



Identifiability of a Binomial Synapse

Camille Gontier^{1*} and Jean-Pascal Pfister^{1,2}

¹ Department of Physiology, University of Bern, Bern, Switzerland, ² Institute of Neuroinformatics and Neuroscience Center Zurich, University of Zurich/ETH Zurich, Zurich, Switzerland

Synapses are highly stochastic transmission units. A classical model describing this stochastic transmission is called the binomial model, and its underlying parameters can be estimated from postsynaptic responses to evoked stimuli. The accuracy of parameter estimates obtained via such a model-based approach depends on the identifiability of the model. A model is said to be structurally identifiable if its parameters can be uniquely inferred from the distribution of its outputs. However, this theoretical property does not necessarily imply practical identifiability. For instance, if the number of observations is low or if the recording noise is high, the model's parameters can only be loosely estimated. Structural identifiability, which is an intrinsic property of a model, has been widely characterized; but practical identifiability, which is a property of both the model and the experimental protocol, is usually only qualitatively assessed. Here, we propose a formal definition for the practical identifiability domain of a statistical model. For a given experimental protocol, this domain corresponds to the set of parameters for which the model is correctly identified as the ground truth compared to a simpler alternative model. Considering a model selection problem instead of a parameter inference problem allows to derive a non-arbitrary criterion for practical identifiability. We apply our definition to the study of neurotransmitter release at a chemical synapse. Our contribution to the analysis of synaptic stochasticity is three-fold: firstly, we propose a quantitative criterion for the practical identifiability of a statistical model, and compute the identifiability domains of different variants of the binomial release model (uni or multi-quantal, with or without short-term plasticity); secondly, we extend the Bayesian Information Criterion (BIC), a classically used tool for model selection, to models with correlated data (which is the case for most models of chemical synapses); finally, we show that our approach allows to perform data free model selection, i.e., to verify if a model used to fit data was indeed identifiable even without access to the data, but having only access to the fitted parameters.

Keywords: model selection, practical identifiability, structural identifiability, binomial, synapse

OPEN ACCESS

Edited by:

Renaud Blaise Jolivet,
Université de Genève, Switzerland

Reviewed by:

Ryota Kobayashi,
University of Tokyo, Japan
Lubomir Kostal,
Institute of Physiology CAS, Czechia
Suhita Nadkarni,
Indian Institute of Science Education
and Research, India

*Correspondence:

Camille Gontier
camille.gontier@unibe.ch

Received: 02 May 2020

Accepted: 18 August 2020

Published: 30 September 2020

Citation:

Gontier C and Pfister J-P (2020)
Identifiability of a Binomial Synapse.
Front. Comput. Neurosci. 14:558477.
doi: 10.3389/fncom.2020.558477

1. INTRODUCTION

Model selection is highly relevant to neuroscience, as neurons, dendrites, and synapses can be represented by models with different levels of complexity and abstraction. When it comes to fitting recorded data, predicting the output of a system to a given stimulus, or making sense of an observed phenomenon, several possible models can be used: this raises the question of what makes a good model. Finding the correct model is a crucial issue in studying the brain.

Firstly, a good model needs to be sufficiently complex to account for observed data, while being simple enough to generalize to future observations. Competing models are typically compared based on their ability to fit an observed data set, while being penalized for their complexity (or number of free parameters) to avoid overfitting. Different model selection tools (Bayesian Information Criterion, Akaike Information Criterion,...) are classically used to determine which model is the best one to fit a data set (Daw et al., 2011).

Secondly, models also differ in their nature, and can be classified as phenomenological, normative, or biophysical. On the one hand, purely phenomenological models are useful for relating the output of a system to its input, and can provide a computationally efficient way to make prediction. However, as they are solely based on the empirical relation between the input and the output of the system, and not on its inner biological principles, they lack interpretability. On the other hand, normative and biophysical models can be computationally challenging to fit on data, but are more realistic. In a normative approach, the output of a system is computed from an objective function which models its high-level functions and principles. As opposed to this top-down approach, biophysical models aim at precisely describing the low-level biological components of the system. An interesting property of these biophysical models is that their parameters correspond to real physical quantities: when the parameters of a system cannot be measured directly, they can be estimated by fitting a corresponding biophysical model on recorded output data of the system, a procedure known as model-based inference. By computing the likelihood of the data as a function of the parameters, it is possible to follow a maximum-likelihood approach to obtain a point estimate of the parameters (Barri et al., 2016), or to compute the full posterior distribution over them (Bird et al., 2016).

Such a parameter inference requires that the model used is identifiable. Structural (i.e., model-based) identifiability is a property of the model, regardless of experimental results. In a structurally identifiable system, the dimension of the output is sufficiently high with respect to the dimension of the parameters vector to uniquely define it: the parameters can be non-ambiguously inferred if the complete distribution of the output is known. Structural identifiability has been widely studied in many fields of physics and biology (Raue et al., 2009, 2011; Komorowski et al., 2011; Koyama, 2012; Hines et al., 2014), and different criteria exist to assess the structural identifiability of a model (Massonis and Villaverde, 2020).

This theoretical property is not equivalent to practical (i.e., experiment-based) identifiability, which is a property of both the model and the experimental protocol: a model which is structurally identifiable might lead to a poor practical identifiability of parameters if data points are noisy or scarce. The accuracy of model-based methods for inferring the values of parameters depends on the experimental protocol used to record the data, as observations need to be sufficiently informative to allow a correct estimation of the parameters. Contrary to structural identifiability, a quantitative criterion is lacking for practical identifiability, which is usually only qualitatively assessed. Non-practical identifiability refers to regimes in which

parameters can only be loosely estimated; but one would need to define what does “loose” mean. Such a definition could be intrinsic to the model: a model could be considered as practically identifiable given a certain experimental protocol if the expected variance of its parameters’ estimate is below a threshold. But this threshold would need to be arbitrarily defined. Here, we propose an extrinsic yet non-arbitrary definition of practical identifiability, by transforming a model identifiability problem into a model selection problem.

A model is said to be practically identifiable when its parameters can be correctly inferred given a certain experimental protocol. But, as explained previously, different possible models can be fitted on a data set. Recorded data need to be sufficiently informative not only to give a correct estimate of the parameters of a model, but also to select the correct model (i.e., the model from which they have been generated). We argue that a model is practically identifiable if and only if it is also correctly identified as the model providing the best fit to the data. For a given experimental protocol, we define the practical identifiability domain of a statistical model as the set of parameters for which the model is correctly identified as the ground truth compared to a simpler alternative submodel.

Our proposed definition of practical identifiability can be applied to any setting where submodels or a nested family of models can be defined. Here, we apply it to the particular problem of estimating the parameters of a chemical synapse. A classical biophysical model used to describe the stochastic release of neurotransmitter at chemical synapses is called the binomial model (Katz, 1969), for which different variants of increasing complexity (in term of the number of free parameters) can be considered.

Different model-based approaches have been proposed (Bykowska et al., 2019) for obtaining an accurate estimate of the parameters describing a synapse (namely, its number of independent release sites, their release probability upon the arrival of a presynaptic spike, the quantum of current elicited by one release event, etc.) These parameters cannot be measured directly, but can be inferred using excitatory postsynaptic currents (EPSCs¹) recorded on the post-synaptic side and elicited by experimental stimulation of the presynaptic cell. By measuring their values before and after a stimulation protocol, it is possible to study the mechanisms and loci of synaptic plasticity (Costa et al., 2015, 2017a,b) and homeostasis (Davis and Müller, 2015; Wentzel et al., 2018). On a more theoretical level, a correct inference of synaptic parameters is necessary to study the computational role of synaptic stochasticity (Levy and Baxter, 2002; Guo and Li, 2012). Finally, an accurate inference of synaptic parameters would allow to clarify the role of synaptic transmission in different diseases (Van Spronsen and Hoogenraad, 2010), such as mental retardation (Pfeiffer and Huber, 2009), schizophrenia (Stephan et al., 2006), Parkinson’s disease (Calabresi et al., 2006), autism (Südhof, 2008), Alzheimer’s disease (Selkoe, 2002), compulsive

¹It is also possible to perform model-based inference of synaptic parameters based on post-synaptic spike trains instead of EPSCs, as in Ghanbari et al. (2017, 2020)

behavior (Welch et al., 2007), and addiction (Kauer and Malenka, 2007).

Our contribution to the analysis of synaptic stochasticity is three-fold. Firstly, we propose a definition for the practical identifiability of a model of synaptic transmission, and compute the identifiability domains of different variants of the binomial release model. Besides, we observe that model selection criteria are classically derived by assuming that recorded data are not correlated, which does not hold for most models of chemical synapse. We extend the Bayesian Information Criterion (BIC), a classically used tool for model selection, to models with correlated data. Finally, a proper description of the model selection step is often missing in studies where a model based-approach is used to infer synaptic parameters. We show that our approach allows to perform *data free* model selection, i.e., to verify if a model used to fit data was indeed identifiable even if a proper model selection step had not been performed.

2. METHODS

2.1. Binomial Models of Neurotransmitter Release

2.1.1. The Classical Binomial Model

The quantal nature of synaptic transmission was first unveiled in Del Castillo and Katz (1954), in which the authors observed that the postsynaptic responses to presynaptic stimulations were all multiples of a small unit of current. They explained how the total response is built up of several of these units, or quanta, each of them arising from a single presynaptic release event. Upon the arrival of an action potential in the presynaptic terminal, vesicles are released with a given probability p . The binomial model (Katz, 1969) assumes that there are N independent release sites and that for each site the release probability p is identical. Therefore, the number of released vesicles k_i after spike i is distributed according to a binomial distribution. This model further assumes that each vesicle release gives rise to a quantal current q , such that the overall excitatory postsynaptic current is given by $e_i = qk_i + \epsilon$, where ϵ models a measurement noise typically drawn from a normal distribution with variance σ^2 . Under the binomial model described by its parameters N , p , q , and σ , the distribution of EPSCs is given by

$$p(e_i) = \sum_{k_i=0}^N p(e_i|k_i)p(k_i)$$

where k_i follows a binomial distribution with parameters N and p , and e_i conditioned on k_i follows a normal distribution with mean qk_i and variance σ^2 . Postsynaptic responses are characterized by their mean Npq and their variance $q^2Np(1-p) + \sigma^2$. A first feature of synaptic transmission is thus its stochasticity. Due to different sources of noise, such as probabilistic vesicles release or recording noise, postsynaptic recordings exhibit trial-to-trial variability.

2.1.2. Full Model of Synaptic Transmission

Although this simple binomial model accounts for synaptic stochasticity, it does not allow to model its dynamics:

postsynaptic responses do not only depend on the parameters of the synapse, but also on its previous activity. On the one hand, successive presynaptic stimulations within a short time interval will lead to a depletion of the readily-releasable vesicle pool, and hence to reduced successive postsynaptic responses, a phenomenon known as short-term depression. This can be modeled by assuming that the number of available vesicles at time i is $n_i \leq N$ (while the simplified binomial model described above assumes that all vesicles are readily releasable, and hence $n_i = N$). On the other hand, successive stimulations will gradually increase the presynaptic calcium concentration, and hence the release probability, which is called short-term facilitation.

Short-term depression and facilitation can be modeled using the Tsodyks-Markram model (Tsodyks et al., 1998; Costa et al., 2015). It consists in two ordinary differential equations, which model the proportion of available vesicles r_i and the release probability u_i at time i . r_i is reduced by an amount $u_i r_i$ after each presynaptic spike, and recovers back to 1 with a depression time constant τ_D between each spike. Similarly, u_i is increased by an amount $p(1 - u_i)$, and decays back to p (its baseline value) with a facilitation time constant τ_F . Different values of the parameters p , τ_D , and τ_F allow to represent different synaptic dynamics (either depression, facilitation, or no plasticity at all).

However, such a deterministic approach to short-term plasticity only allows to model averages, and neglects correlations between successive postsynaptic responses. In recent studies (Barri et al., 2016; Bird et al., 2016), models of synapses incorporating both short-term plasticity and binomial models of vesicles release and refill have been proposed. In these models, the release probability u_i evolves according to the equation of the Tsodyks-Markram model, while each vesicle refills with a probability $1 - \exp(-\Delta t_i / \tau_D)$, where Δt_i is the time interval between two successive presynaptic stimulations. This approach allows to represent both the stochasticity and the dynamics of neurotransmitter release, and to compute the likelihood of a set of recorded data \mathcal{D} given the parameters θ and the presynaptic stimulation protocol Ψ .

We consider a model of chemical synapse which encompasses both short-term depression (STD) and facilitation (STF) (Barri et al., 2016; Bird et al., 2016). Its parameters are (Figure 1A):

- N : the number of independent release sites [-]
- p : their initial release probability [-]
- σ : the recording noise. It encompasses both the noise coming from the experimental apparatus (thermal noise of the amplifier, electric line noise, etc.) and from the recordings per se (such as fluctuations in the membrane potential of the cell) [A]
- q : the quantum of current elicited by one release event [A]²
- τ_D : the time constant of vesicles refilling, and hence of short-term depression [s]
- τ_F : the time constant of Ca^{2+} dynamics, and hence of short-term facilitation [s]

²The outputs of a model of chemical synapse can be either electric postsynaptic currents (EPSC) or potentials (EPSP). In the latter case, σ and q will be expressed in [V] instead of [A].

which defines a vector $\theta = (N, p, q, \sigma, \tau_D, \tau_F)$. A probability conditioned on a parametrization θ is written p_θ .

$\mathbf{n} = \{n_i\}_{1 \leq i \leq T}$, $\mathbf{k} = \{k_i\}_{1 \leq i \leq T}$, and $\mathcal{D} = \{e_i\}_{1 \leq i \leq T}$ represent respectively the number of available vesicles at the moment of spike i , the number of vesicles released after spike i , and the i -th recorded EPSC (Figure 1B). The experimental protocol $\Psi = \{t_1, t_2, \dots, t_T\}$ encompasses the times of presynaptic stimulation: the time of the i -th spike is written t_i and $\Delta t_i = t_i - t_{i-1}$. For simplicity, we will drop the dependency on Ψ from the notations of probabilities.

The probability of recording \mathcal{D} is computed as the marginal of the joint distribution of the observations \mathcal{D} and the hidden variables \mathbf{n} and \mathbf{k} :

$$p_\theta(\mathcal{D}, \mathbf{n}, \mathbf{k}) = p_\theta(e_1 | k_1) p_\theta(k_1 | n_1) p_\theta(n_1) \prod_{i=2}^T p_\theta(e_i | k_i) p_\theta(k_i | n_i) p_\theta(n_i | n_{i-1}, k_{i-1}) \quad (1)$$

where $p_\theta(e_i | k_i)$ is the emission probability, i.e., the probability to record e_i knowing that k_i vesicles released neurotransmitter and assuming a normally distributed recording noise³:

$$p_\theta(e_i | k_i) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(e_i - qk_i)^2}{2\sigma^2}\right) \quad (2)$$

$p_\theta(k_i | n_i)$ is the binomial distribution and represents the probability that, given n_i available vesicles, k_i of them will indeed release neurotransmitter:

$$p_\theta(k_i | n_i) = \binom{n_i}{k_i} u_i^{k_i} (1 - u_i)^{n_i - k_i} \quad (3)$$

where the release probability u_i evolves as

$$u_i = p + u_{i-1}(1 - p) \exp\left(-\frac{\Delta t_i}{\tau_F}\right) \quad (4)$$

with $u_1 = p$. $p_\theta(n_i | n_{i-1}, k_{i-1})$ represents the process of vesicles refilling. During the time interval Δt_i , each empty vesicle can refill with a probability I_i :

$$p_\theta(n_i | n_{i-1}, k_{i-1}) = \binom{N - n_{i-1} + k_{i-1}}{n_i - n_{i-1} + k_{i-1}} I_i^{n_i - n_{i-1} + k_{i-1}} (1 - I_i)^{N - n_i} \quad (5)$$

with

$$I_i = 1 - \exp\left(-\frac{\Delta t_i}{\tau_D}\right) \quad (6)$$

It is usually assumed that, at the beginning of the experiment, all release sites are filled, and hence that $n_1 = N$ (Barri et al., 2016; Bird et al., 2016).

³Other distributions can also be used for the emission probability. Barri et al. (2016) assumed an inverse Gaussian to account for the observed right-skewness of mEPSP (Bekkers et al., 1990; Bhumbra and Beato, 2013)

2.2. Models, Submodels, and Nested Families

Definition 2.1. Model. For a given data set \mathcal{D} and experimental protocol Ψ , a model \mathcal{M} is defined as a triplet $\mathcal{M} = \{\Theta, \pi, \mathcal{L}\}$ where Θ is the space of parameters $\theta \in \Theta$, π is the prior for the parameters $\pi(\theta) = p(\theta | \mathcal{M})$, and \mathcal{L} is the likelihood of the parameters $\mathcal{L}(\theta | \mathcal{D}) = p(\mathcal{D} | \theta, \mathcal{M}, \Psi)$.

Examples: Different models can be derived from Equations (1) to (6). We consider four models of decreasing complexity:

Model \mathcal{M}_3 is the full model with both STD and STF. Its six parameters are N, p, q, σ, τ_D , and τ_F , and hence $\Theta_3 = \mathbb{N}^* \times [0, 1] \times (\mathbb{R}_+)^4$. Its likelihood function \mathcal{L}_3 is obtained by marginalizing out the hidden variables \mathbf{n} and \mathbf{k} :

$$\mathcal{L}_3(\theta | \mathcal{D}) = \sum_{\mathbf{n}, \mathbf{k}} p_\theta(\mathcal{D}, \mathbf{n}, \mathbf{k}) \quad (7)$$

where $p_\theta(\mathcal{D}, \mathbf{n}, \mathbf{k})$ is given by Equation (1).

Model \mathcal{M}_2 has only short-term depression (and no short-term facilitation). Its five parameters are N, p, q, σ , and τ_D , and hence $\Theta_2 = \mathbb{N}^* \times [0, 1] \times (\mathbb{R}_+)^3$. Its likelihood \mathcal{L}_2 can be derived from (7) by further assuming that u_i is a constant equal to p .

Model \mathcal{M}_1 shows no short-term plasticity at all, and can be derived from model \mathcal{M}_2 by further assuming that I_i (defined in 6) is a constant equal to 1 (and hence $n_i = N$). Its four parameters are N, p, q , and σ , and hence $\Theta_1 = \mathbb{N}^* \times [0, 1] \times (\mathbb{R}_+)^2$. In this setting, data points are independent and (7) becomes

$$\mathcal{L}_1(\theta | \mathcal{D}) = \prod_{i=1}^T \left(\sum_{k_i=0}^N p_\theta(e_i | k_i) p_\theta(k_i) \right) \quad (8)$$

with $p_\theta(k_i) = \binom{N}{k_i} p^{k_i} (1 - p)^{N - k_i}$ being the binomial distribution;

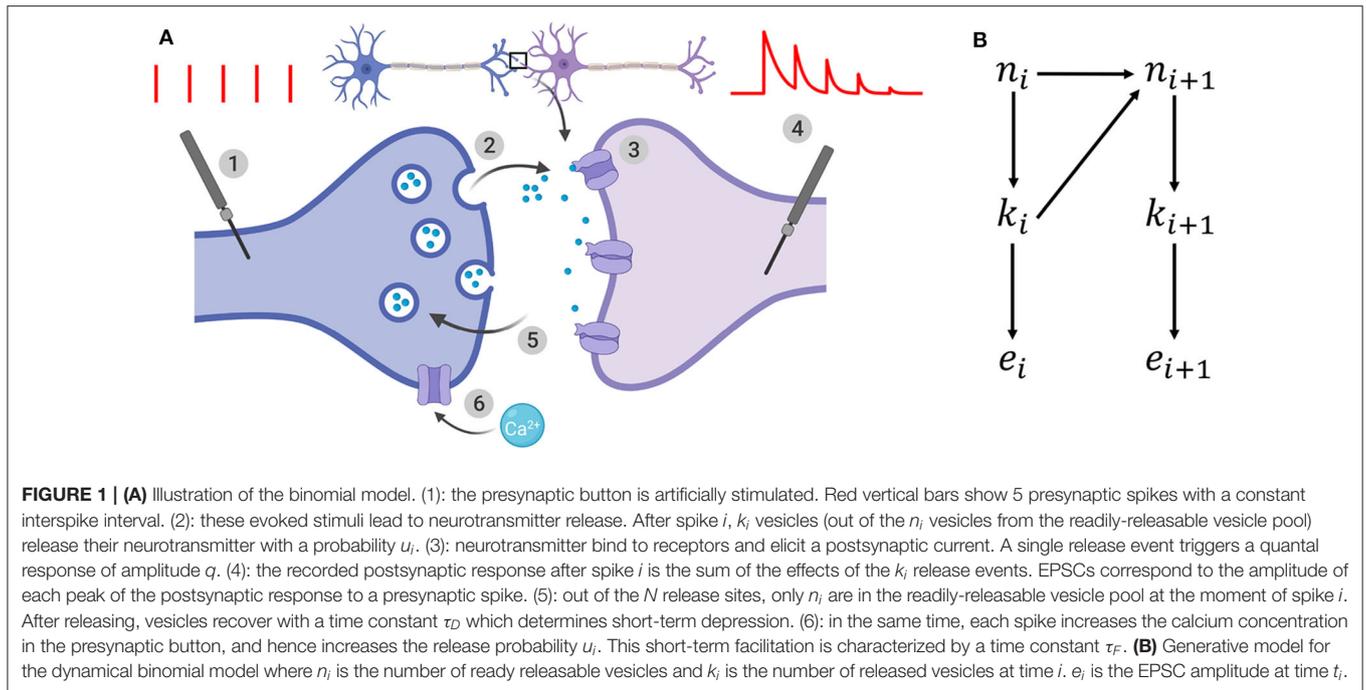
Model \mathcal{M}_0 is a Gaussian model, in which EPSCs are simply generated from a normal distribution parameterized by its mean and variance. Its two parameters are μ and σ^2 , and hence $\Theta_0 = \mathbb{R} \times \mathbb{R}_+$. Its likelihood \mathcal{L}_0 is simply

$$\mathcal{L}_0(\theta | \mathcal{D}) = \prod_{i=1}^T p_\theta(e_i) \quad (9)$$

with $p_\theta(e_i) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(e_i - \mu)^2}{2\sigma^2}\right)$.

To ensure the completeness of the definition of the models, we will assume for each parameter θ a uniform prior between two values θ_{\min} and θ_{\max} (Bird et al., 2016). Note however that the approximate identifiability domain defined in (17) does not depend on the prior.

Definition 2.2. Submodels. Although ubiquitous in statistics (as in the likelihood-ratio test or Pearson's chi-squared test), the notion of submodels (or nested models) is rarely formally defined in the literature (Edwards and MacCallum, 2012). It is usually said that \mathcal{M}_0 is a submodel of \mathcal{M}_1 (or is nested within \mathcal{M}_1)



if \mathcal{M}_0 can be obtained by constraining the parameters of \mathcal{M}_1 (Gottman, 1995). We propose the following formal definition, that encompasses the space of parameters, their priors, and their likelihood.

$\mathcal{M}_0 = \{\Theta_0, \pi_0, \mathcal{L}_0\}$ is said to be a submodel of $\mathcal{M}_1 = \{\Theta_1, \pi_1, \mathcal{L}_1\}$ if

1. $\Theta_0 \subset \Theta_1$ (i.e. the parameters of \mathcal{M}_0 also appear in \mathcal{M}_1)
2. $\pi_0(\theta_0) = \int_{\Theta_1 \setminus \Theta_0} \pi_1(\theta_0, \tilde{\theta}) d\tilde{\theta}$, $\forall \theta_0 \in \Theta_0$ (i.e. \mathcal{M}_0 and \mathcal{M}_1 share the same priors for the parameters they have in common)
3. $\forall \theta_0 \in \Theta_0, \exists \tilde{\theta}$ s.t. $p(\mathcal{D}|\theta_0, \mathcal{M}_0) = p(\mathcal{D}|(\theta_0, \tilde{\theta}), \mathcal{M}_1)$ with $(\theta_0, \tilde{\theta}) \in \Theta_1$ (i.e., \mathcal{M}_0 can be retrieved from \mathcal{M}_1 by constraining its parameters).

We use the notation $\mathcal{M}_0 \preceq \mathcal{M}_1$.

Examples: The model \mathcal{M}_2 with only short-term depression is a submodel of \mathcal{M}_3 (which accounts for both depression and facilitation). Indeed, they have the parameters N, p, q, σ , and τ_D in common, and \mathcal{M}_2 can be retrieved from \mathcal{M}_3 by constraining $\tau_F \rightarrow 0$. Similarly, the model without STP \mathcal{M}_1 is a submodel of \mathcal{M}_2 where $\tau_D \rightarrow 0$; and the uni-quantal model \mathcal{M}_0 is a submodel of the multi-quantal model \mathcal{M}_1 where $p = 1$ and $\mu = Nq$.

We propose the following definitions to characterize the nestedness of a family of models:

Definition 2.3. Nested family. $\mathcal{F} = \{\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_n\}$ is said to be a nested family if

$$\mathcal{M}_i \preceq \mathcal{M}_j, \forall 0 \leq i \leq j \leq n$$

2.3. Structural Identifiability

Definition 2.4. Structural identifiability domain. Consistently with Raue et al. (2009) and Massonis and Villaverde (2020), let the structural identifiability domain Θ_S of a model $\mathcal{M} = \{\Theta, \pi, \mathcal{L}\}$ be defined as:

$$\Theta_S = \{\theta \in \Theta \mid \forall \theta' \in \Theta, \theta \neq \theta' \iff p(\mathcal{D}|\theta, \mathcal{M}) \neq p(\mathcal{D}|\theta', \mathcal{M})\} \quad (10)$$

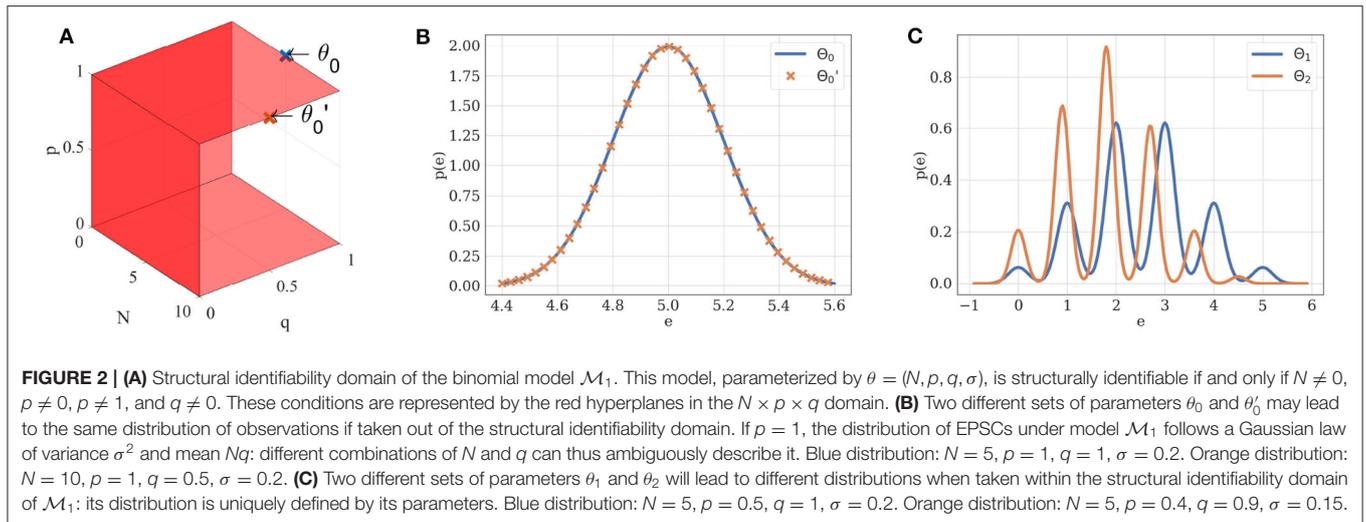
Similarly, \mathcal{M} is said to be structurally identifiable if $\Theta = \Theta_S$.

If θ is in the structural identifiability domain of \mathcal{M} , it can be uniquely identified from $p(\mathcal{D}|\theta, \mathcal{M})$. For instance, a Gaussian distribution of mean μ and variance σ^2 is uniquely defined by its parameters $\theta = (\mu, \sigma^2)$. Its structural identifiability domain is thus $\Theta_S = \mathbb{R} \times \mathbb{R}^+$. Similarly, if $N \neq 0, p \neq 0, p \neq 1$, and $q \neq 0$, the probability density of EPSCs under the binomial model without short-term plasticity \mathcal{M}_1 is structurally identifiable if we restrict Θ_1 to $\mathbb{N}^* \times]0, 1[\times (\mathbb{R}_+^*)^2$ (Figure 2).

2.4. Informative Domain

In some regimes, parameters may not be precisely inferred from observations, even though the model is otherwise structurally identifiable. Indeed, in practice we usually only have access to a finite number of (possibly noisy) observations. *Practical* identifiability thus differs from the *structural* identifiability defined in section 2.3.

A definition for the practical identifiability of a parameter has previously been proposed in Raue et al. (2009), along with an approach for detecting practical non-identifiabilities based on the profile likelihood (Venzon and Moolgavkar, 1988; Murphy and Van der Vaart, 2000). The authors first define the likelihood-based confidence intervals for the estimator $\hat{\theta}_i$ of the i -th



parameter of a model \mathcal{M} :

$$C_{i,\Delta} = \{\theta_i \mid \mathcal{L}(\hat{\theta}_i \mid \mathcal{D}) - \mathcal{L}(\theta_i \mid \mathcal{D}) < \Delta\}$$

where

$$\mathcal{L}(\theta_i \mid \mathcal{D}) = \max_{\theta_{j \neq i}} \mathcal{L}(\theta \mid \mathcal{D})$$

for a given threshold Δ . Then, they propose the following definition: *A parameter estimate $\hat{\theta}_i$ is practically non-identifiable, if the likelihood-based confidence region is infinitely extended in increasing and/or decreasing direction of θ_i , although the likelihood has a unique minimum for this parameter, meaning that the decrease in likelihood compared to the optimal parameters estimate stays below the threshold Δ in direction of θ_i .* When plotting the likelihood as a function of the parameters, practical non-identifiability can be seen as an infinitely extended flat valley, in which the decrease in likelihood stays below Δ . The authors also describe an algorithm for computing the profile likelihood and hence detecting such practical non-identifiabilities: *Structural non-identifiable parameters are characterized by a flat profile likelihood. The profile likelihood of a practically non-identifiable parameter has a minimum, but is not exceeding a threshold Δ for increasing and/or decreasing values of θ_i* (see Figure 3 in Raue et al., 2009).

An important limitation of this definition is to be data-dependent: it only holds for a specific set of recorded data \mathcal{D} . Indeed, likelihood-based confidence intervals, and hence practical identifiability, are defined with respect to a certain data set \mathcal{D} , and may thus vary for different realizations of the experiment. However, an identifiability criterion can be made data-independent by averaging it over all possible realizations of \mathcal{D} , i.e., by computing its expectation with respect to the distribution $p(\mathcal{D} \mid \theta^*, \mathcal{M}, \Psi)$. Such an averaged criterion would correspond to the *a priori* expected identifiability before a specific \mathcal{D} is recorded.

Practical information about θ is a function of the experimental protocol Ψ : for a given Ψ , the informative domain $\Theta_I(\Psi)$ of a model \mathcal{M} could be defined based on the variance of the estimator.

For instance, in a Bayesian setting, the domain $\Theta_I(\Psi)$ could be the set of parameters for which the expected informativeness of the posterior distribution of the parameters (measured as the Kullback-Leibler divergence between the posterior and the prior) is above a threshold Δ :

$$\Theta_I(\Psi) = \{\theta^* \in \Theta \mid \langle D_{KL}(p(\theta \mid \mathcal{D}, \mathcal{M}, \Psi) \parallel p(\theta \mid \mathcal{M})) \rangle_{p(\mathcal{D} \mid \theta^*, \mathcal{M}, \Psi)} \geq \Delta\} \quad (11)$$

Although data-independent, this definition suffers from the same limitation as the one proposed in Raue et al. (2009): it requires to set a specific threshold Δ . Instead of defining an arbitrary criterion Δ on the possible precision of parameters estimate, we will derive our definition from a model selection argument.

2.5. Model Selection

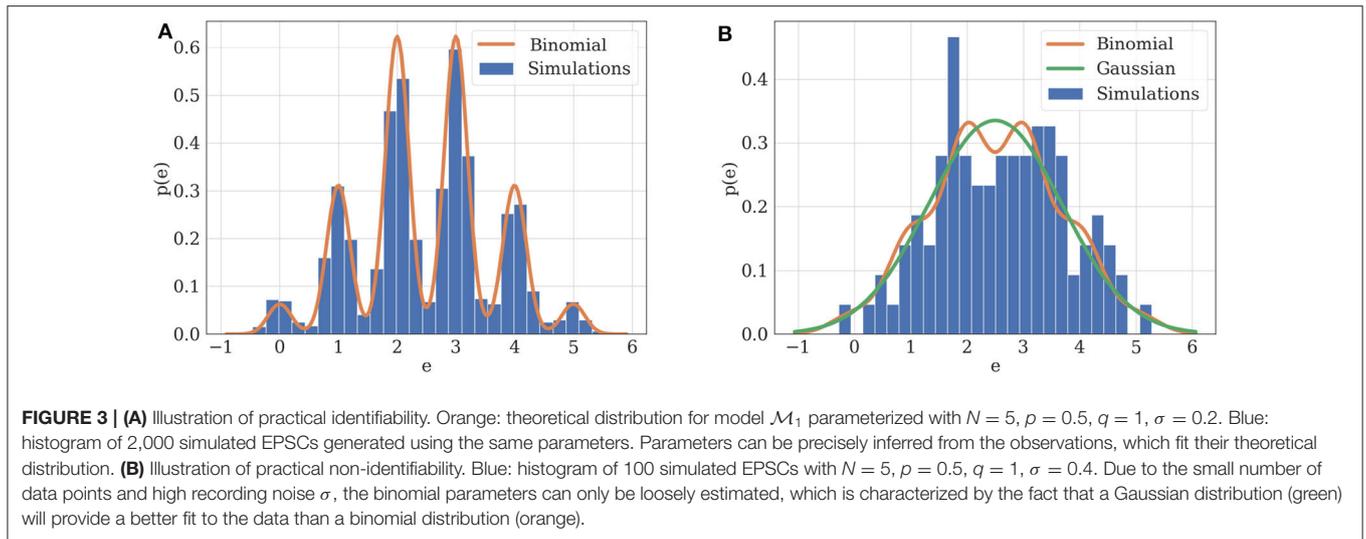
In model selection, the plausibility of two competing models $\mathcal{M} = \{\Theta, \pi, \mathcal{L}\}$ and $\mathcal{M}' = \{\Theta', \pi', \mathcal{L}'\}$ based on observations \mathcal{D} can be assessed using the Bayes Factor (Kass and Raftery, 1995):

$$B_{\mathcal{M}, \mathcal{M}'}(\mathcal{D}) = \frac{p(\mathcal{D} \mid \mathcal{M})}{p(\mathcal{D} \mid \mathcal{M}')} = \frac{\int_{\Theta} \mathcal{L}(\theta \mid \mathcal{D}) \pi(\theta) d\theta}{\int_{\Theta'} \mathcal{L}'(\theta \mid \mathcal{D}) \pi'(\theta) d\theta} \quad (12)$$

If the Bayes Factor is superior to 1, then the evidence for \mathcal{M} is higher than the evidence for \mathcal{M}' . It is worth pointing out that the Bayes Factor will not only favor models which provide a good fit to the data, but also includes a tendency to favor simpler models, a natural form of Occam's Razor (Jefferys and Berger, 1991; MacKay and MacKay, 2003). Indeed, a complex model (i.e., a model with many independent parameters or with a broader prior for its parameters) will be able to explain a larger set of possible observed data than a simple model; but this comes at the price of spreading its likelihood over a larger set of possible outcomes. Hence, if two models fit the observed data equally well, the simpler one will be favored.

2.6. Proposed Definition of Practical Model Identifiability

To compute the identifiability domain of any model \mathcal{M} compared to another model \mathcal{M}' , we introduce the Average Log



Bayes Factor:

$$\mathcal{B}_{\mathcal{M}, \mathcal{M}'}(\theta^*, \Psi) = \langle \log \mathcal{B}_{\mathcal{M}, \mathcal{M}'}(\mathcal{D}) \rangle_{p(\mathcal{D}|\theta^*, \mathcal{M}, \Psi)} \quad (13)$$

For a given parameter θ^* and protocol Ψ , model \mathcal{M} is said to be practically identifiable compared to \mathcal{M}' if $\mathcal{B}_{\mathcal{M}, \mathcal{M}'}(\theta^*, \Psi) \geq 0$. Intuitively, the identifiability domain of \mathcal{M} compared to \mathcal{M}' corresponds to all the settings (parameters and protocols) for which, on average, data generated from the ground truth \mathcal{M} will be better explained by \mathcal{M} than by \mathcal{M}' .

In contrast to the definition in Raue et al. (2009), our proposed definition does not require to set a (possibly arbitrary) threshold Δ . Instead, it is derived from a model selection criterion. We argue that the parameters of a model \mathcal{M} are practically identifiable if \mathcal{M} is itself practically identifiable. In some settings (as for the nested models of chemical synapse described in section 2.1), a family of submodels might naturally arise, while the choice of a threshold Δ would be arbitrary.

Another interest of our approach is to be data-independent, while the definition proposed in Raue et al. (2009) only holds for a specific set of recorded data \mathcal{D} . Indeed, we define practical identifiability as a data-independent and intrinsic property of the model \mathcal{M} and experimental protocol Ψ . As the log-Bayes Factor in (13) is averaged over all possible realizations of \mathcal{D} , it corresponds to the *a priori* expected identifiability before \mathcal{D} is recorded. Our approach thus allows to define practical identifiability domains:

Definition 2.5. Practical identifiability domain. Consider a model $\mathcal{M} = \{\Theta, \pi, \mathcal{L}\}$ and a submodel $\mathcal{M}' = \{\Theta', \pi', \mathcal{L}'\}$ of \mathcal{M} . For a given experimental protocol Ψ , the practical identifiability domain $\Theta_p(\Psi)$ of \mathcal{M} is the set of parameters θ^* for which it is identifiable compared to its submodel:

$$\Theta_p(\Psi) = \{\theta^* \in \Theta \mid \mathcal{B}_{\mathcal{M}, \mathcal{M}'}(\theta^*, \Psi) \geq 0\} \quad (14)$$

Note that in the limit where the priors π and π' are highly peaked (i.e., $\pi(\theta) = \delta(\theta - \hat{\theta})$ and $\pi'(\theta) = \delta(\theta - \hat{\theta}')$), the condition $\mathcal{B}_{\mathcal{M}, \mathcal{M}'}(\theta^*, \Psi) \geq 0$ is always satisfied due to Gibbs' inequality. In this case we have $\Theta_p(\Psi) = \Theta, \forall \Psi$. However, generically the condition $\mathcal{B}_{\mathcal{M}, \mathcal{M}'}(\theta^*, \Psi) \geq 0$ is not always satisfied since $p(\mathcal{D}|\mathcal{M})$ is not equal to $p(\mathcal{D}|\theta^*, \mathcal{M})$. The latter is the probability of observing \mathcal{D} given \mathcal{M} and a certain parametrization θ^* , while the former is the marginal likelihood over all parameters (12).

Two examples can illustrate this correspondence between model selection and parameter inference. Consider first the case of data recorded from \mathcal{M}_1 . If the experimental protocol is not sufficiently informative (i.e., if data are scarce or noisy), not only will the inference of synaptic parameters be poor, but a Gaussian distribution will also provide a better fit than a binomial release model to the data. Indeed, as $[e_i|k_i] \sim \mathcal{N}(qk_i, \sigma)$, in the absence of recording noise (i.e., if $\sigma = 0$), the distribution of EPSCs is a series of Dirac delta functions located at each multiple of the quantal size qk for $k \in \{0, 1, \dots, N\}$. In this ideal case, q is clearly identifiable (Figure 3A). However, upon addition of a recording noise of amplitude σ , EPSCs are normally distributed around qk for $k \in \{0, 1, \dots, N\}$, and the peaks on the histogram corresponding to each multiple of the quantal size might overlap if σ is sufficiently high with respect to q (Figure 3B).

Similarly, we can consider the example of a synapse which shows short-term depression (STD) with a time constant τ_D (model \mathcal{M}_2). If the presynaptic cell is stimulated with an interspike intervals longer than τ_D , no depression will be visible in the recorded data, and the true model with STD will not be identifiable from a simpler binomial model without STD. In the same time, it will impossible to correctly infer the value of τ_D .

Our proposed definition of practical identifiability and of the identifiability domain of a model extend the landscaping technique introduced in Navarro et al. (2004) as well as the framework for testing identifiability of Bayesian models introduced in Acerbi et al. (2014). Especially, comparing the expected supports $\langle \log p(\mathcal{D}|\mathcal{M}) \rangle_{p(\mathcal{D}|\theta^*, \mathcal{M}, \Psi)}$ and $\langle \log p(\mathcal{D}|\mathcal{M}') \rangle_{p(\mathcal{D}|\theta^*, \mathcal{M}, \Psi)}$ of \mathcal{M} and \mathcal{M}' (given that values

are averaged over $\langle \cdot \rangle_{p(\mathcal{D}|\theta^*, \mathcal{M}, \Psi)}$ allows us to define a quantitative criterion for identifiability.

The model evidence $p(\mathcal{D}|\mathcal{M})$ in (12) is often intractable in practice for complex models, as it requires to integrate marginals for each parameter. Different methods have been proposed to approximate it: MCMC computations (Weinberg et al., 2012), Savage-Dickey method (Wagenmakers et al., 2010), supermodels (Mootoovaloo et al., 2016). A practical and time-efficient approximation of the model evidence is given by the Bayesian Information Criterion $\text{BIC}_{\mathcal{M}}(\mathcal{D})$ (Schwarz et al., 1978):

$$\text{BIC}_{\mathcal{M}}(\mathcal{D}) = -2 \log p(\mathcal{D}|\hat{\theta}, \mathcal{M}) + k_{\mathcal{M}} \log(T) \approx -2 \log p(\mathcal{D}|\mathcal{M}) \quad (15)$$

where $\hat{\theta} = \arg \max_{\theta} \mathcal{L}(\theta|\mathcal{D})$ is the maximum likelihood estimator (MLE) of $\mathcal{L}(\theta|\mathcal{D})$, $k_{\mathcal{M}} = \dim(\Theta)$ is the number of independent parameters of \mathcal{M} , and $T = |\Psi|$ is the number of data points in \mathcal{D} . A detailed derivation is provided in **Supplementary Material**. The BIC is the sum of two terms: a likelihood term $-2 \log p(\mathcal{D}|\hat{\theta}, \mathcal{M})$ which represents the ability of the model \mathcal{M} to explain \mathcal{D} , and a penalty term $k_{\mathcal{M}} \log(T)$ which favors simpler models, as explained in section 2.5.

The BIC is commonly used as an approximation of the model evidence $p(\mathcal{D}|\mathcal{M})$ in model selection: the model with the lowest BIC is preferred over the others. The main advantage of using the BIC is to transform a complex integration problem (i.e., the computation of $p(\mathcal{D}|\mathcal{M})$) into a simpler optimization problem (i.e. the computation of $\hat{\theta}$). Besides, it allows to perform model selection without the need to specify a prior for the parameters, and is thus a popular tool for model selection (Daw et al., 2011).

As stated in **Supplementary Material**, the approximation $\text{BIC}_{\mathcal{M}}(\mathcal{D}) \approx -2 \log p(\mathcal{D}|\mathcal{M})$ is only valid under the hypothesis that data points are independent and identically distributed (i.i.d.), which is not the case for models with short-term plasticity. If data are correlated, we are left with the following approximation, which does not simplify in the general case:

$$-2 \log p(\mathcal{D}|\mathcal{M}) \approx -2 \log p(\mathcal{D}|\hat{\theta}, \mathcal{M}) + \log(|H(\hat{\theta})|) \quad (16)$$

where $H(\hat{\theta})$ is the Hessian matrix of $-\log p(\mathcal{D}|\theta, \mathcal{M})$ in Equation (15).

We emphasize that the classical definition of the BIC (15) should not be used if observations are correlated. Here, for models in which output are not independent, we use the approximation given by Equation (16), in which the term $k_{\mathcal{M}} \log(T)$ in the BIC is replaced by $\log(|H(\hat{\theta})|)$. In some settings, the computation of the Hessian matrix can be challenging. However, MCMC methods can be used to approximate $H(\hat{\theta})$, even without an explicit expression for the gradient of the function (Spall, 2005). In our case, a numerical method for computing $|H(\hat{\theta})|$ is detailed in the **Supplementary Material**.

Using approximation (15) in definition (14) yields the following approximation for the practical identifiability domain in case the model evidence $p(\mathcal{D}|\mathcal{M})$ in (12) is intractable:

$$\begin{aligned} \tilde{\Theta}_p(\Psi) &= \{\theta^* \in \Theta \mid \langle \text{BIC}_{\mathcal{M}}(\mathcal{D}) \rangle_{p(\mathcal{D}|\theta^*, \mathcal{M}, \Psi)} \\ &\leq \langle \text{BIC}_{\mathcal{M}'(\mathcal{D})} \rangle_{p(\mathcal{D}|\theta^*, \mathcal{M}, \Psi)}\} \quad (17) \end{aligned}$$

3. RESULTS

3.1. Identifiability Domain of the Binomial Model Without Short-Term Plasticity

We study here the conditions under which a binomial model without short-term plasticity \mathcal{M}_1 can be correctly identified from a Gaussian model having the same mean and variance (\mathcal{M}_0). In order for the binomial model to be identifiable from a Gaussian quantum-less distribution, the recording noise needs to be sufficiently low compared to q for the peaks on the histogram of recorded EPSC to be identified. We will thus plot the identifiability domain as a function of the recording noise of amplitude σ for a fixed q . The identifiability domain corresponds to the points θ in the parameters space Θ_1 for which the average BIC of \mathcal{M}_1 over all possible outputs of \mathcal{M}_1 parameterized with θ is lower than the average BIC of \mathcal{M}_0 .

Per se, the identifiability domain depends on all the parameters of \mathcal{M}_1 , as well as on the experimental protocol. For simplicity and in order to obtain a plot in 2 dimensions, we will only plot it as a function of p and σ while holding other variables to a fixed value. For a given experimental setup Ψ (which encompasses only the number of recorded data points T , the inter-spike intervals playing no role in these models), the following Markov-Chain-Monte-Carlo (MCMC) procedure is implemented:

1. A set of values p^* and σ^* are chosen from the space of possible values for p and σ ;
2. Using p^* and σ^* , 400 independent data sets $(\mathcal{D}_i)_{1 \leq i \leq 400}$ are generated from \mathcal{M}_1 . Each data set consists in T EPSCs;
3. For each \mathcal{D}_i , the BIC of both models are computed; these values are averaged over i to compute an average BIC and identifiability is assessed if \mathcal{M}_1 is preferred over \mathcal{M}_0 , which corresponds to the black dots in **Figure 4A**.

The procedure of plotting a complete identifiability domain can be quite time-consuming. Indeed, it requires to span the entire space of parameters; for each vector θ^* , to generate a large number of independent data sets (\mathcal{D}_i); and for each of these data sets, to compute the maximum likelihood estimator $\hat{\theta}$ using the Expectation-Maximization algorithm (Barri et al., 2016). Details on the computation of $\hat{\theta}$ are available in **Supplementary Material**.

However, as both models \mathcal{M}_0 and \mathcal{M}_1 generate i.i.d. data, and by making the approximation $\hat{\theta} \approx \theta^*$ (i.e., by assuming that the maximum likelihood estimator $\hat{\theta}$ will be close to the true value θ^* from which data were generated), the condition that model \mathcal{M}_1 is identifiable (17) can be approximated as follows:

$$\begin{aligned} &-2T \int p(e|\theta^*, \mathcal{M}_1) \log p(e|\theta^*, \mathcal{M}_1) de + k_{\mathcal{M}_1} \log(T) \leq \\ &-2T \int p(e|\hat{\theta}_{\mathcal{M}_0}, \mathcal{M}_1) \log p(e|\hat{\theta}_{\mathcal{M}_0}, \mathcal{M}_0) de + k_{\mathcal{M}_0} \log(T) \quad (18) \end{aligned}$$

where $\hat{\theta}_{\mathcal{M}_0} = (\mu, \sigma^2)$ represents the mean $\mu \approx N^* p^* q^*$ and the variance $\sigma^2 \approx N^* p^* (1 - p^*) q^{*2} + \sigma^{*2}$ of the data generated from \mathcal{M}_1 .

The condition specified by inequality (18) can be checked for any point θ^* without the need to generate a large number of independent data sets nor to compute the estimator $\hat{\theta}$.

Solving (18) numerically for σ allows to draw the border of the identifiability domain of \mathcal{M}_1 , represented as solid lines in **Figures 4A,B**.

Several points are worth highlighting. Firstly, **Figure 4A** shows a good agreement between the results of the MCMC simulations (black dots) and those from the semi-analytical method (18) (blue line). Secondly, as expected, **Figures 4A,B** illustrate that the identifiability domain increases with the number of data points T : intuitively, a larger data set facilitates the correct identification of a complex model. Besides, irrespective of the values of T and σ , for $p = 0$ and $p = 1$ the model \mathcal{M}_1 is structurally indistinguishable from a Gaussian distribution (see **Figure 2A**). Finally, the maximum noise σ which makes the binomial model \mathcal{M}_1 indistinguishable from a Gaussian distribution \mathcal{M}_0 is larger for extreme values of p (close to 0.9 or 0.1) than for $p = 0.5$. Indeed, in the latter case, the distributions of EPSC will be symmetric (as in the upper panel of **Figure 4C**), and hence just a little increase in recording noise will be enough to cover the inter-peak intervals and make the distribution Gaussian-shaped. In the former case, the distribution will be highly skewed, and thus difficult to approximate with a normal distribution.

The same approach can then be extended to more complicated models, by defining their identifiability domains as the part of the parameters plane where their average BIC will be lower than the BIC of a simpler one.

3.2. Identifiability Domain of the Binomial Model With Short Term Depression

We study here the conditions under which a binomial model with short-term depression (\mathcal{M}_2) can be correctly identified from a model without short-term plasticity (\mathcal{M}_1). In a first example, we assume that the presynaptic cell is stimulated at a constant inter-spike interval (ISI), which needs to be sufficiently short with respect to the time constant τ_D to make depression visible. We thus plot the identifiability domain as a function of both p and τ_D . We use the same method as in 3.1: For each set of parameters p^* and τ_D^* , 400 independent data sets are generated from \mathcal{M}_2 . Both models \mathcal{M}_2 and \mathcal{M}_1 are fitted on them, and black dots in **Figure 5A** correspond to the parameters for which the average BIC of \mathcal{M}_2 is lower than the average BIC of \mathcal{M}_1 .

As expected, we verify that the identifiability of \mathcal{M}_2 is only possible when τ_D is sufficiently long with respect to the inter-spike interval. Besides, if the release probability p is low, correlations between recordings will be weak and the effect of short-term depression will not be detectable. A major difference between models \mathcal{M}_1 and \mathcal{M}_2 is that, in the latter, observations $\{e_i\}_{1 \leq i \leq T}$ are not i.i.d.. The value of the i -th recorded EPSC is a function of the number of available and released vesicles n_i and k_i , which in turn depend on their previous values and on the ISI Δt_i . This has two main consequences. Firstly, using the same approximation as in (18) would lead to a biased estimate of the identifiability domain. Secondly, the classical definition of the BIC (15) should not be used since observations are correlated. Rather, we use Equation (16) to compare the evidence for \mathcal{M}_1 and \mathcal{M}_2 for a given data set.

Plotting the identifiability domain of a model also allows to investigate how the identifiability depends on the experimental protocol. For model \mathcal{M}_1 , we already saw that the identifiability domain increases with the number of data points T (see **Figures 4B,D**): a larger data set is more informative and allows for more reliable inference. In this case, T is the only experimental variable, as observations $\{e_i\}_{1 \leq i \leq T}$ are i.i.d. On the other hand, the identifiability domain of \mathcal{M}_2 will depend not only on the number of data points, but also on the stimulation protocol. We compare the constant stimulation protocol (T data points with a constant inter-spike interval $ISI = 0.05s$) of **Figure 5A** with a more realistic stimulation protocol in **Figure 5B**. In electrophysiological recordings, synaptic transmission is classically studied by stimulating the presynaptic cell with short regular train of spikes at a given frequency, followed by a recovery spike. This protocol is then repeated several times (Costa et al., 2013; Barri et al., 2016; Bird et al., 2016). Such periodic trains are more informative than a constant stimulation protocol, as they allow to probe a broader range of temporal dynamics.

In **Figure 5B**, we use 20 repetitions of a train of 4 spikes at 20Hz ($ISI = 0.05s$), followed by a recovery spike 0.5s later. This protocol entails the same number of data points $T = 100$ as the constant one, but allows to identify STD for a broader range of depression time constants (namely, for $\tau_D < 0.3s$). On the other hand, since there are fewer successive stimulations within a short time interval than in the constant protocol, depression can only be identified when the release probability p is sufficiently high to induce vesicle pool depletion.

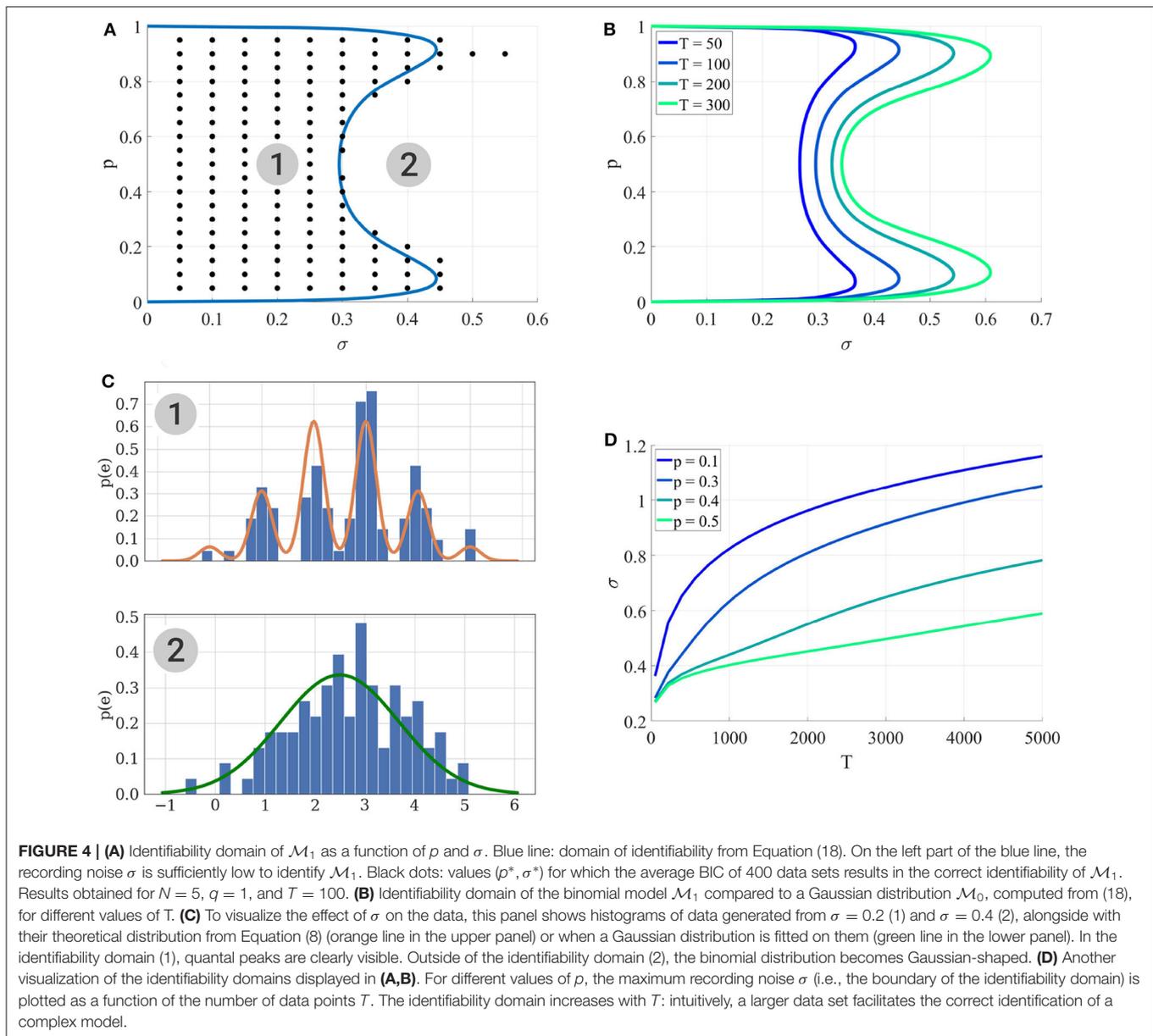
3.3. Data Free Model Selection

In model-based inference of synaptic parameters, a crucial step related to the estimation of the parameters is model selection, which is usually performed in several steps:

1. Data \mathcal{D} are acquired from a synapse using protocol Ψ ;
2. A nested family of $n + 1$ possible models $\mathcal{F} = \{\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_n\}$ is defined;
3. Each of these models is fitted on \mathcal{D} to obtain $n + 1$ MLE $\hat{\theta}_0, \hat{\theta}_1, \dots, \hat{\theta}_n$;
4. A model selection criterion (Bayes Factor, BIC, AIC...) is computed to quantify and rank the fitness of each model on \mathcal{D} ;
5. If \mathcal{M}_i is the selected model, then its MLE $\hat{\theta}_i$ is selected as the inference of synaptic parameters.

However, in many studies (Barri et al., 2016; Bird et al., 2016; Ghanbari et al., 2017), such a model selection step is not described. In this section, we investigate the possibility, having only access to the inferred values $\hat{\theta}$ of the parameters and to the description of the experimental protocol Ψ , to verify that the model used to infer $\hat{\theta}$ was indeed practically identifiable (i.e., to verify if a simpler model would have given a better fit to the data).

We use the notation \mathcal{D} for a set of data generated from a model \mathcal{M} parameterized with θ^* , and $\hat{\theta}$ the inferred parameters obtained by fitting the parameters θ of model \mathcal{M} on \mathcal{D} . If θ^* is within the practical identifiability domain of \mathcal{M} as we defined it, it is then possible to correctly infer it from \mathcal{D} , and hence $\hat{\theta} \approx \theta^*$



will also be within the identifiability domain of \mathcal{M} . Reciprocally, if $\hat{\theta}$ is not in the identifiability domain of \mathcal{M} , then a submodel would have provided a better fit to the data \mathcal{D} than \mathcal{M} . Is it thus possible to verify if \mathcal{M} overfits the data simply by verifying if $\hat{\theta}$ is in its identifiability domain, without having access to the data.

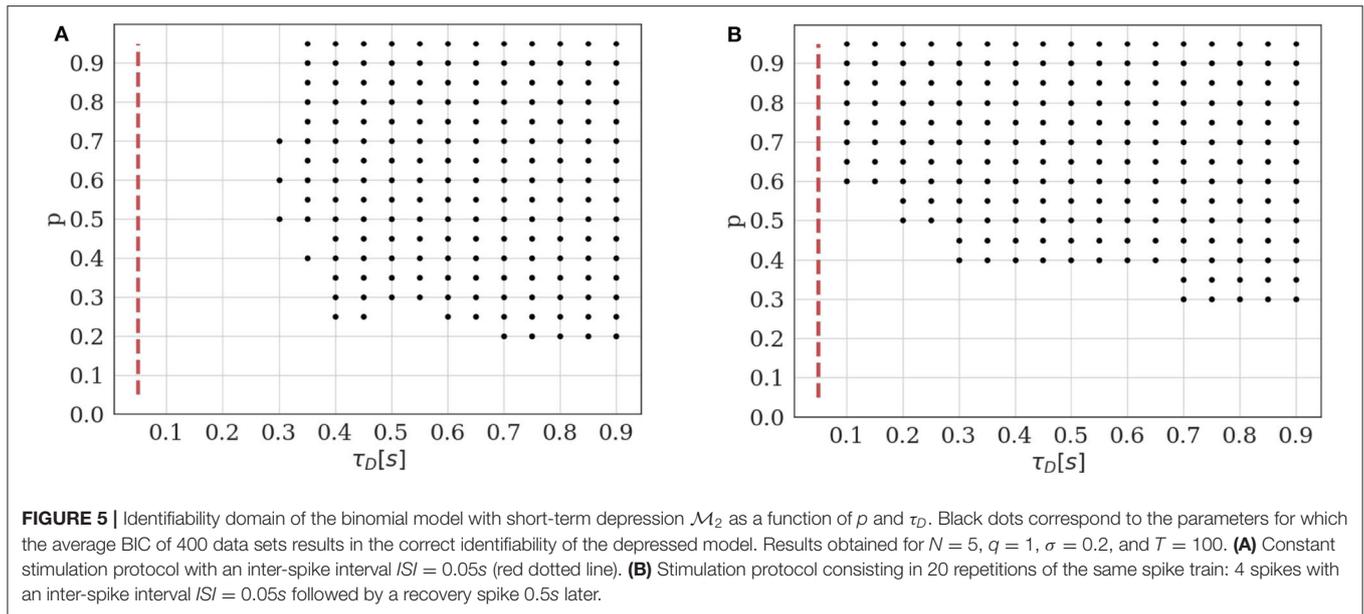
This is illustrated in **Figures 6A,B**, where \mathcal{M}_1 is fitted on data generated from its submodel \mathcal{M}_0 . For six different values of θ_0^* (**Figure 6A**), the inferred parameters are out of the identifiability domain of \mathcal{M}_1 (**Figure 6B**), showing that data are indeed better explained by \mathcal{M}_0 than by \mathcal{M}_1 .

3.3.1. First Example: Application to the Data From Katz et al. (1954)

We first apply our *data free* model selection method to the seminal 1954 paper from Del Castillo and Katz (1954), in

which the quantal nature of neurotransmitter release is identified for the first time. In order to observe mEPSP, they artificially reduced the release probability p by lowering the external calcium concentration. Although the quantal components of postsynaptic potentials are clearly visible and thoroughly analyzed, it would be interesting to verify, using our proposed model identifiability analysis method, that the binomial model (i.e., a multi-quantal distribution) indeed provides a better fit to the data than a simpler Gaussian model (i.e., a uni-quantal distribution).

Data (Fatt and Katz, 1952) consist in 328 EPSPs recorded at the neuro-muscular junction (NMJ) of a frog muscle. Fitting the binomial model and running the Expectation-Maximization algorithm on them yields $\hat{N} = 42$, $\hat{p} = 0.013$, $\hat{q} = 0.875$ mV, and $\hat{\sigma} = 0.15$ mV (and hence $\frac{\hat{\sigma}}{\hat{q}} \approx 17\%$). For this particular example, we have not only access to the inferred parameters $\hat{\theta}$,



but also to the data: it is thus possible to directly compare the BIC of a Gaussian ($BIC_{\mathcal{M}_0} = 764.95$) and of a binomial ($BIC_{\mathcal{M}_1} = 470.37$) distributions, which indeed confirms that data are better explained by the binomial quantal model.

However, even without the data, we can verify that the point in the parameter-protocol space specified by $\hat{\theta}$ (the inferred values of the parameters) and Ψ (the number of data points $T = 328$) is indeed within the identifiability domain of the binomial model \mathcal{M}_1 compared to \mathcal{M}_0 (see **Figure 6C**), thus confirming the multi-quantal nature of the recordings.

3.3.2. Second Example: Application to the Data From Barri et al. (2016)

We then apply our method to the results presented in the 2016 paper from Barri et al. (2016), in which the complete binomial model (with STD and STF) is fitted on recordings from layer 5 pyramidal neurons. They use a slight variation of the binomial release model with short term plasticity described by Equations (1)–(6), in which the emission probability does not follow a Gaussian, but an inverse Gaussian distribution:

$$p_{\theta}(e_i|k_i) = \frac{q^{3/2}k_i}{\sqrt{2\pi\sigma^2e_i^3}} \exp\left(-\frac{q(e_i - qk_i)^2}{2\sigma^2e_i}\right) \quad (19)$$

To verify that the data would not have been better fitted by a simpler model (and hence, that the published estimates of synaptic parameters are reliable), 100 synthetic data sets were generated from the complete binomial model using the stimulation protocol and the inferred values of the parameters described in (Barri et al., 2016):

- 20 repetitions of the same stimulation protocol consisting in 8 presynaptic spikes at 20Hz followed by a recovery spike 500 ms later;

- $N^* = 17$, $p^* = 0.27$, $q^* = 0.18$ mV, $\sigma^* = 0.06$ mV, $\tau_D^* = 202$ ms, and $\tau_F^* = 449$ ms.

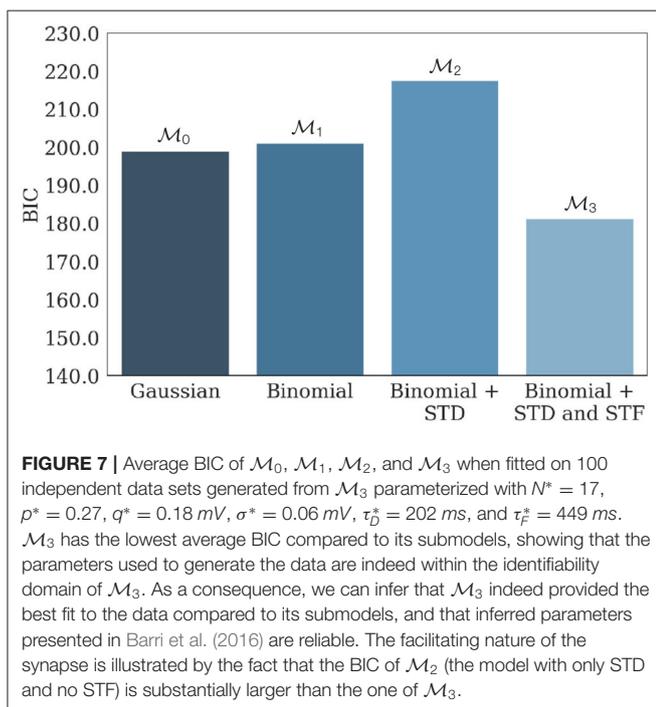
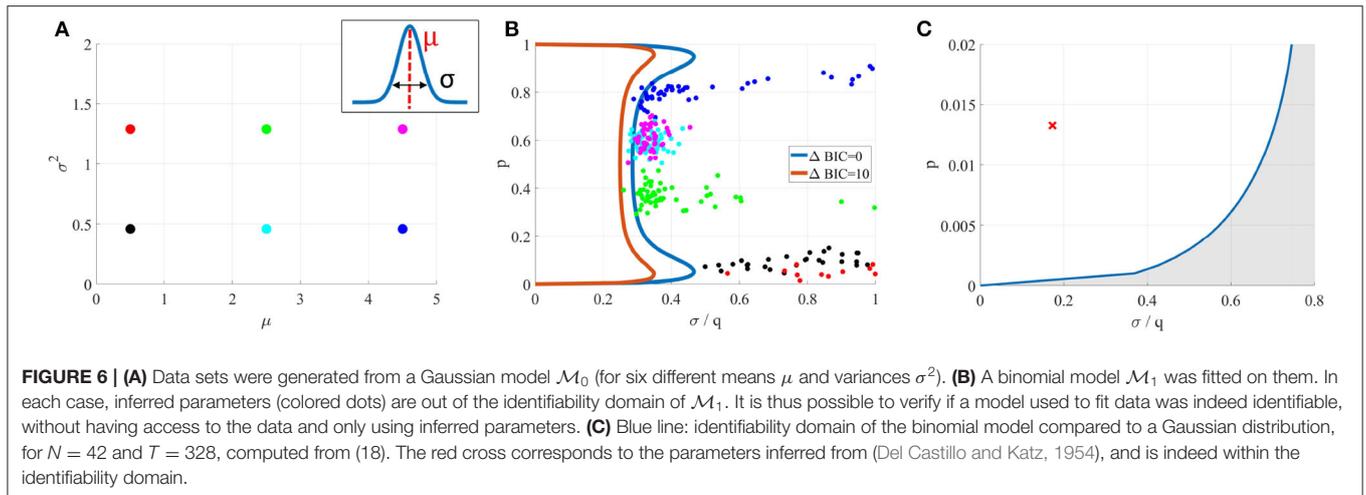
\mathcal{M}_0 , \mathcal{M}_1 , \mathcal{M}_2 , and \mathcal{M}_3 were then fitted on the generated data. Average values of their respective BIC are presented in **Figure 7**, and confirm the identifiability of the model used in the study.

4. DISCUSSION

Obtaining an accurate estimate of the parameters of a system from noisy and scarce observations is a crucial problem in neuroscience. Especially, different methods have been proposed for estimating the parameters describing a synapse (namely, its number of independent release sites, their release probability upon the arrival of a presynaptic spike, the quantum of current elicited by one release event, the time constants of depression and facilitation, etc...). Inferring their values allows to analyze the locus of synaptic plasticity and homeostasis; to study possibly synapse-related diseases; and more generally to investigate learning, memory, and neural dynamics, which are mediated by synaptic transmission.

It is usually impossible to measure directly these parameters. However, they can be estimated by fitting a biophysical model of synapse on currents recorded on the post-synaptic side and elicited by experimental stimulation of the presynaptic cell. This approach for estimating the parameters of a system is referred to as model-based inference. As different competing models may be used to describe the system and explain its output, model-based inference of parameters thus raises the question of what makes a good model.

Prior to any data recording, a required property for competing models is identifiability. Although structural identifiability has been widely studied, no quantitative criterion exists for practical identifiability, which is usually only qualitatively assessed. Here, we propose a definition for the practical identifiability of a



model, based on its expected support given the distribution of the data. We define the practical identifiability domain of a statistical model as the set of parameters for which the model is correctly identified as the ground truth compared to a simpler alternative submodel, and we study the identifiability domains of different models of synaptic release. In the process, we propose an extension of the Bayesian Information Criterion (BIC) for models with correlated data. The BIC is a widely used tool for model selection, but it is derived by assuming that the outputs of the system are mutually independent, which is not the case for models of chemical synapse. Finally, we show that our approach allows to perform *data free* model selection, i.e., to verify the identifiability of a model without having access to the data.

The definition of practical identifiability we introduced here differs from the influential contribution of Raue et al. (2009) in two ways. Firstly, our definition is data-independent: it does not only hold for a specific set of recorded data \mathcal{D} . Indeed, we define practical identifiability as an intrinsic property of the model \mathcal{M} and experimental protocol Ψ . We actually define the *a priori* expected identifiability before a specific \mathcal{D} is recorded, which allows to study how identifiability is affected by different experimental protocols. Secondly, since our definition is derived from a model-selection argument, it does not require to select a possibly arbitrary threshold on the practical identifiability of parameters. Rather, it is defined with respect to a particular submodel. Although the choice of the submodel might itself be arbitrary, we argue that nested models and families naturally arise in commonly used statistical techniques, such as polynomial regression (Edwards and MacCallum, 2012), or Generalized Linear Models (GLM) (Pillow et al., 2008). Especially, the widespread use of phenomenological models in neuroscience (Kobayashi et al., 2009; Melanson et al., 2014; Wang et al., 2016; Levenstein et al., 2020) makes the use of nested families and submodels relevant.

Another limitation of our approach is its practical implementation. As mentioned, the model evidence $p(\mathcal{D}|\mathcal{M})$, on which our definition is based, is often intractable in practice for complex models, and needs to be estimated. For practical purpose, we used the Bayesian Information Criterion (BIC) to compute the identifiability domains of our different models of synapse. However, we acknowledge that the BIC only provides a valid approximation of the model evidence when the number of samples is sufficiently large. A future step would be to study the robustness of our approach to different computations of the model evidence or to other approximations, such as the Akaike Information Criterion (AIC) (Burnham and Anderson, 2004).

Our identifiability domains are similar to the approach adopted in Koyama (2012), in which the authors study under which regime of rate fluctuation are the temporal variations of a neuron firing rate correctly identified. Spike trains are generated from a model of spiking neuron with a fluctuating

firing rate (complex model); but under a certain value of rate fluctuation, this model becomes indistinguishable from a model of spiking neuron with a constant rate (simple model). Plotting the identifiability of the fluctuating-rate model as a function of the amplitude of rate fluctuation allows them to identify which distribution of inter-spike intervals has the broader identifiability domain (and thus maximizes the efficiency of rate fluctuation transmission).

In model-based inference and parameter estimation, one is often interested in obtaining theoretical bounds on the achievable error performance. Such theoretical bounds allow to assess *a priori* the possibility to correctly infer the parameters. A well-known theoretical result is the Cramér-Rao bound (Van Trees, 2004; Van Trees and Bell, 2007), which provides a lower bound on the variance of the parameter estimator. This bound, which depends on the model, its parameters, and the experimental protocol, may actually be too loose in practice, and does not account for the threshold effect described in Kostal et al. (2015). In many cases, as the number of data points increases, the estimate error displays a threshold-like transition, from a region of low performance to a region of high performance where the Cramér-Rao bound is attained. Our definition of practical identifiability also discriminates between regions of low information (for small signal-to-noise ratios and sample size) and high accuracy, provides a quantitative criterion to discriminate them, and can be extended to the case of non-i.i.d. data. An interesting future step would be to verify how the boundaries of our proposed identifiability domains compare with the transition threshold described in Kostal et al. (2015).

An interesting topic would be to study the practical identifiability domain as the number of observations T goes to infinity. In this asymptotic case, practical non-identifiability means that the model cannot be identified, even with an infinite amount of data. We can conjecture that practical identifiability is equivalent to structural identifiability in this asymptotic case, as hinted by **Figure 4**: the identifiability domain increases with T . A future step would be to verify if the practical identifiability domain of a model is included in its structural identifiability domain, and how it behaves when the number of observations T goes to infinity.

We applied our analysis to four variants of the binomial model, of increasing complexity: a Gaussian model (i.e., a uni-quantal distribution); a binomial model without short-term plasticity; a binomial model with only short-term depression; and a binomial model with both short-term depression and facilitation. A future step would be to extend our analysis to further generalizations of the binomial model, in order to account for parameters heterogeneity. Especially, the binomial model assumes that the release probability and the quantal amplitude are identical for each release site. It is however possible to hypothesize that there are several pools of vesicles, each having different parameters (for instance a fast depleting pool and a slow depleting pool). There will be regimes in which those sub-pools can be detected and other in which the noise is too high or the experimental protocol not informative enough to identify them, which can be

quantified using our definition of identifiability. Another possible generalization of the binomial model is to assume that the postsynaptic response to one vesicle release is not fixed, but follows for instance a Gamma distribution (Bhumbra and Beato, 2013) to account for variability in vesicles size and neurotransmitter content.

Model selection is not only a first step in model-based inference of synaptic parameters (as it is necessary to have a reliable estimates of the parameters), but also a tool to study the mechanisms of neurotransmitter release at a chemical synapse. An alternative hypothesis (e.g., “*this synapse shows short-term plasticity*”) can be compared to a null hypothesis (“*this synapse does not show short-term plasticity*”) by computing how well the complex model (i.e. with short-term plasticity) explains the behavior of the synapse compared to the simple model (i.e., without short-term plasticity). Testing models of growing complexity allows to study the nature of the synapse and to identify mechanisms of neurotransmitter release. But the possibility to correctly select the model that corresponds to the true behavior of the synapse will depend on its parameters and on the experimental protocol used to record data: there are regimes in which the specific features of a model do not appear in the data. Such regimes correspond to the identifiability domain of the model, and studying them allows to draw conclusions on the nature of the synapse.

As stated previously, the problem of inferring parameters from noisy and scarce observations is not restricted to synaptic parameters estimation, but is a crucial question in neuroscience. Our proposed methodology could also be applied to models of single neurons (Koch, 2004; Jolivet et al., 2008; Gerstner and Naud, 2009; Mensi et al., 2012), neural population dynamics (René et al., 2020), or calcium-driven vesicles fusion (Schneggenburger and Neher, 2000; Lou et al., 2005; Sun et al., 2007).

On a broader scale, instead of seeing parametric non-identifiability as a statistical problem, we could consider it as a biophysical feature. The total synaptic strength between two cells is a function of both presynaptic (N, p) and postsynaptic (q) parameters. Different combinations of these parameters could lead to the same average postsynaptic response: a presynaptic modification of the number of release sites N can be compensated by an inverse modification of the postsynaptic number of receptors affecting q . This combined effect of presynaptic and postsynaptic plasticity has been shown to enable reliable and flexible learning (Costa et al., 2015) and homeostatic modulation (Davis and Müller, 2015). More generally, the question of degeneracy, defined as the ability of different elements to perform the same function, could be addressed within the framework of identifiability analysis (Drion et al., 2015; Rathour and Narayanan, 2019).

Finally, our proposed definition of model identifiability is paving the way toward Optimal Experiment Design (OED) for model selection and parameter inference. The information conveyed by the data about the ground truth model and its

parameters depends on the experimental protocol: number of recorded data points, stimulation frequency, etc. The goal of OED is to optimize the experimental protocol in order to maximize the possibility to discriminate between competing models (Vanlier et al., 2014; Baliatti et al., 2018) and the precision of the inference of their parameters. An OED for inferring the parameters of a given model maximizes the mutual information between the data and the parameters $\mathcal{I}(\mathcal{D}, \theta)$ (Huan and Marzouk, 2013). This quantity turns out to be equal to the expected gain in information about θ (defined as the Kullback-Leibler divergence between its prior and its posterior), on which our proposed definition of the informative domain (11) is based. Similarly, maximizing the Average Log Bayes Factor (13) is equivalent to maximizing the discriminability between the two models \mathcal{M} and \mathcal{M}' , and hence finding an OED for model selection. As a thorough theoretical preliminary analysis of the properties of the competing models is a first step prior to model selection and parameter inference (Asprey and Macchietto, 2000), we believe that our theoretical contribution to model analysis will contribute to the development of OED techniques for synaptic transmission study.

DATA AVAILABILITY STATEMENT

Matlab and Python files are available in the following GitHub repository: https://github.com/camillegontier/identifiability_binomial.git.

REFERENCES

- Acerbi, L., Ma, W. J., and Vijayakumar, S. (2014). "A framework for testing identifiability of Bayesian models of perception," in *Advances in Neural Information Processing Systems* (Montréal, QC), 1026–1034.
- Asprey, S., and Macchietto, S. (2000). Statistical tools for optimal dynamic model building. *Comput. Chem. Eng.* 24, 1261–1267. doi: 10.1016/S0098-1354(00)00328-8
- Baliatti, S., Klein, B., and Riedl, C. (2018). Fast model-selection through adapting design of experiments maximizing information gain. *arXiv preprint arXiv:1807.07024*.
- Barri, A., Wang, Y., Hansel, D., and Mongillo, G. (2016). Quantifying repetitive transmission at chemical synapses: a generative-model approach. *eNeuro* 3, 1–21. doi: 10.1523/ENEURO.0113-15.2016
- Bekkers, J., Richerson, G., and Stevens, C. (1990). Origin of variability in quantal size in cultured hippocampal neurons and hippocampal slices. *Proc. Natl. Acad. Sci. U.S.A.* 87, 5359–5362. doi: 10.1073/pnas.87.14.5359
- Bhumbra, G. S., and Beato, M. (2013). Reliable evaluation of the quantal determinants of synaptic efficacy using Bayesian analysis. *J. Neurophysiol.* 109, 603–620. doi: 10.1152/jn.00528.2012
- Bird, A. D., Wall, M. J., and Richardson, M. J. (2016). Bayesian inference of synaptic quantal parameters from correlated vesicle release. *Front. Comput. Neurosci.* 10:116. doi: 10.3389/fncom.2016.00116
- Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.* 33, 261–304. doi: 10.1177/0049124104268644
- Bykowska, O. S., Gontier, C., Sax, A.-L., Jia, D. W., Llera-Montero, M., Bird, A. D., et al. (2019). Model-based inference of synaptic transmission. *Front. Synapt. Neurosci.* 11:21. doi: 10.3389/fnsyn.2019.00021
- Calabresi, P., Picconi, B., Parnetti, L., and Di Filippo, M. (2006). A convergent model for cognitive dysfunctions in Parkinson's disease: the

AUTHOR CONTRIBUTIONS

CG and J-PP wrote the paper and derived the equations. CG performed the numerical simulations, made the figures, and wrote the MATLAB code and the Python code. J-PP conceived of the presented idea and supervised the project. All authors contributed to the article and approved the submitted version.

FUNDING

The work presented in this paper was supported by the Swiss National Science Foundation under grant number 31003A_175644 entitled Bayesian Synapse.

ACKNOWLEDGMENTS

We thank Simone Carlo Surace and Mihai A. Petrovici for the fruitful discussions. Figures were created using BioRender. Calculations were performed on UBELIX (<http://www.id.unibe.ch/hpc>), the HPC cluster at the University of Bern.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2020.558477/full#supplementary-material>

- critical dopamine-acetylcholine synaptic balance. *Lancet Neurol.* 5, 974–983. doi: 10.1016/S1474-4422(06)70600-7
- Costa, R. P., Froemke, R. C., Sjöström, P. J., and van Rossum, M. C. (2015). Unified pre- and postsynaptic long-term plasticity enables reliable and flexible learning. *Elife* 4:e09457. doi: 10.7554/eLife.11988
- Costa, R. P., Mizusaki, B. E., Sjöström, P. J., and van Rossum, M. C. (2017a). Functional consequences of pre- and postsynaptic expression of synaptic plasticity. *Philos. Trans. R. Soc. B Biol. Sci.* 372:20160153. doi: 10.1098/rstb.2016.0153
- Costa, R. P., Padamsey, Z., D'Amour, J. A., Emptage, N. J., Froemke, R. C., and Vogels, T. P. (2017b). Synaptic transmission optimization predicts expression loci of long-term plasticity. *Neuron* 96, 177–189. doi: 10.1016/j.neuron.2017.09.021
- Costa, R. P., Sjöström, P. J., and Van Rossum, M. C. (2013). Probabilistic inference of short-term synaptic plasticity in neocortical microcircuits. *Front. Comput. Neurosci.* 7:75. doi: 10.3389/fncom.2013.00075
- Davis, G. W., and Müller, M. (2015). Homeostatic control of presynaptic neurotransmitter release. *Annu. Rev. Physiol.* 77, 251–270. doi: 10.1146/annurev-physiol-021014-071740
- Daw, N. D. (2011). "Trial-by-trial data analysis using computational models," in *Decision Making, Affect, and Learning: Attention and Performance XXIII*, eds M. R. Delgado, E. A. Phelps, and T. W. Robbins (Oxford: Oxford University Press), 3–38. doi: 10.1093/acprof:oso/9780199600434.003.0001
- Del Castillo, J., and Katz, B. (1954). Quantal components of the end-plate potential. *J. Physiol.* 124, 560–573. doi: 10.1113/jphysiol.1954.sp005129
- Drion, G., O'Leary, T., and Marder, E. (2015). Ion channel degeneracy enables robust and tunable neuronal firing rates. *Proc. Natl. Acad. Sci. U.S.A.* 112, E5361–E5370. doi: 10.1073/pnas.1516400112
- Edwards, M. C., and MacCallum, R. C. (2012). *Current Topics in the Theory and Application of Latent Variable Models*. New York, NY: Routledge. doi: 10.4324/9780203813409

- Fatt, P., and Katz, B. (1952). Spontaneous subthreshold activity at motor nerve endings. *J. Physiol.* 117, 109–128.
- Gerstner, W., and Naud, R. (2009). How good are neuron models? *Science* 326, 379–380. doi: 10.1126/science.1181936
- Ghanbari, A., Malyshev, A., Volgushev, M., and Stevenson, I. H. (2017). Estimating short-term synaptic plasticity from pre- and postsynaptic spiking. *PLoS Comput. Biol.* 13:e1005738. doi: 10.1371/journal.pcbi.1005738
- Ghanbari, A., Ren, N., Keine, C., Stoelzel, C., Englitz, B., Swadlow, H. A., et al. (2020). Modeling the short-term dynamics of in vivo excitatory spike transmission. *J. Neurosci.* 40, 4185–4202. doi: 10.1523/jneurosci.1482-19.2020
- Gottman, J. M. (1995). *The Analysis of Change*. Mahwah, NJ: Psychology Press.
- Guo, D., and Li, C. (2012). Stochastic resonance in Hodgkin-Huxley neuron induced by unreliable synaptic transmission. *J. Theor. Biol.* 308, 105–114. doi: 10.1016/j.jtbi.2012.05.034
- Hines, K. E., Middendorf, T. R., and Aldrich, R. W. (2014). Determination of parameter identifiability in nonlinear biophysical models: a Bayesian approach. *J. Gen. Physiol.* 143, 401–416. doi: 10.1085/jgp.201311116
- Huan, X., and Marzouk, Y. M. (2013). Simulation-based optimal Bayesian experimental design for nonlinear systems. *J. Comput. Phys.* 232, 288–317. doi: 10.1016/j.jcp.2012.08.013
- Jefferys, W. H., and Berger, J. O. (1991). *Sharpening Ockham's Razor on a Bayesian Strop*. Technical Report, Department of Statistics, Purdue University.
- Jolivet, R., Kobayashi, R., Rauch, A., Naud, R., Shinomoto, S., and Gerstner, W. (2008). A benchmark test for a quantitative assessment of simple neuron models. *J. Neurosci. Methods* 169, 417–424. doi: 10.1016/j.jneumeth.2007.11.006
- Kass, R. E., and Raftery, A. E. (1995). Bayes factors. *J. Am. Stat. Assoc.* 90, 773–795. doi: 10.1080/01621459.1995.10476572
- Katz, B. (1969). *The Release of Neural Transmitter Substances*. Liverpool: Liverpool University Press.
- Kauer, J. A., and Malenka, R. C. (2007). Synaptic plasticity and addiction. *Nat. Rev. Neurosci.* 8, 844–858. doi: 10.1038/nrn2234
- Kobayashi, R., Tsubo, Y., and Shinomoto, S. (2009). Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold. *Front. Comput. Neurosci.* 3:9. doi: 10.3389/fncom.2009.10.009.2009
- Koch, C. (2004). *Biophysics of Computation: Information Processing in Single Neurons*. Oxford: Oxford University Press.
- Komorowski, M., Costa, M. J., Rand, D. A., and Stumpf, M. P. (2011). Sensitivity, robustness, and identifiability in stochastic chemical kinetics models. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8645–8650. doi: 10.1073/pnas.1015814108
- Kostal, L., Lansky, P., and Pilarski, S. (2015). Performance breakdown in optimal stimulus decoding. *J. Neural Eng.* 12:036012. doi: 10.1088/1741-2560/12/3/036012
- Koyama, S. (2012). “Coding efficiency and detectability of rate fluctuations with non-poisson neuronal firing,” in *Advances in Neural Information Processing Systems* (Stateline), 180–188.
- Levenstein, D., Alvarez, V. A., Amarasingham, A., Azab, H., Gerkin, R. C., Hasenstaub, A., et al. (2020). On the role of theory and modeling in neuroscience. *arXiv preprint arXiv:2003.13825*.
- Levy, W. B., and Baxter, R. A. (2002). Energy-efficient neuronal computation via quantal synaptic failures. *J. Neurosci.* 22, 4746–4755. doi: 10.1523/JNEUROSCI.22-11-04746.2002
- Lou, X., Scheuss, V., and Schneggenburger, R. (2005). Allosteric modulation of the presynaptic Ca²⁺ sensor for vesicle fusion. *Nature* 435, 497–501. doi: 10.1038/nature03568
- MacKay, D. J., and MacKay, D. J. (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge: Cambridge University Press.
- Massonis, G., and Villaverde, A. F. (2020). Finding and breaking lie symmetries: implications for structural identifiability and observability in biological modelling. *Symmetry* 12:469. doi: 10.3390/sym12030469
- Melanson, A., Mejias, J. F., Jun, J. J., Maler, L., and Longtin, A. (2014). A phenomenological model for self-initiated movement in electric fish. *BMC Neurosci.* 15:P112. doi: 10.1186/1471-2202-15-S1-P112
- Mensi, S., Naud, R., Pozzorini, C., Avermann, M., Petersen, C. C., and Gerstner, W. (2012). Parameter extraction and classification of three cortical neuron types reveals two distinct adaptation mechanisms. *J. Neurophysiol.* 107, 1756–1775. doi: 10.1152/jn.00408.2011
- Mootoovaloo, A., Bassett, B. A., and Kunz, M. (2016). Bayes factors via savage-dickey supermodels. *arXiv preprint arXiv:1609.02186*.
- Murphy, S. A., and Van der Vaart, A. W. (2000). On profile likelihood. *J. Am. Stat. Assoc.* 95, 449–465. doi: 10.1080/01621459.2000.10474219
- Navarro, D. J., Pitt, M. A., and Myung, I. J. (2004). Assessing the distinguishability of models and the informativeness of data. *Cogn. Psychol.* 49, 47–84. doi: 10.1016/j.cogpsych.2003.11.001
- Pfeiffer, B. E., and Huber, K. M. (2009). The state of synapses in fragile X syndrome. *Neuroscientist* 15, 549–567. doi: 10.1177/1073858409333075
- Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E., et al. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454, 995–999. doi: 10.1038/nature07140
- Rathour, R. K., and Narayanan, R. (2019). Degeneracy in hippocampal physiology and plasticity. *Hippocampus* 29, 980–1022. doi: 10.1002/hipo.23139
- Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., et al. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* 25, 1923–1929. doi: 10.1093/bioinformatics/btp358
- Raue, A., Kreutz, C., Maiwald, T., Klingmüller, U., and Timmer, J. (2011). Addressing parameter identifiability by model-based experimentation. *IET Syst. Biol.* 5, 120–130. doi: 10.1049/iet-syb.2010.0061
- René, A., Longtin, A., and Macke, J. H. (2020). Inference of a mesoscopic population model from population spike trains. *Neural Comput.* 1–51. doi: 10.1162/neco_a_01292
- Schneggenburger, R., and Neher, E. (2000). Intracellular calcium dependence of transmitter release rates at a fast central synapse. *Nature* 406, 889–893. doi: 10.1038/35022702
- Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136
- Selkoe, D. J. (2002). Alzheimer's disease is a synaptic failure. *Science* 298, 789–791. doi: 10.1126/science.1074069
- Spall, J. C. (2005). Monte Carlo computation of the fisher information matrix in nonstandard settings. *J. Comput. Graph. Stat.* 14, 889–909. doi: 10.1198/106186005X78800
- Stephan, K. E., Baldeweg, T., and Friston, K. J. (2006). Synaptic plasticity and dysconnection in schizophrenia. *Biol. Psychiatry* 59, 929–939. doi: 10.1016/j.biopsych.2005.10.005
- Südhof, T. C. (2008). Neurotrophins and neurexins link synaptic function to cognitive disease. *Nature* 455, 903–911. doi: 10.1038/nature07456
- Sun, J., Pang, Z. P., Qin, D., Fahim, A. T., Adachi, R., and Südhof, T. C. (2007). A dual-Ca²⁺-sensor model for neurotransmitter release in a central synapse. *Nature* 450, 676–682. doi: 10.1038/nature06308
- Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835. doi: 10.1162/089976698300017502
- Van Spronsen, M., and Hoogenraad, C. C. (2010). Synapse pathology in psychiatric and neurologic disease. *Curr. Neurol. Neurosci. Rep.* 10, 207–214. doi: 10.1007/s11910-010-0104-8
- Van Trees, H. L. (2004). *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. New York, NY: John Wiley & Sons.
- Van Trees, H. L., and Bell, K. L. (2007). Bayesian bounds for parameter estimation and nonlinear filtering/tracking. *AMC* 10:12. doi: 10.1109/9780470544198
- Vanlier, J., Tiemann, C. A., Hilbers, P. A., and van Riel, N. A. (2014). Optimal experiment design for model selection in biochemical networks. *BMC Syst. Biol.* 8:20. doi: 10.1186/1752-0509-8-20
- Venzon, D., and Moolgavkar, S. (1988). A method for computing profile-likelihood-based confidence intervals. *J. R. Stat. Soc.* 37, 87–94. doi: 10.2307/2347496
- Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., and Grasman, R. (2010). Bayesian hypothesis testing for psychologists: a tutorial on the savage-dickey method. *Cogn. Psychol.* 60, 158–189. doi: 10.1016/j.cogpsych.2009.12.001
- Wang, T., Yin, L., Zou, X., Shu, Y., Rasch, M. J., and Wu, S. (2016). A phenomenological synapse model for asynchronous neurotransmitter release. *Front. Comput. Neurosci.* 9:153. doi: 10.3389/fncom.2015.00153

- Weinberg, M. D. (2012). Computing the Bayes factor from a Markov chain Monte Carlo simulation of the posterior distribution. *Bayesian Anal.* 7, 737–770. doi: 10.1214/12-BA725
- Welch, J. M., Lu, J., Rodriguiz, R. M., Trotta, N. C., Peca, J., Ding, J.-D., et al. (2007). Cortico-striatal synaptic defects and OCD-like behaviours in SAPAP3-mutant mice. *Nature* 448, 894–900. doi: 10.1038/nature06104
- Wentzel, C., Delvendahl, I., Sydlik, S., Georgiev, O., and Müller, M. (2018). Dysbindin links presynaptic proteasome function to homeostatic recruitment of low release probability vesicles. *Nat. Commun.* 9, 1–16. doi: 10.1038/s41467-017-02494-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Gontier and Pfister. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.