) and then solve the system for the change in ISI density due to the perturbation, (44)–(46). Note that knowledge of $\mu^1_{\mathrm{ISI}}(s)$ for $s \geq 0$ is required to calculate $\hat{h}$. If the input perturbations are given by delta pulses (as for the network in Results section 4) the calculation of $\hat{h}$ is greatly simplified; e.g., for a pulse at $s = s_0$, $\mu^1_{\mathrm{ISI}}(s) = \delta(s - s_0)$ we have $\hat{h}(V;\omega) = e^{-\mathrm{i}\omega s_0} p^0_V(V, s_0)$.

S3. Related I&F-based method for benchmarks

We considered the approach from [62] for comparison. This method infers the parameters of a generalized leaky I&F model with spike history dependence in response to a fluctuating background input and a time-varying stimulus by maximizing the spike train likelihood. For this purpose an approximation based on the Fokker-Planck equation is used [61]. The method is applicable to the model scenarios in Results sections 2, 3 and 5. We ensured equal conditions for the benchmarks in terms of model dynamics, parameters to be estimated, amount of observed data and computational resources. An efficient implementation in Matlab with integrated C/C++ code for accelerated program execution was used for the method from [62]; the code is available at http://pillowlab.princeton.edu/code_GIF.html. All tests were performed on the same personal computer.

Specifically, the model from [62] adapted for our benchmarks is described by Eqs. (5) and (7), where the time-varying mean input $\mu(t)$ consists of three additive components: a constant baseline μ, perturbations given by the convolution of the kernel $k(t)$ with the sequence of perturbation pulses $x(t) = \sum_l \delta(t - \tilde{t}_l)$, i.e., $k * x(t) := \int_0^\infty k(\tau)x(t-\tau)d\tau$, and spike history-dependent input given by the convolution of kernel $h(t)$ with the spike train $s(t) = \sum_k \delta(t - t_k)$, i.e., $h * s(t)$. The kernels $k(t)$ and $h(t)$ were parametrically described using a particular set of basis functions (raised cosines with log scaling of the time axis; for details see [62]). To ensure equal numbers of estimated parameters for the different methods we used two basis functions for each kernel, whose weights were the parameters to be inferred. The basis functions were selected to yield, when equally weighted, a kernel shape that is similar to the shape of the respective equivalent kernel in our description. The selected basis functions and resulting kernels are shown in Fig. S3e and S8d.

For benchmarks on inference of background inputs (cf. Results section 2.1) we omitted the perturbation and spike history components, $k(t) = h(t) = 0$; for inference of input perturbations (cf. Results section 3.1) we set $h(t) = 0$ and measured detection sensitivity as for methods **1a** and **2**; for inference of neuronal adaptation (cf. Results section 5.1) we set $k(t) = 0$.

S4. Ground truth connection labels from in-vivo data

(Mono)synaptic connections were assumed to produce excess synchrony above baseline co-modulation in a short interval following PYR spikes. To generate estimated ground truth connection labels we adopted the approach in [51]. A CCG (0.4 ms binning) was computed for each PYR-INT pair using the evoked PYR spikes. For positive labels the peak of the CCG in the interval $[0.8, 2.8]$ ms needed to exceed that from the slowly co-modulated baseline, and it needed to be significantly larger than the largest peak in the anticausal direction (short negative lags). The lower frequency baseline was computed by convolving the observed CCG with a "partially hollow" Gaussian kernel, with a standard deviation of 10 ms and a hollow fraction of 60%. We estimated the probability of obtaining a spike count sc of n as observed (or higher) in the mth time lag within $[0.8, 2.8]$ ms of the CCG, given the expected, low frequency baseline rate $\lambda_{\text{s}}(m)$ using the Poisson distribution with a continuity correction,

$$P_{\text{fast}}(sc(m) \geq n|\lambda_{\text{s}}(m)) = 1 - \sum_{k=0}^{n-1} \frac{\lambda_{\text{s}}^{k}(m)e^{-\lambda_{\text{s}}(m)}}{k!} - \frac{\lambda_{\text{s}}^{n}(m)e^{-\lambda_{\text{s}}(m)}}{n!\,2}. \tag{50}$$

Similarly, we estimated the probability of obtaining the observed spike count n (or higher) in the mth time lag within $[0.8, 2.8]$ ms of the CCG as expected from the maximum rate λ_{a} across negative lags within $[-2, 0]$ ms using the Poisson distribution with a continuity correction,

$$P_{\text{causal}}(sc(m) \geq n|\lambda_{\text{a}}) = 1 - \sum_{k=0}^{n-1} \frac{\lambda_{\text{a}}^{k}e^{-\lambda_{\text{a}}}}{k!} - \frac{\lambda_{\text{a}}^{n}e^{-\lambda_{\text{a}}}}{n!\,2}. \tag{51}$$

Connections were labeled as synapses if $P_{\text{fast}} < 0.001$ and $P_{\text{causal}} < 0.0026$ for all (binned) lags in $[0.8, 2.8]$ ms, according to the rigorous criterion defined in [51].

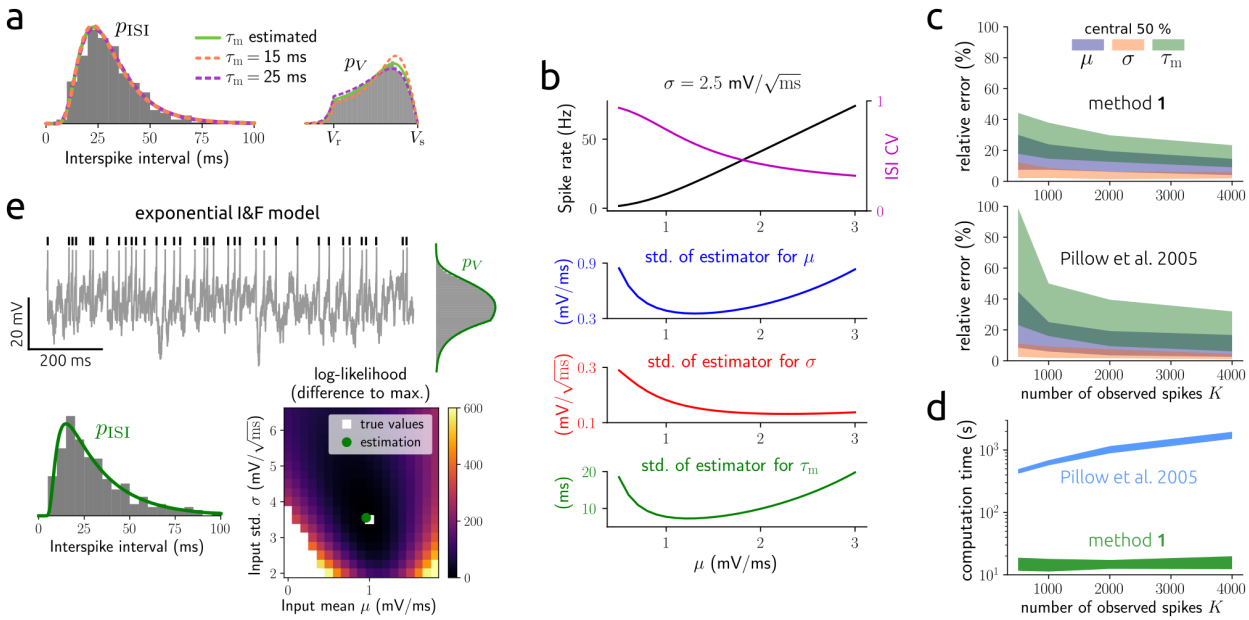

Figure S1. Estimation and benchmark results using simulated data of single-cell spontaneous activity. **a**: ISI and membrane voltage histograms corresponding to the example in Fig. 1a, together with the densities p_{ISI} and p_V for τ_{m} estimated or set to wrong values (true $\tau_{\mathrm{m}} = 20$ ms). **b**, top: spike rate and ISI CV as a function of μ; below: standard deviation of estimates for μ, σ and τ_{m}, respectively, according to the Cramer-Rao bound for $K = 1000$. **c**: 25th-75th percentile (central 50 %) of relative errors between estimated and true parameter values for μ, σ and τ_{m} as a function of number of spikes; top: results from method **1**, bottom: results from the method described in [62]. **d**: central 50 % of computation times corresponding to the estimations in **c** (for details see SI section S3). **e**: same as Fig. 1a,b for the exponential I&F model instead of the leaky I&F model. Membrane voltage histogram and p_V reflect non-refractory times. Parameter values were $V_{\mathrm{s}} = 30$ mV, $V_{\mathrm{r}} = 0$ mV, $\tau_{\mathrm{m}} = 20$ ms, $\Delta_{\mathrm{T}} = 1.5$ mV, $V_{\mathrm{T}} = 15$ mV, $\mu = 1$ mV/ms, $\sigma = 3.5$ mV/$\sqrt{\mathrm{ms}}$, $\Delta_{\mathrm{T}} = 1.5$ mV, $T_{\mathrm{ref}} = 3$ ms and $K = 400$.

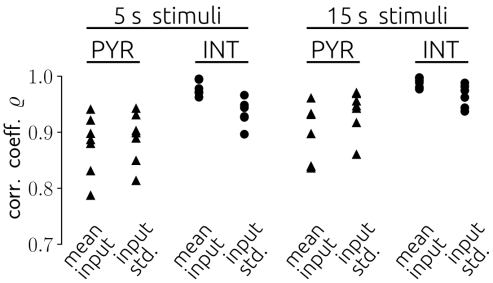

Figure S2. Correlation between estimated input parameters and empirical input statistics from in-vitro ground truth data. Pearson correlation coefficient ϱ between estimated and empirical input statistics for all PYRs and INTs, and for two stimulus durations: 5 s-long simuli (cf. Fig. 2c) and 15 s-long stimuli (corresponding to 3 repetitions of 5 s stimuli with equal statistics).

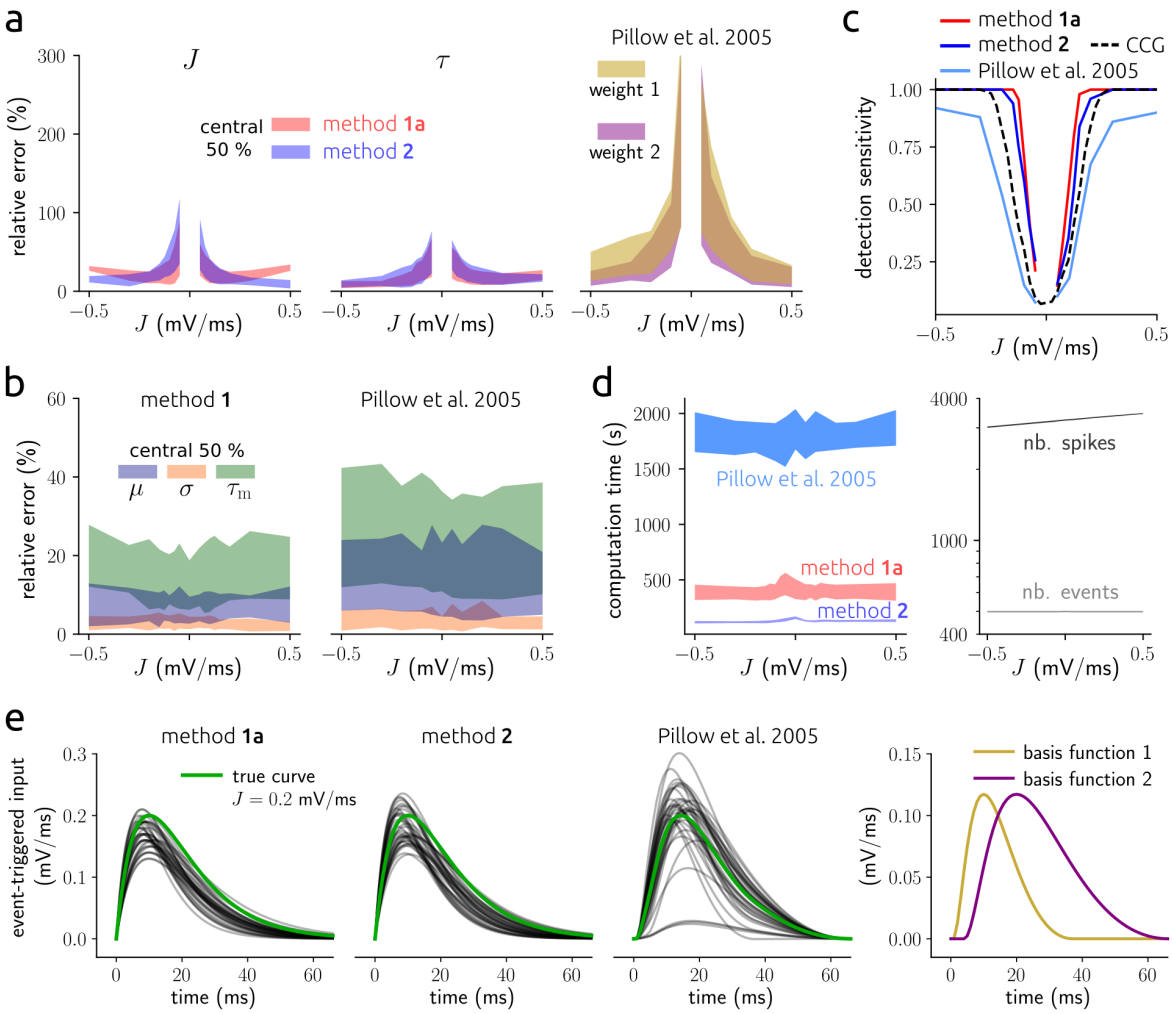

Figure S3. Estimation and benchmark results for input perturbations using simulated data. a: 25th-75th percentile (central 50 %) of relative errors between estimated and true parameter values for J (left) and τ (center) as a function of true J using methods **1a** (red) and **2** (blue), and for the two perturbation parameters using the method from [62] (right). Estimates are based on spike trains of length 100 s (for details see SI section S3). True $\tau = 10$ ms. **b**: central 50 % of relative errors between estimated and true parameter values for μ, σ and τ_{m} from our approach (left) and the one from [62] (right) using the same data as in **a**. Note that the parameters of input perturbations as well as μ, σ and τ_{m} were simultaneously inferred. **c**: detection sensitivity as a function of J for the methods used in **a** (solid lines) and a CCG-based method (dashed line, cf. Fig. 3d). **d**: central 50 % of computation times (left) and numbers of observed spikes and events (right) as a function of J, corresponding to the estimations in **a**–**c**. Note that slight asymmetries in **a**–**d** for excitatory compared to inhibitory perturbations may be due to differences in the number of observed spikes. **e**, left: true time series of mean input perturbation triggered at time 0 for $J = 0.2$ mV/ms (green) and reconstructions from the methods used in **a** (gray; best 75% according to the combined relative errors in **a**); right: selected basis functions for the method from [62].

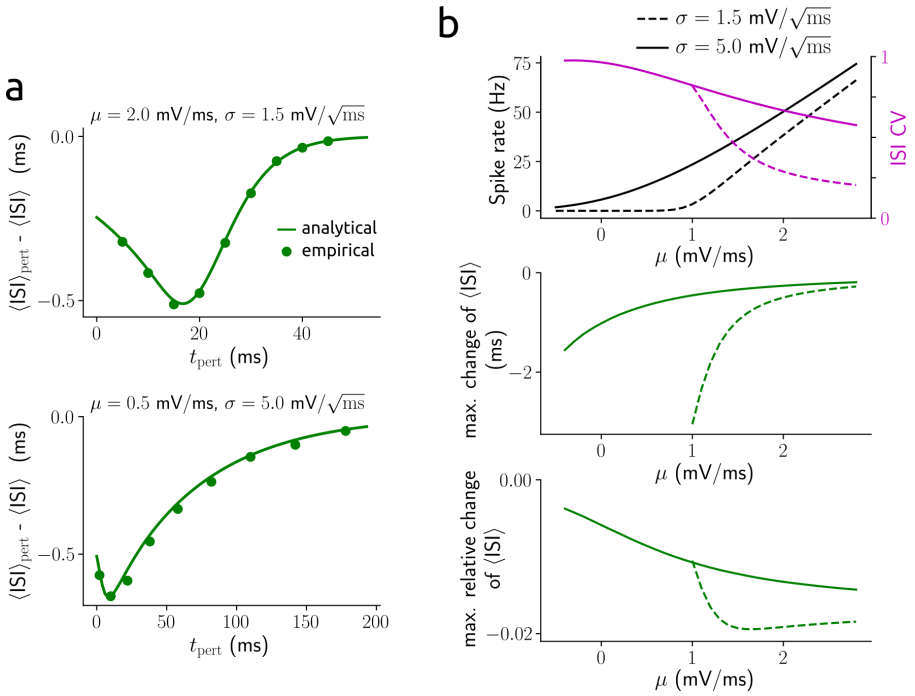

Figure S4. Impact of weak input perturbations on the expected ISI. a: change of expected ISI due to a weak input perturbation that causes a sudden deflection of 0.5 mV in the membrane voltage ($\langle \text{ISI} \rangle_{\text{pert}}$ - $\langle \text{ISI} \rangle$) as a function of time t_{pert} within the ISI at which the perturbation occurs, for different background input statistics. **b**, top: spike rate and ISI CV as a function of μ for different values of σ (cf. Fig. 1e, symbols indicate the parametrizations in **a**); below: corresponding maximal (relative) change of expected ISI caused by an input perturbation (as in **a**).

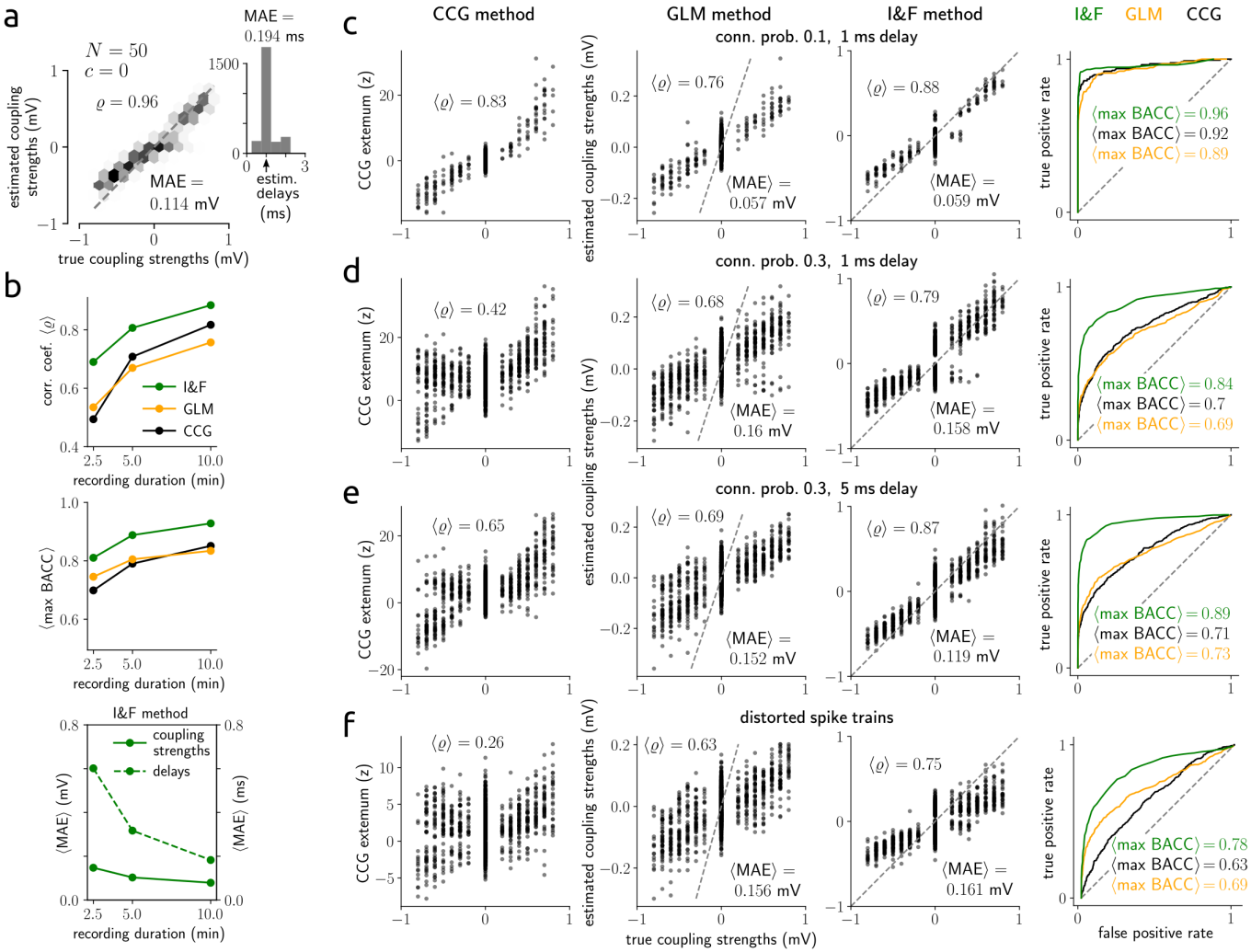

Figure S5. Additional estimation results for synaptic coupling strengths using simulated data. **a**: density of estimated vs. true coupling strengths from a network of $N = 50$ neurons, with Pearson correlation coefficient ϱ and mean absolute error MAE between true and estimated coupling strengths indicated; otherwise as in Fig. 5b. **b**: average correlation coefficient $\langle\varrho\rangle$ (top) and average maximal BACC ⟨maxBACC⟩ (center) across the 5 networks in Fig. 5g as a function of recording duration for the three methods (I&F, GLM, CCG); average mean absolute error ⟨MAE⟩ between true and estimated coupling strengths and delays, respectively, for the I&F method (bottom). **c**–**f**: detailed inference results from the three methods (as shown in Fig. 5d, 5e right and 5f) corresponding to the networks in Fig. 5h(i)–(iv) as indicated. Results from each 5 networks are shown; ROC curves represent averages.

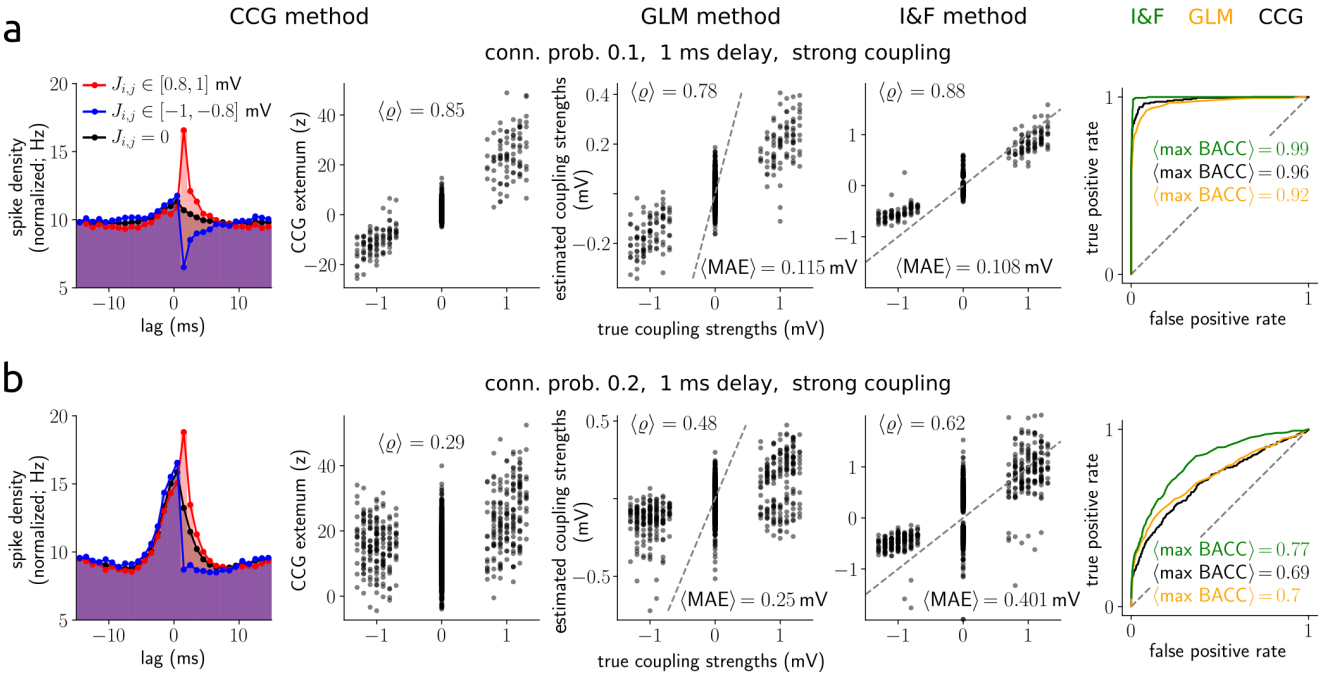

Figure S6. Estimation results for strong synaptic coupling using simulated data. a,**b**: inference results from the three methods (as shown in Fig. 5d, 5e right, 5f and 5h left) for subsampled networks of $N_{\text{tot}} = 1000$ (500 excitatory, 500 inhibitory) neurons with connection probability of 0.1 (**a**) and 0.2 (**b**); otherwise as in Fig. 5h(i). Results from each 5 networks are shown; ROC curves represent averages.

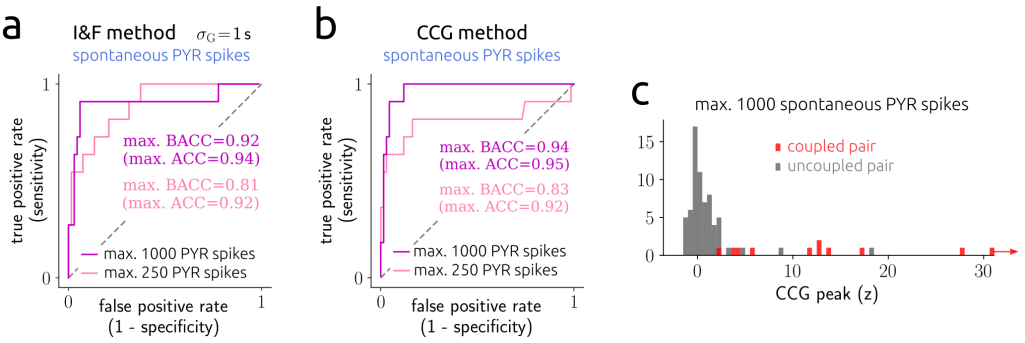

Figure S7. Additional estimation results for synaptic couplings using in-vivo data. a: ROC curves across 78 PYR-INT pairs using the spontaneous PYR spikes for $\sigma_G = 1$ s. Maximal (balanced) accuracy values are indicated. **b**: ROC curves as in **a**, using the model-free CCG method instead of the I&F method. **c**: histogram of CCG peak z-score values for all 78 PYR-INT pairs with true labels (derived using the evoked spikes) indicated.

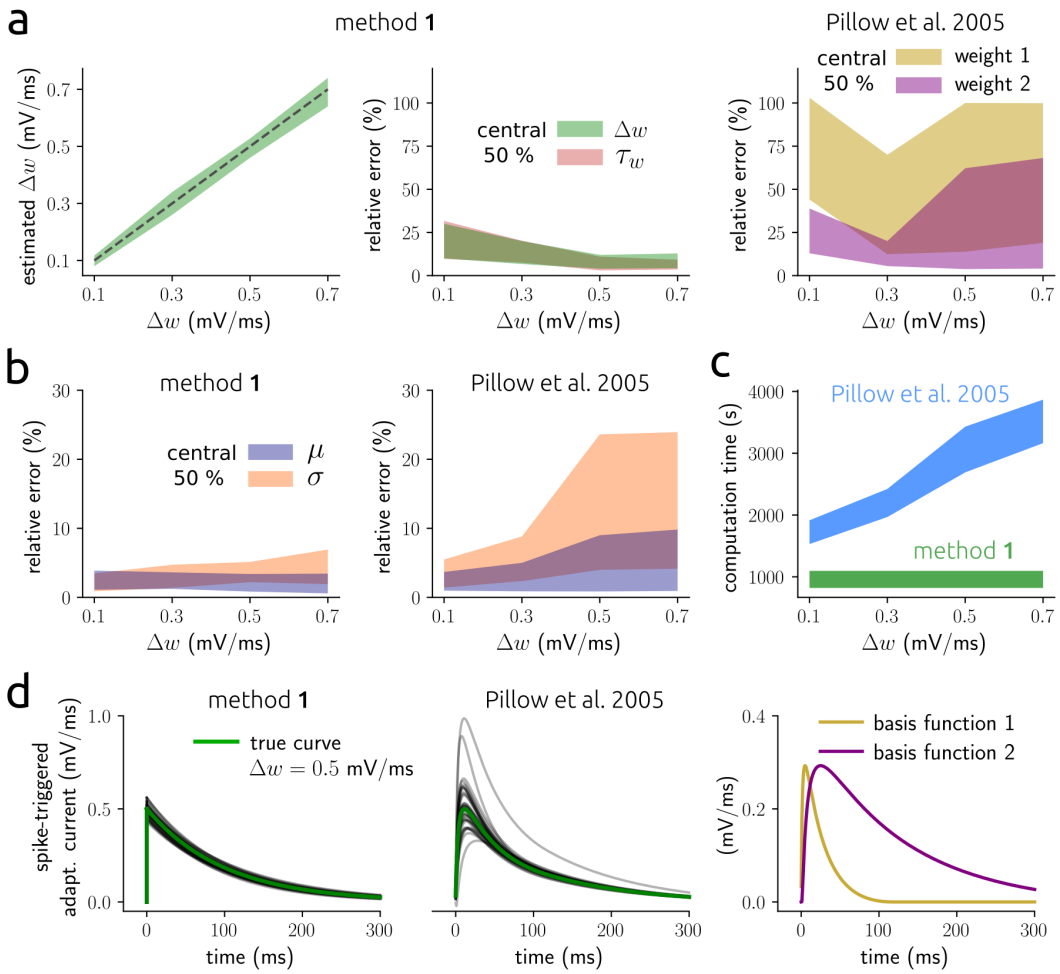

Figure S8. Estimation and benchmark results for adaptive model neurons using simulated data. a, left: 25th-75th percentile (central 50 %) of estimated values for Δw as a function of true Δw using method **1**; center, right: central 50 % of relative errors between estimated and true parameter values for Δw and τ_w (center), and for the two adaptation parameters using the method from [62] (right). Estimates are based on 2000 spikes (for details see SI section S3). True $\tau_w = 100$ ms. **b**: central 50 % of relative errors between estimated and true parameter values for μ and σ from method **1** (left) and the method from [62] (right) using the same data as in **a**. Note that the adaptation parameters as well as μ and σ were simultaneously inferred. **c**: central 50 % of computation times as a function of Δw, corresponding to the estimations in **a** and **b**. **d**, left: true time series of adaptation current triggered by a spike at time 0 for $\Delta w = 0.5$ mV/ms (green) and reconstructions from the methods used in **a** (gray; best 75% according to the combined relative errors in **a**); right: selected basis functions for the method from [62].

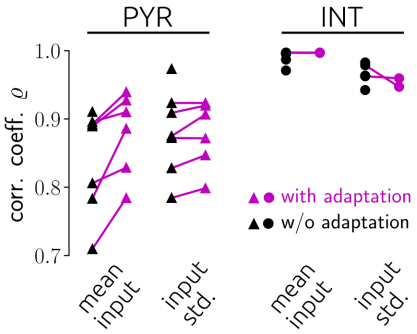

Figure S9. Correlation between estimated input parameters and empirical input statistics for adaptive and nonadaptive models. Pearson correlation coefficient ϱ between estimated and empirical input statistics for the I&F model with adaptation (magenta symbols) and without adaptation (black symbols) for all PYRs and INTs.

Table S1. Computation times for estimation examples.

Figure	1a	3a†		5a‡	5c‡	7b
Method	**1**	**1a**	**2**	**1a**	**1a**	**1**
Multiprocessing	no	no	no	yes	yes	no
Estimation time (min)	< 0.2	< 10	< 2	< 17	< 117	< 7

All computations were performed on a hexa-core personal computer. †Applies to each of the two examples. ‡Computation time scaled sub-linearly with recording time (hence with number of observed spikes) and supra-linearly with number of observed neurons.