Reconstruction of post-synaptic potentials by reverse modeling of local field potentials

Yochum Maxime[1], Modolo Julien[1*], Benquet Pascal[1], Wendling Fabrice[1],

1 LTSI, INSERM, U1099, Université Rennes 1, Rennes, France

* julien.modolo@inserm.fr

Abstract

Among electrophysiological signals, Local Field Potentials (LFPs) are extensively used to study brain activity, either *in vivo* or *in vitro*. LFPs are recorded with extracellular electrodes implanted in brain tissue. They reflect intermingled excitatory and inhibitory processes in neuronal assemblies. In cortical structures, LFPs mainly originate from the summation of post-synaptic potentials (PSPs), either excitatory (ePSPs) and inhibitory (iPSPs) generated at the level of pyramidal cells. The challenging issue, addressed in this paper, is to estimate, from a single extracellularly-recorded signal, both ePSP and iPSP components of the LFP. The proposed method is based on a model-based reverse engineering approach in which the measured LFP is fed into a physiologically-grounded neural mass model (mesoscopic level) in order to estimate the synaptic activity of a sub-population of pyramidal cells interacting with local GABAergic interneurons. The method was first validated using simulated LFPs for which excitatory and inhibitory components are known *a priori* and can thus serve as a ground truth. It was then evaluated on *in vivo* data (PTZ-induced seizures, rat; PTZ-induced excitability increase, mouse; epileptiform discharges, mouse) and on *in clinico* data (human seizures recorded with depth-EEG electrodes). Under these various conditions, results showed that the proposed reverse engineering method provides a reliable estimation of the average excitatory and inhibitory post-synaptic potentials at the origin of the measured LFPs. They also indicated that the method allows for monitoring of the excitation/inhibition ratio. The method has potential for multiple applications in neuroscience, typically when a time tracking of local excitability changes is required.

Author summary

The measurement of excitatory and inhibitory processes which control brain excitability is a key issue in brain research. Indeed, alteration of excitation/inhibition balance is involved in both normal (like plasticity-related changes in response to learning) and pathological (like epileptic seizures) conditions. We report a novel method to estimate the two main components of the local field potential, namely the excitatory and inhibitory post-synaptic potentials. This method starts from extracellularly-recorded signals and uses a physiologically-relevant neural mass model to constrain the local field potential decomposition into post-synaptic potentials. To our knowledge, this neural mass model-based reverse engineering method is novel as it constitutes the first attempt to combine neural mass modeling with optimization techniques to extract hidden excitatory and inhibitory components from local field potentials recorded by extracellular macroelectrodes. A second progress beyond the state-of-the-art is the

integration of more accurate physiology in the neural mass model as it now accounts for specific input-output functions that characterize glutamatergic principal neurons and GABAergic interneurons. Results demonstrated that the proposed method enables real-time tracking of neural excitability with essential information about dynamic changes of the excitation/inhibition ratio possibly impaired in the epileptic tissue. New epilepsy treatment can benefit from this method as it can help finding optimal control strategies based, for instance, on deep brain stimulation.

Introduction

Local field potentials (LFPs) refer to extracellularly-recorded electrophysiological signals generated by neuron assemblies [1]. LFPs are extensively used to monitor and analyze brain function in brain research as well as in neurological clinical studies, Since LFP recordings require minimal equipment and provide critical information on local neuronal activity, they are used in a wide spectrum of applications, from basic neuroscience to clinics [2] (e.g., using the LFP as a control signal to trigger stimulation in closed-loop [3]).

Physiologically, LFP signals reflect ongoing excitation- and inhibition-related processes in recorded networks [4]. Indeed, it is known that the major contribution to LFPs arises from post-synaptic potentials (PSPs) generated at the level of pyramidal cells within spatially extended neuronal assemblies. As described by bioelectromagnetic models [5, 6] and experimental studies [4], the summation of PSPs, either excitatory (ePSPs) or inhibitory (iPSPs) generated at the level of pyramidal cells located in the cerebral cortex is the major contribution to LFPs. This is essentially explained by two factors: (1) synaptic activation leads to the formation of a sink and a source at the level of neurons, which can then be viewed as elementary current dipoles [4], and (2) when neurons are geometrically aligned [4, 7] (like pyramidal cells organized "in palissade" in cortical structures), then dipole contributions tend to sum up instead of cancelling out [4]. In addition, due to the frequency-filtering properties of cerebral tissue, slow post-synaptic currents are more likely to be recorded further away from the electrode contacts, as opposed to fast transmembrane ionic currents (involved in action potentials) which attenuate much faster with space [8]. Overall, LFPs mainly result from an averaging process of synaptic currents, both excitatory (glutamatergic) and inhibitory (GABAergic).

Therefore, LFPs may, in theory, give access to underlying synaptic currents and, subsequently, provide insights about the excitability level of recorded neuronal local networks. To reach this goal, the challenging issue is to estimate from a single extracellularly-recorded signal both ePSP and iPSP components of the LFP. In this paper, we propose a novel model-based approach to solve this problem. The method is based on model-based reverse engineering approach in which a neural mass model (NMM) is used to deconstruct the extracellularly-recorded field activity and reveal its main excitatory and inhibitory components. This reverse engineering method introduces two main steps. First, the classical NMM was revisited to solve an ill-posed problem, since two parameters (ePSPs and iPSPs amplitude) are estimated from a single signal (LFP recording) measured over a finite time window. Second, an optimization procedure was developed allowing for time tracking of ePSPs and iPSPs over a short-duration sliding window.

As demonstrated from both simulated and experimental data, the proposed approach can reliably estimate and monitor the average post-synaptic potentials and the densities of action potentials arising from underlying excitatory and inhibitory subsets of neurons. To our knowledge, this is the first study solving the inverse problem of LFP to estimate its main synaptic components. Since this method also provides access to the

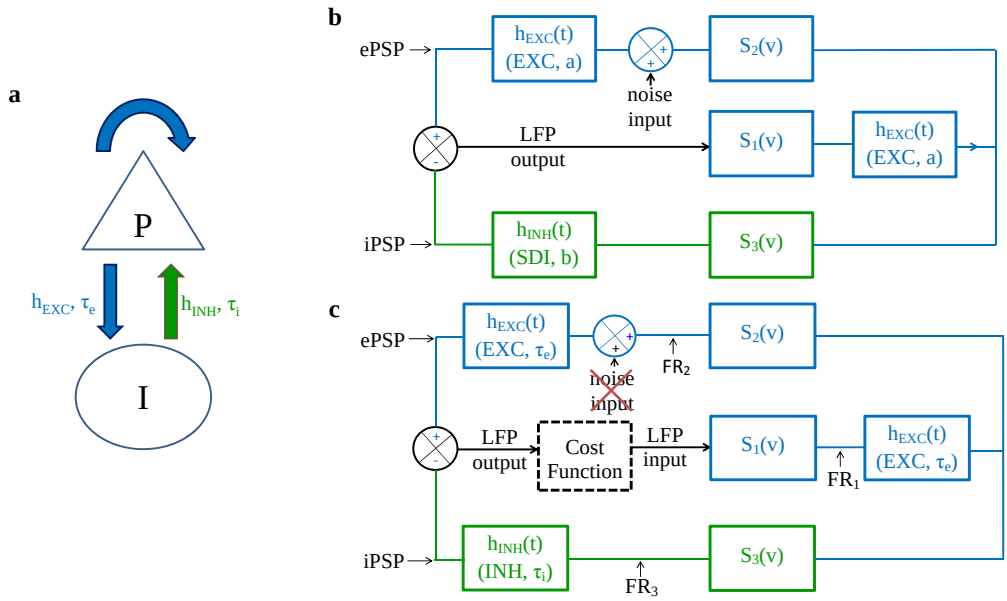


Fig 1. Illustration of the reverse modeling approach. (**a**) Neural mass model block diagram including two sub-populations: pyramidal cells (P) and interneurons (I). Excitatory (blue) and inhibitory (green) links between sub-populations are represented with arrows. (**b**) Neural mass model diagram where sigmoid functions were modified and connectivity parameters from the original NMM were removed. Excitatory and inhibitory loops are represented respectively in blue and green. (**c**) Reverse modeling approach where the new neural mass model is used to estimate PSPs contained in an LFP signal, the latter being set as an input of the model. A cost function based on the RMS error between estimated $\widehat{LFP}$ (LFP output) and input LFP was introduced to improve identification of EXC and INH gain parameters.

excitatory/inhibitory balance, the applications encompass diverse areas of neuroscience, ranging from basic neuroscience (tracking of excitability changes in neuronal networks) to clinics (analysis of depth-EEG recordings in patients with epilepsy).

Materials and methods

Principle of the model-based LFP reconstruction

The method exploits *a priori* knowledge about the processes (type and kinetics) underlying LFP generation in order to constrain LFP decomposition. LFPs are mostly composed of two sub-components (excitatory and inhibitory), and their extraction is challenging since 1) reconstructing two signals from only one is by definition an ill-posed problem; and 2) their frequency range is similar. In order to overcome this roadblock, we used a physiologically relevant NMM as a constraint to enable this decomposition process. NMM are a well-established computational class of models of neuronal population activity considering synaptic interactions between the neuronal assemblies involved in each LFP component: pyramidal cells (excitation) and interneurons (inhibition). We then used an LFP (experimental or simulated) to constrain the NMM response, and exploited this NMM to extract information contained in the LFP, especially to reconstruct excitatory and inhibitory processes.

Pyramidal cells			
References	Sigmoid function parameters		
	e0	r	v0
Reid et al. [9]	48.9	0.534	5.12
Chen et al. [10]	26.7	0.721	5.81
Malik et al. [11]	20.5	0.435	5.81
Malik et al. [11]	112	0.356	6.62
Wang et al. [12]	19.1	0.550	6.67
Mean	45.4	0.519	6.00

Interneurons			
References	Sigmoid function parameters		
	e0	r	v0
Gibson et al. [13]	173	0.166	17.7
Ferguson et al. [14]	128	0.239	12.0
Fidzinsk et al. [15]	195	0.320	11.1
Fidzinski et al. [15]	154	0.210	13.1
Erisir et al. [16]	66.3	0.376	10.6
Mean	143	0.262	12.9

Table 1. Identified sigmoid function parameters.

Revised version of the neural mass model (NMM)

The NMM used in this study is an established and validated neural population model [17] including two sub-populations corresponding to pyramidal cells for excitation, and interneurons for inhibition. The corresponding block diagram of the model is presented in Fig. 1, where excitatory/inhibitory processes are represented in blue/green, respectively. In order to improve the physiological relevance of the NMM, we adapted the parameters of the "wave-to-pulse" sigmoid functions, converting the mean depolarization of sub-populations into an output firing rate, according to available neurophysiological data from the literature. We also reduced the number of parameters to be identified through the reverse modeling process by removing the connectivity parameters, which were instead integrated in the new sigmoid functions and in the synaptic gain parameters of the transfer functions. The main difference with the original Jansen-Rit's model [17] is therefore the absence of connectivity parameters. Table 2 summarizes model parameters values. The resulting equations were derived following the method presented by Touboul et al. [18]. The synaptic impulse response functions used in the model were under the standard form: $h_{EXC}(t) = EXC \cdot t \cdot e^{\frac{-t}{\tau_e}}$ for excitation, and $h_{INH}(t) = INH \cdot t \cdot e^{\frac{-t}{\tau_i}}$ for inhibition, leading to a set of 6 differential

Parameter	Interpretation	Value
EXC	Average excitatory synaptic gain	60
INH	Average inhibitory synaptic gain	15
1/a	Dendritic average time constant in the feedback excitatory loop	1/100 s
1/b	Dendritic average time constant in the feedback inhibitory loop	1/35 s
V_{0e}, e_{0e}, r_e	Parameters of the sigmoid function for pyramidal cells	V_{0e} = 6 mV e_{0e} = 45,4 Hz r_e=0.519 mV^{-1}
V_{0i}, e_{0i}, r_i	Parameters of the sigmoid function for interneurons	V_{0i} = 12,9 mV e_{0i} = 143 Hz r_i = 0.262 mV^{-1}
p(t)	Excitatory input noise (Gaussian white noise)	m = 90 Hz s = 30 Hz

Table 2. Model parameters.

equations forming the full NMM:

$$\begin{aligned}
\dot{y}_0(t) &= y_3(t) \\
\dot{y}_1(t) &= y_4(t) \\
\dot{y}_2(t) &= y_5(t) \\
\dot{y}_3(t) &= EXC \cdot aS_1\left(y_1(t) - y_2(t)\right) - 2ay_3(t) - a^2 y_0(t) \\
\dot{y}_4(t) &= EXC \cdot a\left[p(t) + S_2\left(y_0(t)\right)\right] - 2ay_4(t) - a^2 y_1(t) \\
\dot{y}_5(t) &= INH \cdot bS_3\left(y_0(t)\right) - 2by_5(t) - b^2 y_2(t)
\end{aligned} \tag{1}$$

Table 1 lists the studies that were used to identify sigmoid function parameters corresponding to pyramidal cells and interneurons. Since, in these studies, the pulse-to-wave relation was studied with respect to current, the first step was to convert the data with respect to voltage. To do so, a current-to-voltage conversion was used according to Destexhe and Paré's study [19], which used Ohm's law with the membrane resistance. Then, a non-linear least square (Levenberg-Marquardt algorithm) was used to fit data from the literature, and an average was made between the two available sets to obtain sigmoid function parameters used in our model. The three sigmoid functions that were included in the NMM are:

$$\begin{aligned}
S_1 = S_2 &= \frac{e_{0_e}}{1 + e^{r_e\left(v_{0_e} - v\right)}}, with \begin{cases} e_{0_e} = 45.4 \\ r_e = 0.519 \\ v_{0_e} = 6 \end{cases} \\
S_3 &= \frac{e_{0_i}}{1 + e^{r_i\left(v_{0_i} - v\right)}}, with \begin{cases} e_{0_i} = 143 \\ r_i = 0.262 \\ v_{0_i} = 12.9 \end{cases}
\end{aligned} \tag{2}$$

Estimation of excitatory and inhibitory post-synaptic currents

We verified that this novel NMM could reproduce the four possible types of activity of the original model, namely: background activity, sporadic epileptic spikes, rhythmic spikes and narrow band activity at theta and alpha frequencies (see S1 Fig. in Supporting information). The reverse modeling approach uses LFPs (simulated or recorded experimentally) that are injected in the NMM. The originality in our approach is to force the NMM state at the node of the associated block diagram (see Fig. 1c),

corresponding to where LFPs are generated ("LFP input"), forcing other model components to use the LFP that has been injected as an input. By comparing the block diagrams in Fig. 1b and 1c, the LFP input of the reverse modeling approach (Fig. 1c) corresponds to the LFP output of the NMM (Fig. 1b).

The injected LFP was first normalized to ensure that its magnitude was similar to simulated LFPs. The physical unit of LFPs, as well as ePSP and iPSP, is in terms of Volts, and the normalization allows not considering differences in hardware amplification gain use for recording LFPs. The reverse modeling approach involved removing the noise input (s=0) to avoid adding a stochastic component to the process, and also to correctly adapt $\widehat{\mathrm{EXC}}$ and $\widehat{\mathrm{INH}}$ parameters to fit optimally the input LFP with the estimated $\widehat{\mathrm{LFP}}$. Third, in order to ensure that NMM dynamics corresponds to the injected LFP, a key step consists in comparing the NMM output $\widehat{\mathrm{LFP}}$ (estimated LFP computed in the reverse modeling approach) with the injected LFP, and to identify NMM gain parameters maximizing the match between LFP and $\widehat{\mathrm{LFP}}$. The identified gain parameters are $\widehat{\mathrm{EXC}}$ for the excitatory loop and $\widehat{\mathrm{INH}}$ for the slow inhibitory loop. This identification was performed using a gradient descent, where the cost function was chosen as the root mean square (RMS) error between LFP and $\widehat{\mathrm{LFP}}$ in magnitude (M_{RMS}) and derivative (d_{RMS}):

$$M_{RMS} = \sqrt{\frac{\sum_{n=1}^{N}\left(LFP_n - \widehat{LFP}_n\right)^2}{N}}, \tag{3}$$

$$d_{RMS} = \sqrt{\frac{\sum_{n=1}^{N}\left(\frac{dLFP_n}{dt} - \frac{d\widehat{LFP}_n}{dt}\right)^2}{N}}, \tag{4}$$

$$RMSE = M_{RMS} + d_{RMS}, \tag{5}$$

where N is the number of sample in the LFP.

Using both the difference in magnitude and derivative between LFP and $\widehat{\mathrm{LFP}}$ in the cost function enables evaluating the similarity in magnitude, but also in tendency. In the gradient descent method, both $\widehat{\mathrm{EXC}}$ and $\widehat{\mathrm{INH}}$ parameters were set to initial values (in a 2D parameter space) and the estimated $\widehat{\mathrm{LFP}}$ was computed for different set of parameters around the initial ones (on an ellipse where both radii were defined on the 2D parameter space d_1 and d_2). The gradient descent algorithm ended when no error lower than the current parameter set was found, and current $\widehat{\mathrm{EXC}}$ and $\widehat{\mathrm{INH}}$ parameters were then taken as the optimal identification. It should be noted that this method does not identify all the basins of attraction in a 2D parameter space. However, a comparison with a 2D grid parameters space (where errors are computed for each set of parameter in 2D grid, which becomes significantly more computationally extensive) led to the conclusion that only one basin of attraction was found for the great majority of tested LFP signals (not shown).

Custom software was designed in Python, integrating a Qt interface (See S1 video in Supporting information), dedicated to fitting at best NMM parameters on an injected LFP along short sliding windows on the time axis (i.e., NMM gain parameters are identified once per sliding window). Excitatory ($\widehat{\mathrm{EXC}}$) and inhibitory ($\widehat{\mathrm{INH}}$) gains could therefore be tracked during the LFP time course, along with the excitation to inhibition ratio ($\widehat{\mathrm{EIR}}$, computed as the $\widehat{\mathrm{EXC}}/\widehat{\mathrm{INH}}$ ratio). Obviously, the NMM provides full access at any time to the intermediate states for each sub-population, such as the average firing rate of action potentials and average post-synaptic potentials at the level of pyramidal cells and interneurons. Such signals are challenging to extract from LFPs with standard signal processing methods, especially since LFPs originate from the

summation of PSPs onto pyramidal cells, and are not directly related with APs at the soma level. Using our approach, FRs can be directly deduced from a LFP with the reverse modeling approach, since FRs are intermediate states in the NMM.

Goodness-of-fit

The method used to compute the goodness of fit between an LFP and a reconstructed LFP was the zero normalized cross correlation (ZNCC [20]). This method provides an indication of similarity between two signals, named γ, which is bound [-1;1]. This indicator is 1 when both signals are identical, -1 when one signal is the opposite of the other, and is close to 0 when both signals are far from similar. This indicator does not depend neither on the length of the signals nor on the amplitude scales of different LFP measures (different amplifier gains for instance). Therefore, the ZNCC is especially suited to compare results for different LFP time/amplitude scales.

In vivo recordings

Electrophysiological recordings from N=7 mice were performed in accordance with the European Community Council Directive of November 24th 1986 (86/609/EEC), and were approved by the local ethics committee from the University of Rennes (agreement No 7872-2017031711448150). In all cases, bipolar electrodes (tips 400 μm apart) were implanted bilaterally in the hippocampus (-2 mm anteroposterior, -1.5 mm mesio-lateral, -2 mm dorso-ventral from the Bregma) of C57B6j/Rj mice (80 days old). A reference electrode (monopolar) was placed at the cerebellum level. Mice were made epileptic following an intra-hippocampal injection of kainic acid that triggered an approx. 4-week epileptogenesis phase preceding the chronic epileptic phase. LFPs were sampled at 2048 Hz and hardware highpass filtered (0.16 Hz cutoff frequency). Recordings were made after the 4-week epileptogenesis period (chronic epilepsy stage). After protocol completion, mice were euthanized using the CO2 gradient method.

Electrophysiological recordings in rats (N=1) were obtained from the company Biotrial (http://www.biotrial.com). Ethics approval was obtained for these recordings, which involved injection of convulsive PTZ doses. A single bipolar electrode was implanted in the rat cortex, while a reference electrode was placed at the cerebellum level. A surgically implanted wireless transmitter enabled recording throughout the entire experiment without any physical manipulation of the rat, avoids movement artifacts. PTZ injection was performed through perfusion at the dose of 75 mg/kg.

In clinico recordings

In the context of pre-surgical planning, epileptic patients candidate to surgery undergo stereoencephalography (sEEG), involving the implantation of multiple intracranial electrodes aiming at improving the identification of epileptogenic zones. sEEG recordings from N=5 patients in the context of clinical evaluation were obtained from the Epileptology Unit of Rennes University Hospital (CHU Pontchaillou, Rennes, France). sEEG recordings were sampled at 2048 Hz with a Deltamed EEG acquisition system. Seizure epochs, involving several minutes pre- and post-seizures, were extracted for each patient to enable LFP reconstruction pre-, per- and post-seizure.

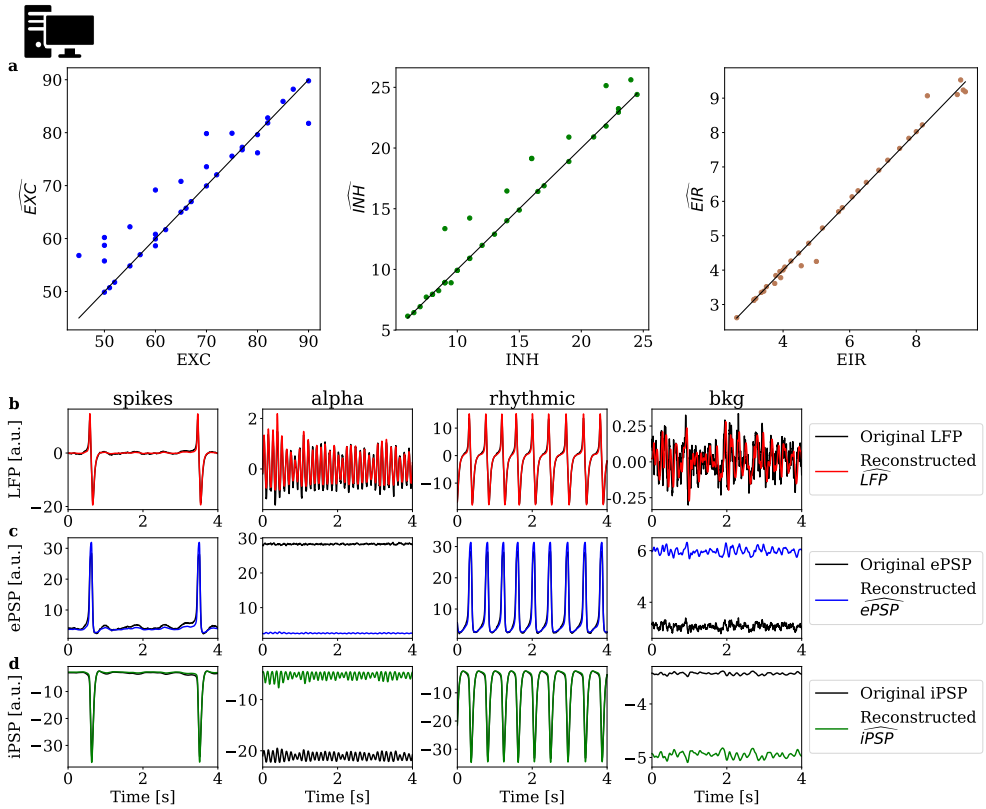


Fig 2. Results for simulated LFPs obtained using the same model as the reverse modeling approach. (**a**) Scatter plots for 35 simulated LFPs where ideal parameters vs identified parameters are displayed. Columns of combined panels b to e represent four types of activities. (**b**) LFP signals generated with the model (black) and $\widehat{\text{LFP}}$ signals reconstructed with the reverse modeling approach (red). Goodness-of-fit indexes are: γ=0.999 for spikes, γ=0.962 for alpha, γ=0.993 for rhythmic and γ=0.909 for background (bkg), (**c**) ePSP signals generated with the model (black) and $\widehat{\text{ePSP}}$ signals extracted from the reverse modeling approach (blue). (**d**) iPSP signals generated with the model (black) and $\widehat{\text{iPSP}}$ signals extracted from the reverse modeling approach (green).

Results

In silico validation

The reverse modeling method was first tested on simulated data. The NMM used in the reverse modeling algorithm used as an input the simulated LFPs generated by the same NMM. In this first validation step, a unique advantage is the ground-truth: estimated parameters (denoted by $\widehat{\text{ePSP}}$ and $\widehat{\text{iPSP}}$) can be directly compared to actual components (ePSP and iPSP) of the simulated LFP used as an input signal.

Fig. 2 presents the method performance on simulated data. Scatter plots in Fig. 2a display estimated parameter values ($\widehat{\text{EXC}}$, $\widehat{\text{INH}}$ and $\widehat{\text{EIR}}$: excitation to inhibition ratio) versus reference parameter values (denoted as EXC, INH and EIR) for a set of 35 simulated LFPs. The reverse modeling approach provides an excellent identification of parameters, since points in the scatter plot are close to the ideal result (represented by black lines when identified and reference parameters are equal). The method also

provides an excellent $\widehat{\text{EIR}}$ estimation, represented by the $\widehat{\text{EXC}}/\widehat{\text{INH}}$ ratio. Panels in Fig. 2b to 2d present 4 qualitatively different types of LFPs (sporadic spiking, alpha, ictal and background activity) generated from the model, and the corresponding reconstruction computed from the proposed approach. Panels in Fig. 2b display LFPs, while panels in Fig. 2c display ePSPs, and panels in Fig. 2d display iPSPs. As depicted in Fig. 2b, results revealed an excellent LFP reconstruction from estimated synaptic components, as shown by the goodness-of-fit indexes close to 1 (γ =0.999 for spikes, γ=0.962 for alpha, γ=0.993 for rhythmic and γ=0.909 for background). For background activity, the goodness-of-fit index was lower (γ=0.909) due to the strong contribution of the stochastic noise input (p(t)) for this particular type of activity. Reconstructed LFPs were obtained from estimated postsynaptic potentials, either excitatory ($\widehat{\text{ePSP}}$) or inhibitory ($\widehat{\text{iPSP}}$), respectively shown in panels in Fig. 2c and 2d. As depicted, these estimated components perfectly fit the reference components (ePSPs and iPSPs) that were summed up to obtain the simulated LFP signal used as an input for the reverse modeling approach. In the following section, we move from the validation on simulated data to an experimental validation of the proposed method in the specific context of epileptic activity. This choice was motivated by the fact that the major qualitative changes observed in LFPs (during the interictal to ictal transition for example) involve drastic changes in neuronal excitability, potentially quantified by the proposed approach. Typically, seizures are associated with a significant increase of the excitation to inhibition ratio [21].

In vivo and *in clinico* evaluation of the proposed approach

PTZ-induced seizure (rat)

In vivo results obtained in rats following PTZ (pentylenetetrazol) intra-peritoneal injection are reported in Fig. 3. Results indicated that the proposed method leads to a reconstructed $\widehat{\text{LFP}}$ signal (red line, Fig. 3a) that was in excellent agreement (normalized goodness-of-fit between 0.85 and 0.97) with the experimentally-recorded LFP (black line, Fig. 3a) for the entire recording duration (130 seconds), as depicted on the 4 selected epochs (Fig. 3a). This illustrates the robustness of our estimation method, which provides accurate estimates for epochs featuring qualitatively different dynamics. In addition to close replication of LFP dynamics, the estimation of the excitation/inhibition ratio ($\widehat{\text{EIR}}$) presented in Fig. 3b, was in agreement with the expected excitability changes during the seizure time course: the $\widehat{\text{EIR}}$ transiently increased at the onset of the PTZ-induced seizure, and then dramatically increased again during the seizure. It gradually decreased back to baseline as the seizure terminated. Interestingly, a decrease in the excitation level is observed during the seizure (Fig. 3b), which is however more modest than the drastic drop in inhibition, therefore still resulting in an increased EIR. Furthermore, since $\widehat{\text{PSP}}$s are computed in the NMM, they can be estimated through the proposed method as presented in Fig. 3c. The gradual decrease in inhibitory post-synaptic potentials associated with the seizure is clearly visible from epoch 1 to epoch 3, with a slight recovery in epoch 4 (Fig. 3c). Overall, there is satisfactory qualitative and quantitative agreement of LFP dynamics on the one hand, and of estimated $\widehat{\text{EIR}}$ and $\widehat{\text{PSP}}$s on the other hand. This supports the validity of our model-based reconstruction of excitatory and inhibitory components in data obtained *in vivo*. In addition, our method points at a result that is non-intuitive, namely the slight excitation decrease at the seizure onset.

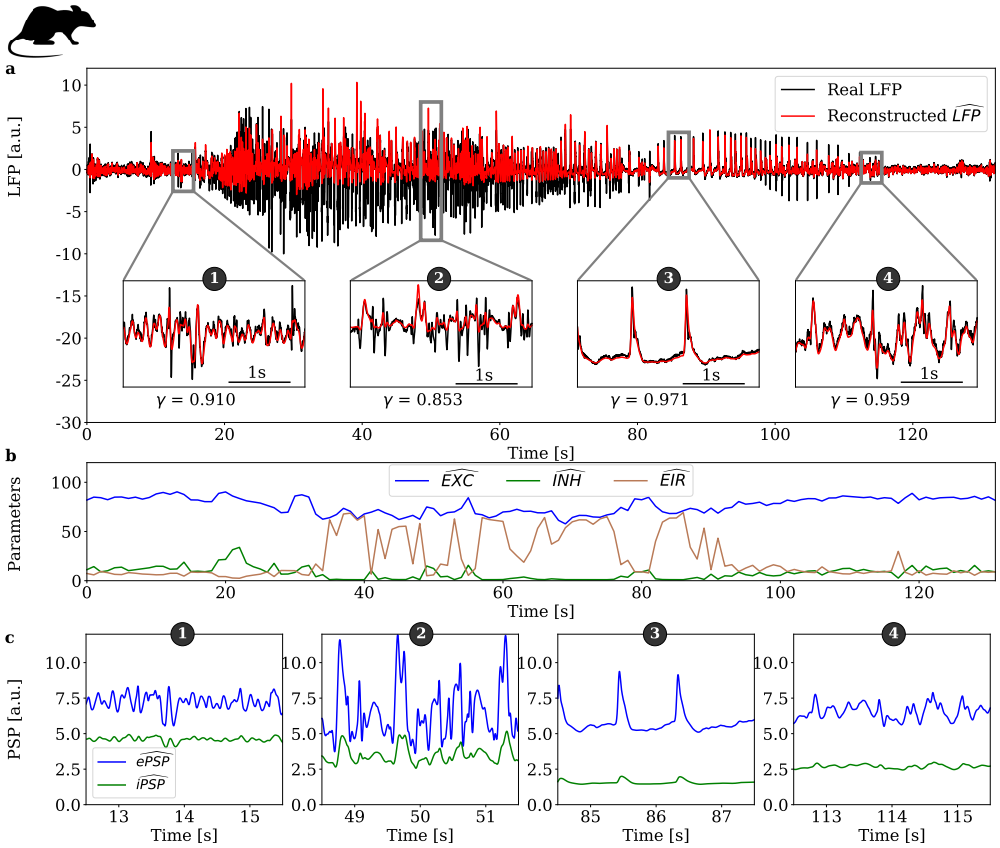


Fig 3. *In vivo* validation of the model-based LFP reconstruction in a PTZ-induced seizure (rat, N=1). (**a**) LFP signals recorded during a PTZ-induced seizure in rat (black line) and $\widehat{\text{LFP}}$ signals reconstructed using the reverse modeling approach (red line) for different types of activity (panels numbered from 1 to 4). The normalized goodness-of-fit index γ is presented for each activity type. (**b**) Identified model parameters: excitatory $\widehat{\text{EXC}}$ (blue line), inhibitory $\widehat{\text{INH}}$ (green line), and the excitability ratio $\widehat{\text{EIR}}$ (brown line). (**c**) Reconstructed $\widehat{\text{ePSP}}$ and $\widehat{\text{iPSP}}$ extracted from the NMM for each numbered panel presented in (**a**).

Interictal activity in the kainate model of epilepsy (mouse)

The proposed method was also tested in epileptic mice (kainate model, see Materials and Methods section). Based on our method applied to 53 epochs of interictal events (hippocampal paroxysmal discharges, HPD) in N=7 mice, a reconstruction of excitatory and inhibitory components along with the $\widehat{\text{EIR}}$ was performed. As depicted in Fig. 4a, the model-based reconstruction of the $\widehat{\text{LFP}}$ (red line) was in excellent agreement (goodness-of-fit index γ between =0.89 and 0.94) with the experimental LFP (black line) for the four considered epochs (see numbered panels in Fig. 4a). The reconstructed $\widehat{\text{LFP}}$ accurately matched the experimental LFP not only during background activity (panel 3 in Fig. 4a), but also during HPD (panel 4 in Fig. 4a). In addition, the time course of the estimated $\widehat{\text{EIR}}$ (Fig. 4b) was consistent with the occurrence of HPD within the signal: the $\widehat{\text{EIR}}$ remained low during background activity, transiently increases during HPD, before returning rapidly to background values (Fig. 4b). Interestingly, as can be seen in Fig. 4b, there was both a reduction of $\widehat{\text{EXC}}$ and $\widehat{\text{INH}}$

components (and not a decrease of $\widehat{\text{INH}}$ only as could be intuitively deducted in this case), which was however greater for $\widehat{\text{INH}}$, still resulting in a major $\widehat{\text{EIR}}$ increase. This is similar to the results obtained in the case of the PTZ-induced seizure (see Section). In addition, Fig. 4c presents reconstructed post-synaptic potentials ($\widehat{\text{ePSP}}$ and $\widehat{\text{iPSP}}$) corresponding to each of the four LFP epochs in Fig. 4a. The reconstructed iPSP points at a drastic inhibition decrease during a spike (epoch 2 in Fig. 4c) or HPD (epochs 1 and 4 in Fig. 4c), as compared to background (epoch 3 in Fig. 4c). The slight decrease in the reconstructed excitatory PSPs during interictal events (spikes or HPD) is also apparent (panels 1, 2 and 4 as compared to panel 3 in Fig. 4c). Fig. 4d presents results at the group level (N=53 epochs) regarding the excitation and inhibition levels during HPD as compared to background activity epochs. We observed a significant decreased of excitation and inhibition ($p < 0.001$) during HPD as compared to background activity, with a lesser excitation decrease as compared to the inhibition decrease, which still results in a drastic $\widehat{\text{EIR}}$ increase.

LFP recordings after injection of a sub-convulsive dose of PTZ (mouse)

In the third *in vivo* experiment, a non-epileptic mouse was injected with a sub-convulsive dose of PTZ, aiming at validating our model-based method when LFP changes are subtler (i.e., where no seizures are present). The LFP recorded over the entire experiment (approx. 2200 seconds long) is presented in Fig. 5a. Estimated $\widehat{\text{EXC}}$ and $\widehat{\text{INH}}$ components of the LFP are presented in Fig. 5b and 5c, and illustrate a significant increase ($p < 0.001$) of the $\widehat{\text{EIR}}$ (Fig. 5d) shortly after injection, as compared to pre-injection levels. Consistently with the PTZ-induced rat seizure, the $\widehat{\text{EIR}}$ increased (Fig. 5d) due to a joint decrease of $\widehat{\text{EXC}}$ and $\widehat{\text{INH}}$ components, which is more pronounced for the $\widehat{\text{INH}}$ component than $\widehat{\text{EXC}}$. Furthermore, the $\widehat{\text{EIR}}$ followed an exponential decay past the first 500 seconds after injection (the 500 s immediately after injection are discarded to avoid contamination by movements of the mouse due to manipulation by the experimenter), which is consistent with a pharmacokinetic response. The decay curve of the $\widehat{\text{EIR}}$ was fitted with an exponentially decaying curve, and the time constant optimizing the fit was estimated to be 370 seconds, in agreement with the PTZ pharmacokinetic time constant found in the literature [22, 23].

In clinico validation of the method in epileptic patients

Our method was validated using intracranial recordings from N=5 epileptic patients undergoing stereoencephalography (sEEG) in the context of pre-surgical evaluation. Among the available recordings, we extracted a total of 47 spontaneous seizures, which featured artifact-free epochs pre- and post-seizures, so that the $\widehat{\text{EIR}}$ evolution could be tracked over the entire seizure time course. Fig. 6 presents an example of human seizure and associated $\widehat{\text{LFP}}$ reconstruction (Fig. 6a) along with $\widehat{\text{EIR}}$ estimation (Fig. 6b), and reconstruction of PSPs (Fig. 6c). $\widehat{\text{LFP}}$ reconstruction is illustrated in Fig. 6a, with the four zoomed panels emphasizing qualitatively different portions of the signal, illustrating the excellent agreement between the recorded intracranial signal (black line) and the reconstructed $\widehat{\text{LFP}}$ (red line), as quantified by the goodness-of-fit index ranging between 0.92 and 0.98. In the seizure presented as an example in Fig. 6a, the estimated $\widehat{\text{EIR}}$ over the seizure time course (see Fig. 6b) was physiologically plausible, with a drastic $\widehat{\text{EIR}}$ increase during the seizure itself, with a return to pre-ictal values after seizure cessation. The reconstructed post-synaptic components (Fig. 6c) for each of the four panels presented in Fig. 6a also pointed at a gradual decrease in the inhibition level as the seizure progresses, consistently with the *in vivo* PTZ-induced seizure studied (Fig. 3). Fig. 6d presents the group results (N=5) and compares reconstructed

excitation and inhibition levels during interictal and ictal epochs. The estimated $\widehat{\text{EIR}}$ increase during seizure was driven by a significant $\widehat{\text{EXC}}$ increase ($p < 0.001$), combined with a significant $\widehat{\text{INH}}$ decrease ($p < 0.001$). Reconstructed $\widehat{\text{LFP}}$s for this database matched closely experimental LFPs, with an average goodness-of-fit γ=0.893 $\pm$0.56.

Discussion

We developed a novel model-based method enabling the plausible reconstruction of excitatory and inhibitory post-synaptic currents involved in LFP generation as measured experimentally by intracranial electrodes. This method was validated at several levels: from in silico (using the method on LFPs generated using a NMM, providing a ground truth), to *in vivo* (rat and mouse LFPs), and finally to *in clinico* (sEEG recordings in epileptic patients). In all model-based LFP reconstructions based on experimental data, we found plausible and relevant results emphasizing an $\widehat{\text{EIR}}$ increase during the generation of interictal events and during seizures. Furthermore, the reconstructed $\widehat{\text{LFP}}$, using our model-based decomposition, matched accurately experimentally recorded LFPs, suggesting that the assumption that the LFP is mostly generated by excitatory and slow inhibitory post-synaptic currents is reasonable. In addition, using a subconvulsive dose of PTZ in mice, we were able to provide an estimate of the PTZ time constant consistent with the literature. Therefore, a major advantage of our technique is that it enables the detection of subtle variations in neuronal excitability; such as those induce by sub-convulsive doses of PTZ. In addition, it is worth noting that the proposed method consistently pointed at a counter-intuitive decrease of excitation (minor as compared to the decrease of inhibition) in all *in vivo* recordings (PTZ-induced seizure, epileptiform events in the kainate model). However, we did not identify a similar decrease of excitation during human seizures from our available database. A possible explanation is that slightly different pathophysiological mechanisms are involved in the generation of epileptiform activity in our *in vivo* models as compared to human seizures.

The possibility to reconstruct, from an experimentally recorded LFP, the time course of excitatory (glutamatergic) and inhibitory (slow GABAergic) post-synaptic contributions, is especially appealing. Our method provides indeed full access to fundamental variables such as ePSPs and iPSPs for each subpopulation. These signals are challenging to extract from LFPs, since the problem is underdetermined. The identification of NMM parameters to optimize the fit of simulated LFPs with experimental LFPs has been addressed in several studies, using methods based on genetic algorithms [24], Kalman filter [25], synaptic conductance estimation [26], or Hodgkin-Huxley parameters estimation [27]. However, none of these methods provides access to PSP signals from experimentally recorded LFPs. Our method should have applications in numerous areas of neuroscience, since LFP is a very common electrophysiological recording modality, from *in vivo* to *in clinico*, and ranging from neurophysiology to diagnostic applications. In addition, the method is sensitive enough to capture subtle changes in excitability, as shown in the case of the sub-convulsive PTZ dose experiment, and was able to estimate successfully PTZ pharmacokinetics.

One limitation is that our method currently features only slow, dendrite-targeting GABAergic synapses, and neglects fast soma-targeting synapses; which limits the possible frequency range of LFPs that can be reconstructed. This simplification was done for three reasons. First, we aimed at providing a proof-of-concept that a method based on a physiologically-based NMM can provide meaningful excitatory and inhibitory components from an experimentally recorded LFP. Second, due to its very generic structure, the NMM used should be general enough to analyze LFPs from any brain region. Third, if a sub-population of fast GABAergic neurons was added in the

model, that would considerably increase the time required for parameter identification, since that would involve a 3D parameter space instead of 2D. This absence of fast GABAergic activity explains why the epochs where higher frequencies (beta band -13 to 30 Hz- and above) were present in the LFP are more challenging to reconstruct using our approach. This specific point explains why the goodness-of-fit is better in the case of HPD in mice (Fig. 4) as compared to epochs of relatively fast activity during the rat PTZ-induced seizure (Fig. 3): HPD are mainly characterized by slow activity, as compared to the relatively fast activity occurring at the onset of seizures. Therefore, agreement between reconstructed and experimentally measured LFPs is lower when higher frequencies (beta range and above, i.e. > 13 Hz) are present in the signal.

Using our model-guided approach to unveil the time course of excitatory and inhibitory post-synaptic currents is promising, especially since the algorithm provides excitability tracking with a performance in real-time (S1 video available as Supporting Information): on average, 2 s of LFP sampled at 1024 Hz requires 0.7 s using a single core on a PC equipped with an Intel Xeon E5-2637 3.5 GHz CPU and 64 GB of 1866 MHz RAM. Furthermore, there is still ample room for computation time optimization, since the algorithm is currently programmed in Python, which is known for not being optimal in terms of computation time. The possibility to automatically track neuronal excitability in real-time could be implemented on-chip, for example, in a neuromodulation device, to trigger stimulation when excitability exceeds a certain threshold. Another immediate application would be the monitoring of epileptic patients undergoing sEEG for pre-surgical evaluation.

Supporting information

S1 Fig. Activity maps of the original vs. updated neural mass model. (a) Activity maps for the classical Jansen-Rit model where the four possible regimes are colored. (b) Activity maps for the new neural mass model where the same four possible regimes than the Jansen-Rit model are colored. (c) An example of the four possible regimes for the Jansen-Rit model (to the left) and their frequency spectrum (to the right). Background: EXC = 3.75 and INH = 39; spike: EXC = 4.2 and INH = 30.5; Alpha: EXC = 4.4 and INH = 10.5; Rhythmic: EXC = 4.3 and INH = 24.5. (d) An example of the four possible regimes for the updated neural mass model (to the left) and their frequency spectrum (to the right). Background: EXC = 60 and INH = 20; spike: EXC = 60 and INH = 13.9; Alpha: EXC = 60 and INH = 6.7; Rhythmic: EXC = 60 and INH = 13.

S1 Video. Software demonstration.

Author Contributions

Conceived and designed the experiments: P.B., F.W. Computational model: M.Y., J.M. F.W. Performed the experimental study: P.B., J.M. Manuscript preparation: M.Y., P.B., J.M., F.W.

References

1. Buzsáki G. Large-scale recording of neuronal ensembles. Nature neuroscience. 2004;7(5):446.

2. Destexhe A, Bedard C. Local field potential. Scholarpedia. 2013;8(8):10713.

3. Priori A, Foffani G, Rossi L, Marceglia S. Adaptive deep brain stimulation (aDBS) controlled by local field potential oscillations. Experimental neurology. 2013;245:77–86.

4. Buzsáki G, Anastassiou CA, Koch C. The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. Nature reviews neuroscience. 2012;13(6):407.

5. Malmivuo P, Malmivuo J, Plonsey R. Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields. Oxford University Press, USA; 1995.

6. Ramachandran VS. Encyclopedia of the Human Brain, Four-Volume Set. Academic Press; 2002.

7. Harris KD, Csicsvari J, Hirase H, Dragoi G, Buzsáki G. Organization of cell assemblies in the hippocampus. Nature. 2003;424(6948):552.

8. Bédard C, Kröger H, Destexhe A. Modeling extracellular field potentials and the frequency-filtering properties of extracellular space. Biophysical journal. 2004;86(3):1829–1842.

9. Reid CA, Leaw B, Richards KL, Richardson R, Wimmer V, Yu C, et al. Reduced dendritic arborization and hyperexcitability of pyramidal neurons in a Scn1b-based model of Dravet syndrome. Brain. 2014;137(6):1701–1715.

10. Chen C, Bazan NG. Endogenous PGE2 regulates membrane excitability and synaptic transmission in hippocampal CA1 pyramidal neurons. Journal of neurophysiology. 2005;93(2):929–941.

11. Malik R, Chattarji S. Enhanced intrinsic excitability and EPSP-spike coupling accompany enriched environment-induced facilitation of LTP in hippocampal CA1 pyramidal neurons. Journal of neurophysiology. 2011;107(5):1366–1378.

12. Wang HY, Hsieh PF, Huang DF, Chin PS, Chou CH, Tung CC, et al. RBFOX3/NeuN is required for hippocampal circuit balance and function. Scientific reports. 2015;5:17383.

13. Gibson JR, Bartley AF, Huber KM. Role for the Subthreshold Currents I Leak and IH in the Homeostatic Control of Excitability in Neocortical Somatostatin-Positive Inhibitory Neurons. Journal of neurophysiology. 2006;96(1):420–432.

14. Ferguson KA, Huh CY, Amilhon B, Williams S, Skinner FK. Experimentally constrained CA1 fast-firing parvalbumin-positive interneuron network models exhibit sharp transitions into coherent high frequency rhythms. Frontiers in computational neuroscience. 2013;7:144.

15. Fidzinski P, Korotkova T, Heidenreich M, Maier N, Schuetze S, Kobler O, et al. KCNQ5 K+ channels control hippocampal synaptic inhibition and fast network oscillations. Nature communications. 2015;6:6254.

16. Erisir A, Lau D, Rudy B, Leonard C. Function of specific K+ channels in sustained high-frequency firing of fast-spiking neocortical interneurons. Journal of neurophysiology. 1999;82(5):2476–2489.

17. Jansen BH, Rit VG. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biological cybernetics. 1995;73(4):357–366.

18. Touboul J, Wendling F, Chauvel P, Faugeras O. Neural mass activity, bifurcations, and epilepsy. Neural computation. 2011;23(12):3232–3286.

19. Destexhe A, Paré D. Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo. Journal of neurophysiology. 1999;81(4):1531–1547.

20. Nakhmani A, Tannenbaum A. A new distance measure based on generalized image normalized cross-correlation for robust video tracking and image recognition. Pattern recognition letters. 2013;34(3):315–321.

21. Fritschy JM. Epilepsy, E/I balance and GABAA receptor plasticity. Frontiers in molecular neuroscience. 2008;1:5.

22. Wendling F, Gerber U, Cosandier-Rimele D, Nica A, De Montigny J, Raineteau O, et al. Brain (Hyper) Excitability Revealed by Optimal Electrical Stimulation of GABAergic Interneurons. Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation. 2016;9(6):919–932.

23. Mandhane SN, Aavula K, Rajamannar T. Timed pentylenetetrazol infusion test: a comparative analysis with sc PTZ and MES models of anticonvulsant screening in mice. Seizure-European Journal of Epilepsy. 2007;16(7):636–644.

24. Wendling F, Hernandez A, Bellanger JJ, Chauvel P, Bartolomei F. Interictal to ictal transition in human temporal lobe epilepsy: insights from a computational model of intracerebral EEG. Journal of Clinical Neurophysiology. 2005;22(5):343.

25. Bellanger J, Hernandez A, Wendling F, Frogerais P. Mesure de l'excitation et de l'inhibition dans le tissu neuronal en épilepsie par identification d'un modèle dynamique non linéaire d'activité EEG. In: 20ème Colloque sur le traitement du signal et des images, FRA, 2005. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images; 2005.

26. Rudolph M, Piwkowska Z, Badoual M, Bal T, Destexhe A. A method to estimate synaptic conductances from membrane potential fluctuations. Journal of neurophysiology. 2004;91(6):2884–2896.

27. Zheng Y, Luo JJ, Harris S, Kennerley A, Berwick J, Billings SA, et al. Balanced excitation and inhibition: Model based analysis of local field potentials. NeuroImage. 2012;63(1):81–94.

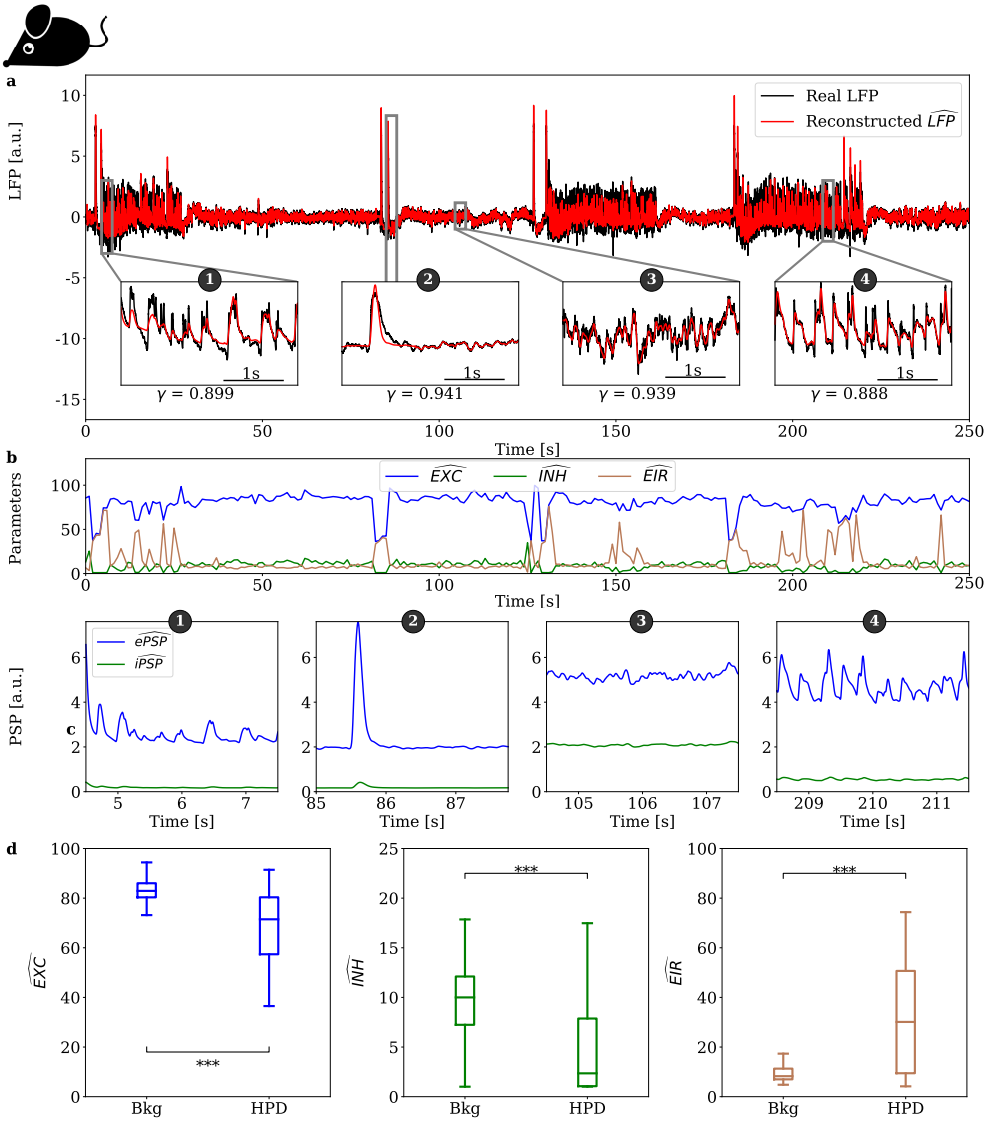


Fig 4. *In vivo* validation of the model-based LFP reconstruction in kainate mice (N=7). (**a**) LFP signals recorded in mouse where HPD are present (black line) and $\widehat{\text{LFP}}$ signals reconstructed with the reverse modeling approach (red line) for different types of activity (panels numbered from 1 to 4). The normalized goodness-of-fit index γ is presented for each activity type. (**b**) Identified model parameters: excitatory $\widehat{\text{EXC}}$ (blue line), inhibitory $\widehat{\text{INH}}$ (green line), and the excitability ratio $\widehat{\text{EIR}}$ (brown line). (**c**) Reconstructed $\widehat{\text{ePSP}}$ and $\widehat{\text{iPSP}}$ extracted from the NMM for each numbered panel presented in (a). (**d**) Boxplot of $\widehat{\text{EXC}}$, $\widehat{\text{INH}}$, and $\widehat{\text{EIR}}$ over background activity vs. HPD segments of 53 LFPs ($\approx$ 500s long) in seven mice. $***$ stands for p<0.001.

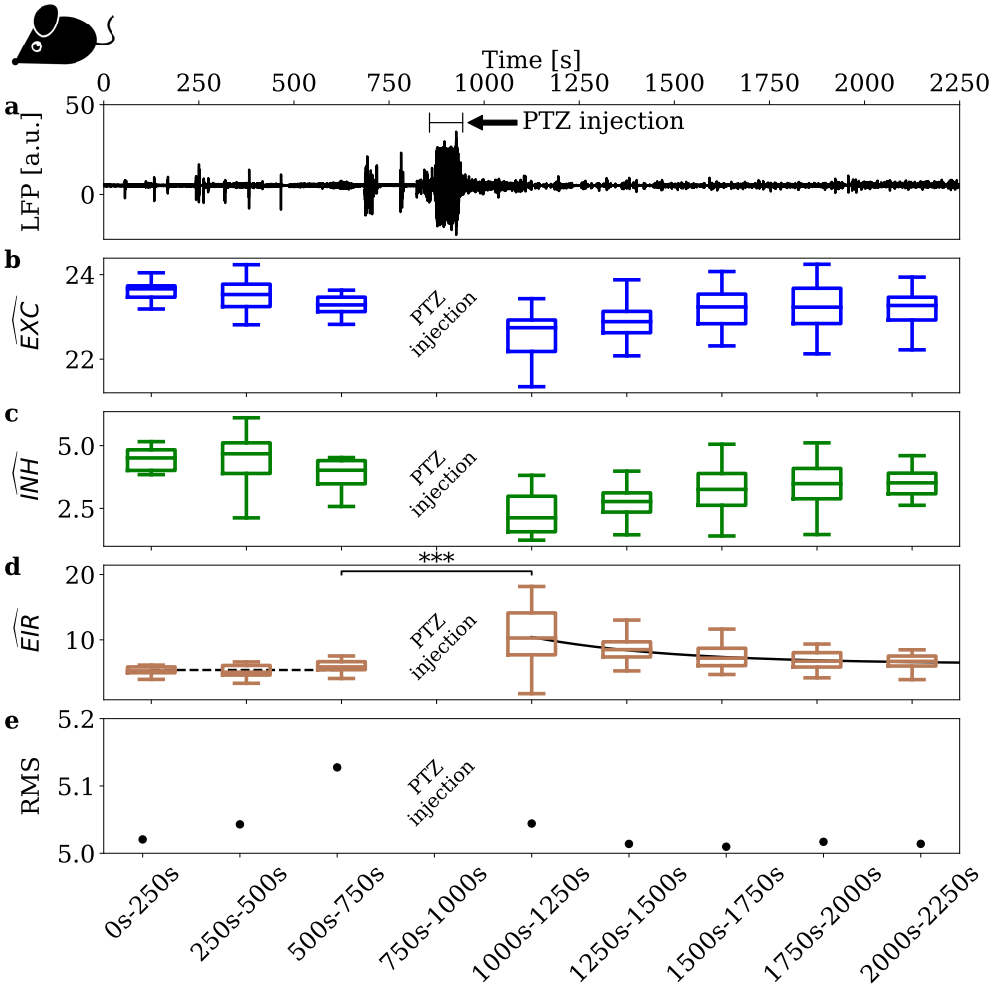


Fig 5. *In vivo* estimation of the excitation to inhibition ratio ($\widehat{\text{EIR}}$) after injection of a subconvulsive dose of PTZ (mouse, N=1). (**a**) LFP recording. (**b**) Estimation of the $\widehat{\text{EXC}}$ parameter. (**c**) Estimation of the $\widehat{\text{INH}}$ parameter. (**d**) $\widehat{\text{EIR}}$. Dashed line corresponds to the average of three first boxplots. The continuous line corresponds to an exponential decay model ($ae^{\frac{t-1000}{b}} + c$, with a=4.1, b=-370 and c=6.3) on the five last boxplots. (**e**) RMS values computed for each time window. $***$ stands for p<0.001.

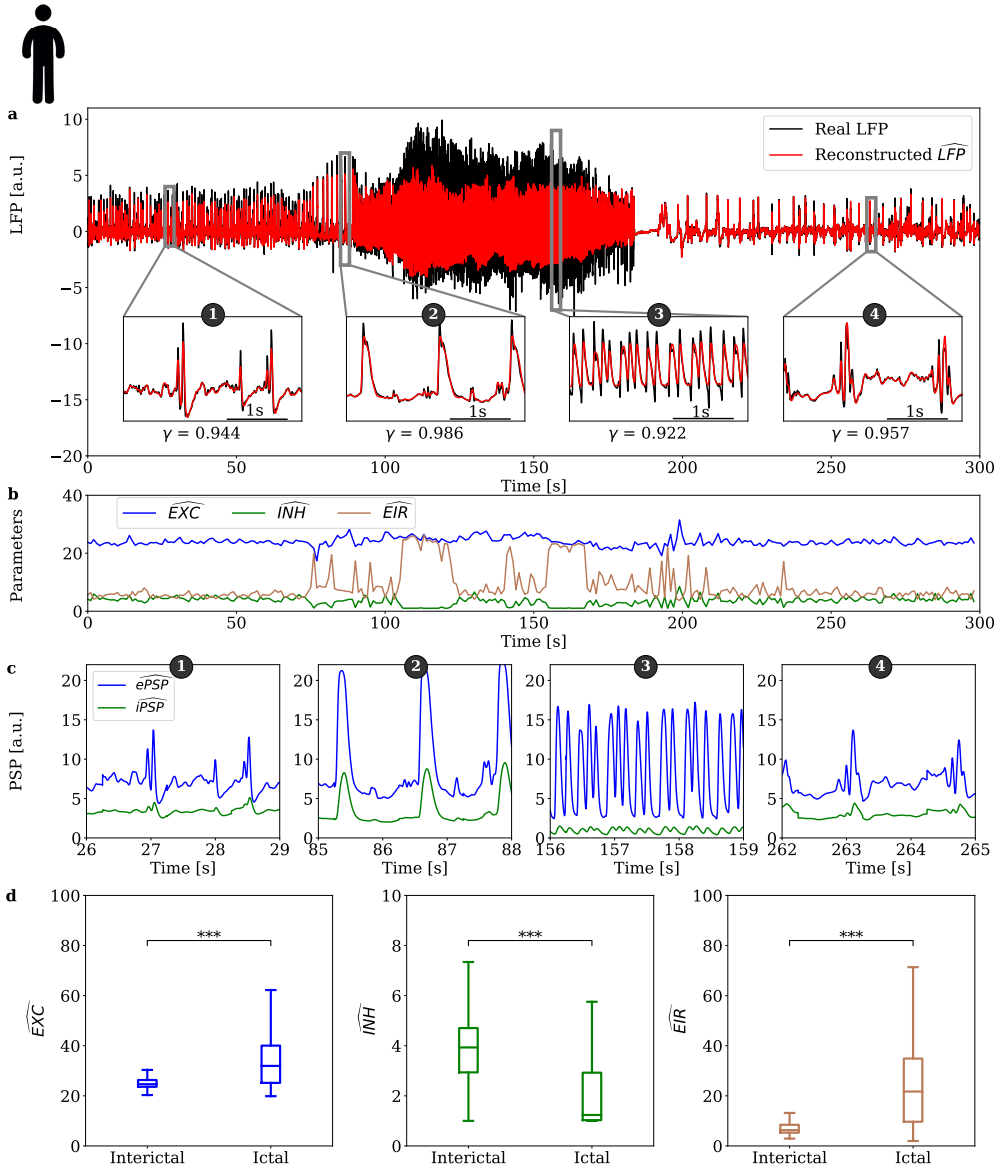
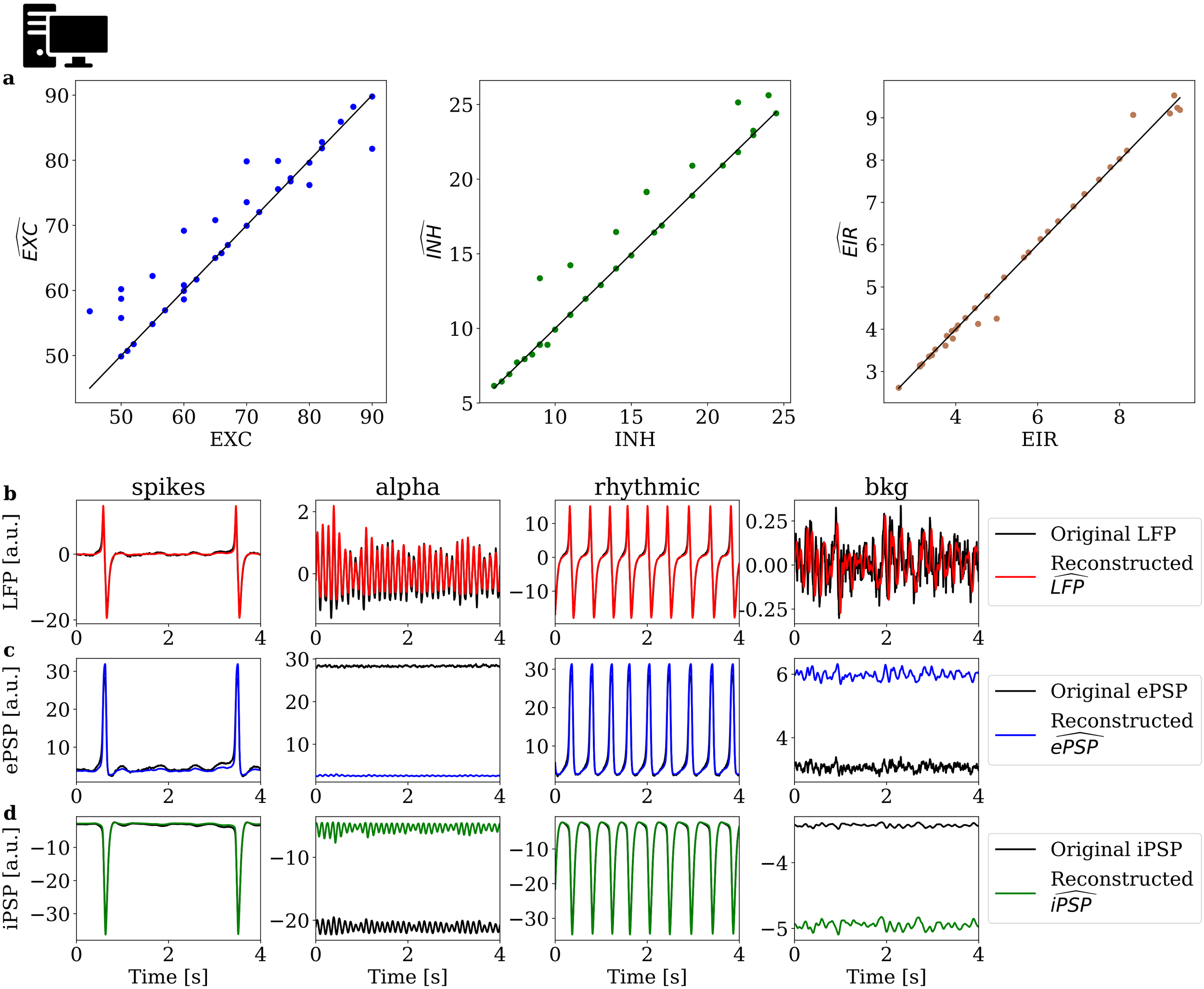
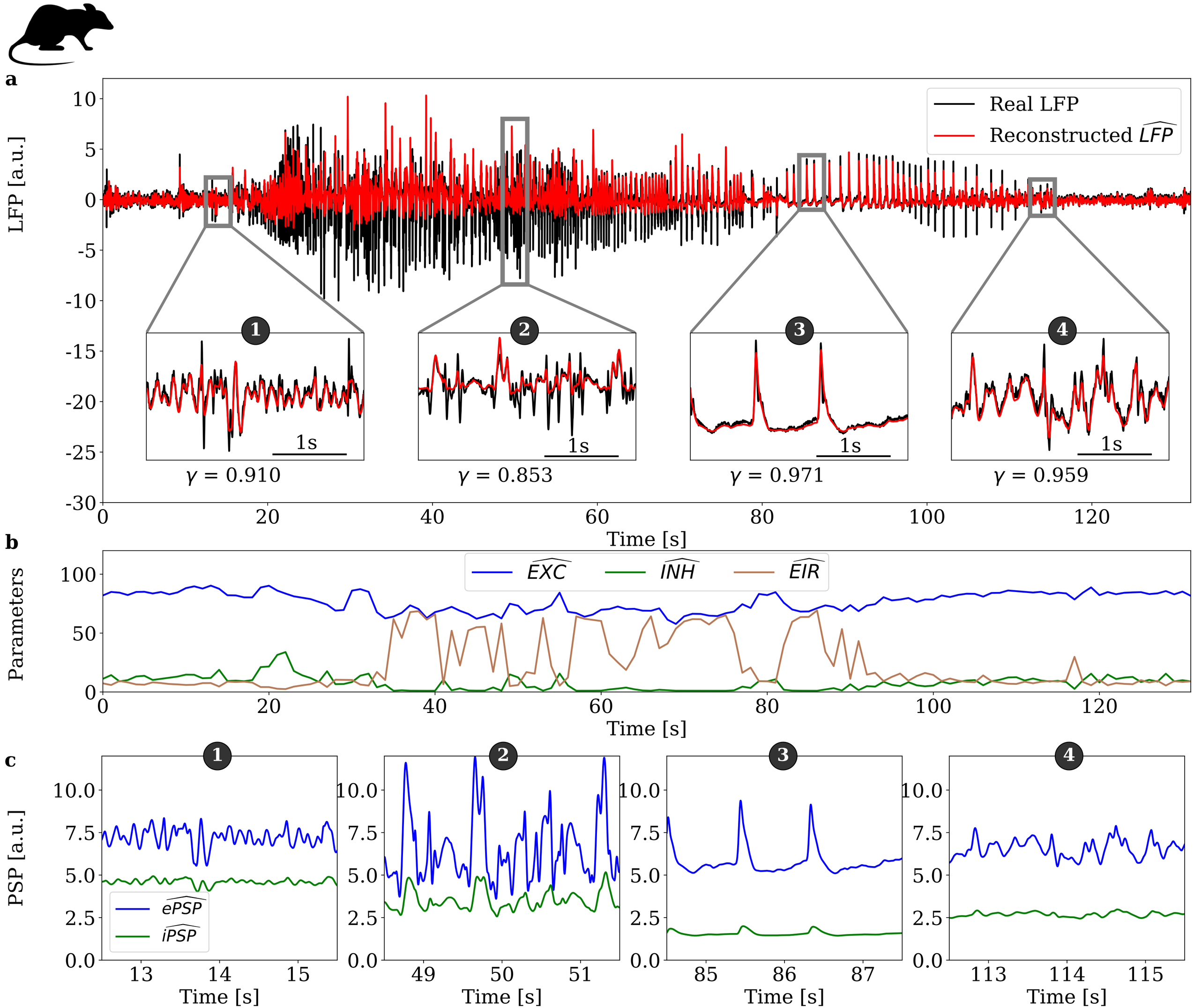
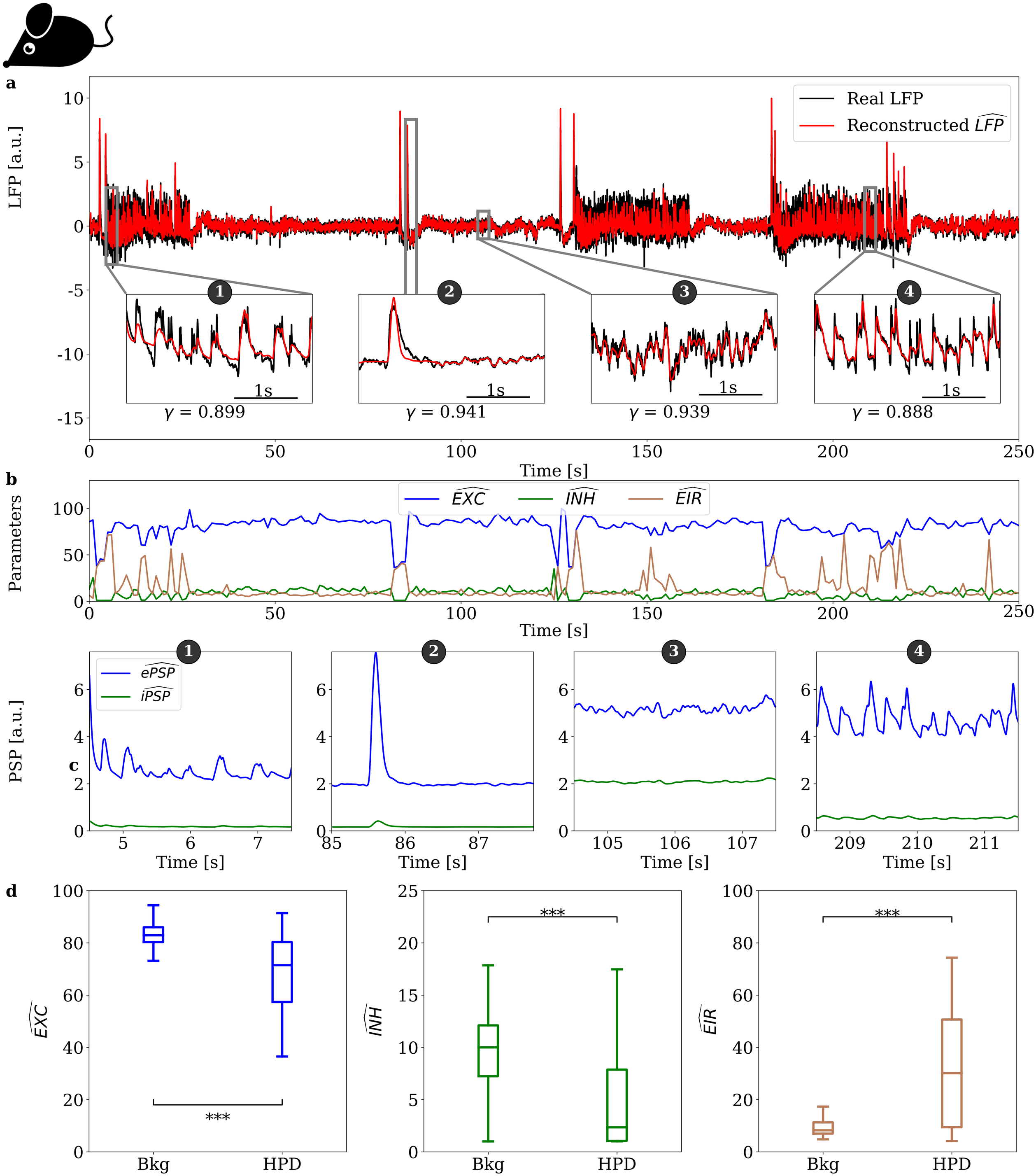
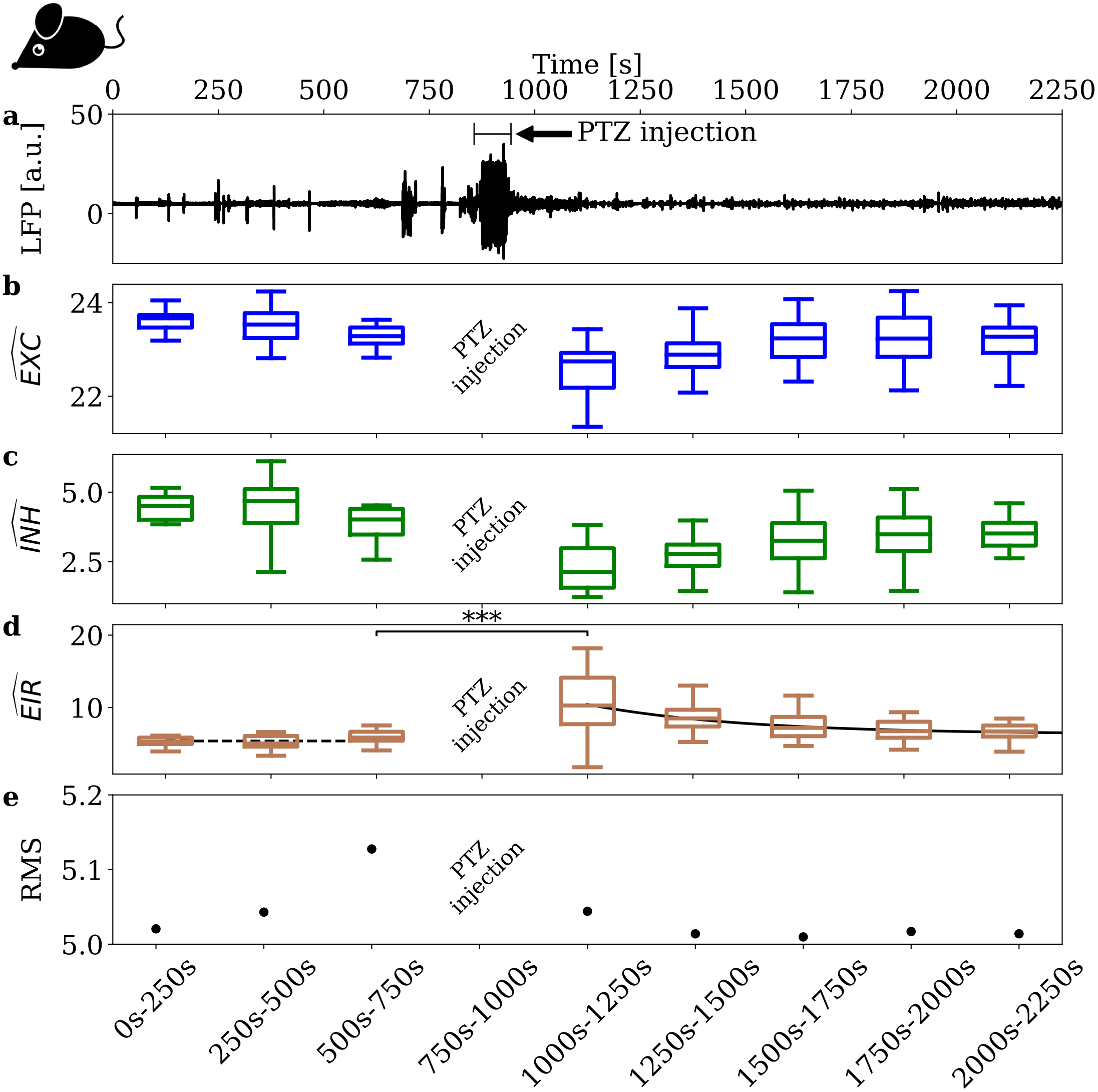


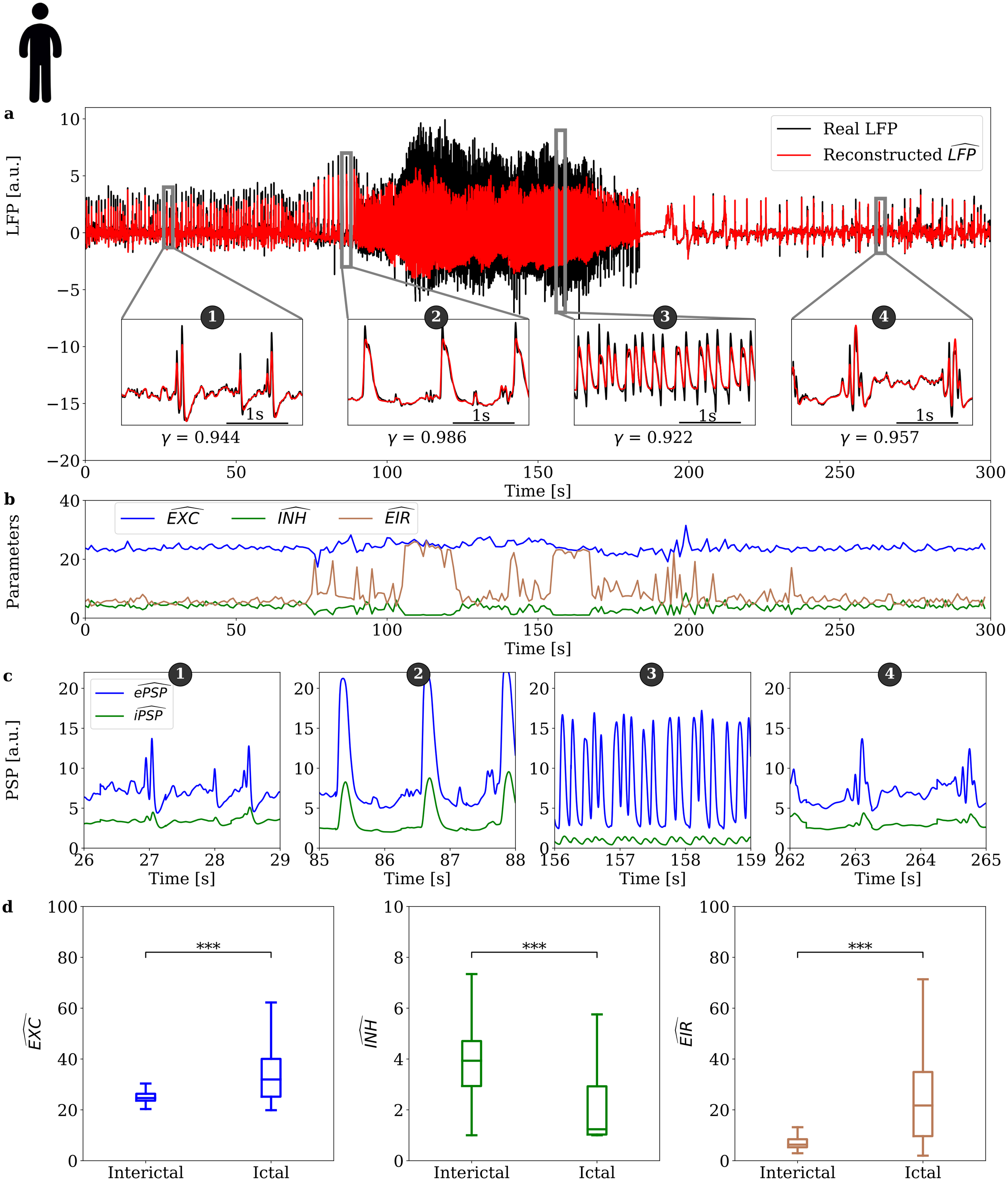
Fig 6. Validation on human seizures data (epileptic patients, N=5) obtained through intracranial recordings. (**a**) LFP signals recorded in human where a seizure is present (black line) and $\widehat{\text{LFP}}$ signals reconstructed with the reverse modeling approach (red line) for different types of activity (panels numbered from 1 to 4). The normalized goodness-of-fit index γ is presented for each activity type. (**b**) Identified model parameters: excitatory $\widehat{\text{EXC}}$ (blue line), inhibitory $\widehat{\text{INH}}$ (green line), and the excitability ratio $\widehat{\text{EIR}}$ (brown line). (**c**) Reconstructed $\widehat{\text{ePSP}}$ and $\widehat{\text{iPSP}}$ extracted from the NMM for each numbered panel presented in (a). (**d**) Boxplot of $\widehat{\text{EXC}}$, $\widehat{\text{INH}}$, and $\widehat{\text{EIR}}$ ratio over interictal vs ictal segments of 47 LFPs in five human patients. $***$ stands for $p<0.001$.

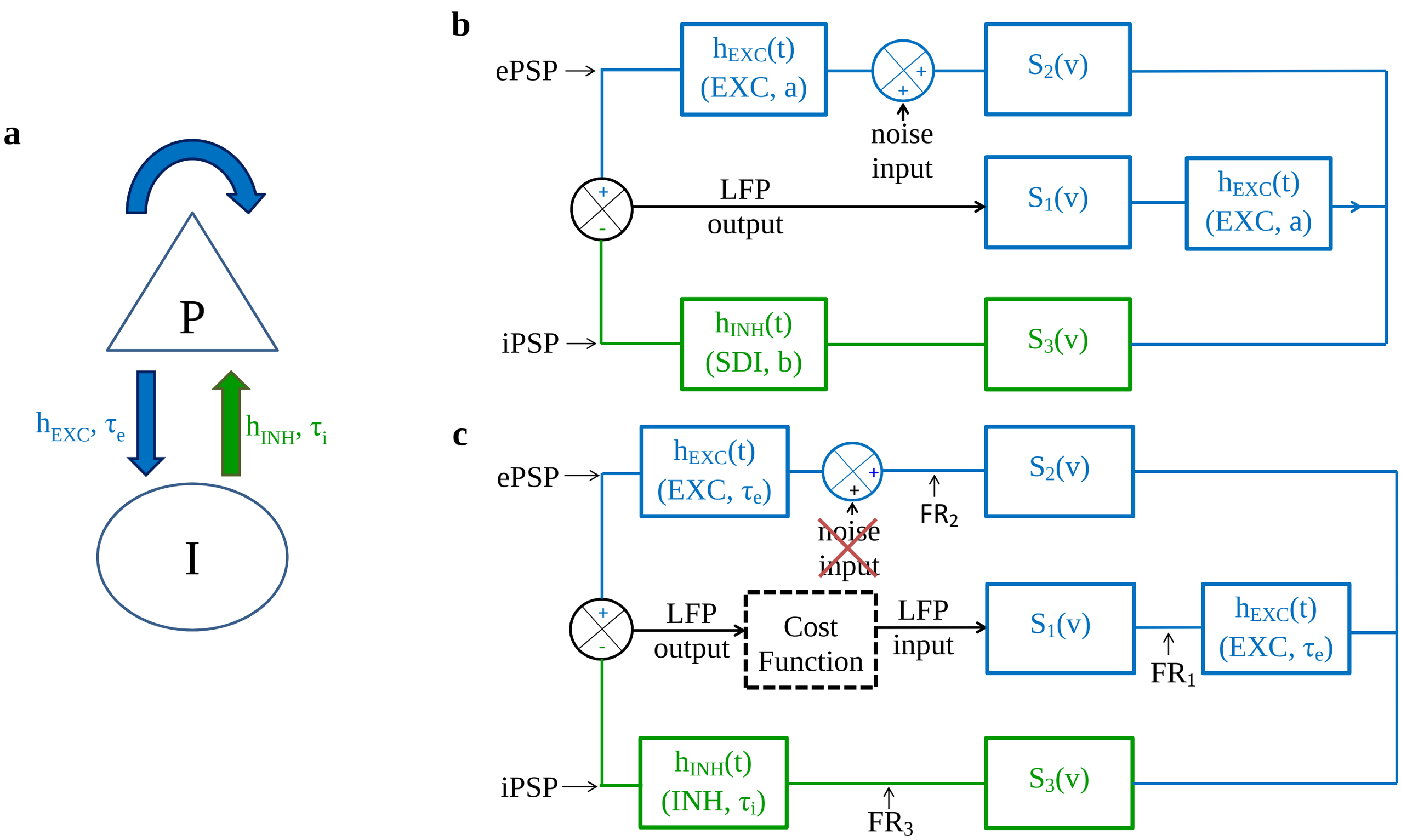

a
b
c
d
EXC
INH
EIR
spikes
alpha
rhythmic
bkg
LFP [a.u.]
ePSP [a.u.]
iPSP [a.u.]
Time [s]
Original LFP
Reconstructed
Original ePSP
Reconstructed
Original iPSP
Reconstructed


a
LFP [a.u.]
Real LFP
Reconstructed $\widehat{LFP}$
1
2
3
4
1s
γ = 0.910
γ = 0.853
γ = 0.971
γ = 0.959
Time [s]
b
Parameters
$\widehat{EXC}$
$\widehat{INH}$
$\widehat{EIR}$
Time [s]
c
PSP [a.u.]
$\widehat{ePSP}$
$\widehat{iPSP}$
Time [s]


a
LFP [a.u.]
Real LFP
Reconstructed $\widehat{LFP}$
1
2
3
4
1s
γ = 0.899
γ = 0.941
γ = 0.939
γ = 0.888
Time [s]
b
Parameters
$\widehat{EXC}$
$\widehat{INH}$
$\widehat{EIR}$
Time [s]
c
PSP [a.u.]
$\widehat{ePSP}$
$\widehat{iPSP}$
Time [s]
d
$\widehat{EXC}$
$\widehat{INH}$
$\widehat{EIR}$
***
Bkg
HPD


Time [s]
0
250
500
750
1000
1250
1500
1750
2000
2250
a
LFP [a.u.]
50
0
PTZ injection
b
$\widehat{EXC}$
24
22
PTZ injection
c
$\widehat{INH}$
5.0
2.5
PTZ injection
d
$\widehat{EIR}$
20
10
***
PTZ injection
e
RMS
5.2
5.1
5.0
PTZ injection
0s-250s
250s-500s
500s-750s
750s-1000s
1000s-1250s
1250s-1500s
1500s-1750s
1750s-2000s
2000s-2250s


a
LFP [a.u.]
Real LFP
Reconstructed $\widehat{LFP}$
1
2
3
4
1s
γ = 0.944
γ = 0.986
γ = 0.922
γ = 0.957
Time [s]
b
Parameters
$\widehat{EXC}$
$\widehat{INH}$
$\widehat{EIR}$
Time [s]
c
PSP [a.u.]
$\widehat{ePSP}$
$\widehat{iPSP}$
Time [s]
d
$\langle EXC \rangle$
$\langle INH \rangle$
$\langle EIR \rangle$
***
Interictal
Ictal


a
P
I
h_{EXC}, τ_e
h_{INH}, τ_i
b
ePSP
iPSP
$h_{EXC}(t)$ (EXC, a)
$h_{INH}(t)$ (SDI, b)
noise input
LFP output
$S_2(v)$
$S_1(v)$
$S_3(v)$
$h_{EXC}(t)$ (EXC, a)
c
ePSP
iPSP
$h_{EXC}(t)$ (EXC, τ_e)
$h_{INH}(t)$ (INH, τ_i)
noise input
FR_2
LFP output
Cost Function
LFP input
FR_1
FR_3
$S_2(v)$
$S_1(v)$
$S_3(v)$
$h_{EXC}(t)$ (EXC, τ_e)