Behavioral/Cognitive

The Synaptic Properties of Cells Define the Hallmarks of Interval Timing in a Recurrent Neural Network

Oswaldo Pérez and Hugo Merchant

Institute of Neurobiology, Universidad Nacional Autónoma de México, Campus Juriquilla, Department of Cognitive Neuroscience, Queretaro 76230, México

Extensive research has described two key features of interval timing. The bias property is associated with accuracy and implies that time is overestimated for short intervals and underestimated for long intervals. The scalar property is linked to precision and states that the variability of interval estimates increases as a function of interval duration. The neural mechanisms behind these properties are not well understood. Here we implemented a recurrent neural network that mimics a cortical ensemble and includes cells that show paired-pulse facilitation and slow inhibitory synaptic currents. The network produces interval selective responses and reproduces both bias and scalar properties when a Bayesian decoder reads its activity. Notably, the interval-selectivity, timing accuracy, and precision of the network showed complex changes as a function of the decay time constants of the modeled synaptic properties and the level of background activity of the cells. These findings suggest that physiological values of the time constants for paired-pulse facilitation and GABAb, as well as the internal state of the network, determine the bias and scalar properties of interval timing.

Key words: interval timing; long-term synaptic properties; recurrent neural network; spontaneous activity

---

Significance Statement

Timing is a fundamental element of complex behavior, including music and language. Temporal processing in a wide variety of contexts shows two primary features: time estimates exhibit a shift toward the mean (the bias property) and are more variable for longer intervals (the scalar property). We implemented a recurrent neural network that includes long-lasting synaptic currents, which cannot only produce interval-selective responses but also follow the bias and scalar properties. Interestingly, only physiological values of the time constants for paired-pulse facilitation and GABAb, as well as intermediate background activity within the network can reproduce the two key features of interval timing.

Introduction

Two critical independent variables are used to characterize behavior in a wide range of timing tasks: constant error and temporal variability. Constant error is defined as estimated duration minus target duration; hence, it is the parameter that corresponds to the accuracy of the estimate (how close is the subjective interval to the actual interval; Woodrow, 1934). In contrast, temporal variability corre-

sponds to the precision of the subjective estimate, namely, how reproducible are the estimates for a specific duration (Gibbon et al., 1997). In computational terms, bias is the constant error and SD is the temporal variability (Jazayeri and Shadlen, 2010). A myriad of psychophysical studies across timing tasks and species have shown a systematic shift in constant error across durations, with an overestimation and underestimation of intervals for shorter and longer intervals in a global set of durations, and with an indifference interval that corresponds to the intermediate duration associated with constant error equal to zero. We called this the bias property (Woodrow, 1934; Jones and McAuley, 2005; Gu et al., 2011). In addition, temporal variability (i.e., SD) increases linearly as a function of duration. This phenomenon has been called the scalar property of interval timing and is a form of Weber's law (Gibbon et al., 1997; Merchant et al., 2008b; Mendez et al., 2011; García-Garibay et al., 2016). These two parameters are critical for describing temporal processing, yet few modeling studies have addressed the question of how neural networks process temporal information in give rise to bias and scalar properties.

Previous modeling work has shown that interval timing can arise from changes in the network state. One of the critical aspects

Figure 1. Neural-network configuration. The network consisted of 800 excitatory and 200 inhibitory integrate-and-fire neurons. **A**, Each cell in the network received a sequence of four stimuli with a constant interstimulus duration d_s, defining three serial-order elements So (red, green, and blue lines). **B**, The stimuli modulated the activity of each cell in the network through two input induced currents: the AMPA with paired-pulse facilitation (associated with an increase in current for consecutive action potentials) and the GABAb (long-term hyperpolarization). Dashed line corresponds to the induced current level after the second stimulus, as a reference. **C**, The recurrent network activity was driven by two recurrent network currents (AMPA and GABAa). **D**, The excitatory cells of the network showed an output activity $r_i(d_s)$, which was 1 if the neuron generated an action potential in response to an input stimulus d_s, or 0 otherwise. Dashed lines correspond to the reset voltage (V_r) and the threshold voltage (V_T).

of these state-dependent models is the fact that the simulated neurons develop interval-selectivity (Laje and Buonomano, 2013). This selectivity is due to the delicate balance between excitation and inhibition inputs defined by paired-pulse synaptic excitatory plasticity and slow synaptic currents (i.e., paired-pulse facilitation and GABAb, respectively; Buonomano and Merzenich, 1998; Buonomano, 2000; Karmarkar and Buonomano, 2007). The notion of interval-selectivity is similar to interval tuning, where cells present an orderly change in activity reaching a peak in the preferred interval. Recently, it has been shown that cells in medial premotor areas and putamen are tuned to the duration of produced intervals (Merchant et al., 2013b, 2015; Bartolo et al., 2014; Crowe et al., 2014). The task performance of monkeys in these neurophysiological studies followed the bias and scalar properties of interval timing (Zarco et al., 2009), suggesting a relation between interval tuning and the two hallmark properties of temporal processing. Now, both paired-pulse facilitation (Varela et al., 1997; Markram et al., 1998) and GABAb (Benardo, 1994; Wang et al., 2010) are time-dependent properties that extend from hundreds of milliseconds to a second and can be modified by experience (Schulz et al., 1994; Misgeld et al., 1995). Yet, little is known about the role of the time constants of these neuronal synaptic properties in interval-selectivity and the two hallmarks of timing. Here, we provide a formal framework on how constant error and temporal variability depend on interval-selectivity, background activity, and synaptic properties of neurons in recurrent neural networks. A Bayesian decoding method was used to optimally read the neural network activity and estimate the constant error and the temporal variability across durations. We found that physiological values of the time constants for paired-pulse facilitation and GABAb can determine the bias and scalar properties.

Materials and Methods

Network model. We simulated a network of 800 excitatory and 200 inhibitory integrate-and-fire neurons that were sparsely ($p = 0.05$; Buonomano, 2000) and randomly connected (Fig. 1). Each neuron (Amit and Brunel, 1997a; Brunel and van Rossum, 2007) was characterized by a membrane potential V that obeyed the equation:

$$\tau \frac{dV}{dt} = -V + I_{Fac} - I_{GABAb} + I_{AMPA} - I_{GABAa} + N \quad (1)$$

where τ is the time constant of the neural membrane, which was $\tau = 10$ and $\tau = 5$ for excitatory and inhibitory neurons, respectively. When V reached the threshold value of $V_T = 20$ mV, an action potential was triggered, which was followed by a membrane potential of $V_r = 0$ mV during a refractory period of $t_r = 1$ ms. Both I_{Fac}, I_{GABAb} were the input induced currents that provided information about the interval duration

and sequence order to each neuron (Fig. 1B). I_{Fac} was an excitatory current that included paired-pulse facilitation, whereas I_{GABAb} was a slow inhibitory current. The internal dynamics of the recurrent network were driven by I_{AMPA} and I_{GABAa}, corresponding to fast excitation and inhibition, respectively (Amit and Brunel, 1997a; Fig. 1C). Hence, the facilitation and GABAb plasticity occurred at input and not at recurrent synapses. Finally, $N(t)$ is white noise with a SD σ_N and zero mean.

Input stimuli. Input I_{Fac}, I_{GABAb} currents were activated by four pulses separated by a duration d_s, generating an isochronous sequence with three serial-order (So) elements. d_s covered a range from 100 to 1500 ms, on steps of 50 ms (29 values; Figs. 1A, 3B). This input mimics the sensory metronome used to drive the motor response of subjects during a metronome synchronization task (Merchant et al., 2011). Because the recurrent network properties were time-varying and depended on stimulus history (i.e., depended on d_s, the neuronal responses to the first pulse were not included in the analysis. Consequently, the responses to the stimulus across the serial-orders $So1$, $So2$, and $So3$ corresponded to the activity linked to the second, third, and fourth input stimuli, respectively. Last, $r_i(d_s)$ was the response of neuron i to the input stimulus d_s, and had a value of 1 if the neuron generated at least one action potential within 20 ms after each input pulse, or 0 otherwise (Fig. 1D). Overall, this is a neural network that processed temporal sensory information and produced interval selective responses.

Input currents. Four input currents (I_s) changed the membrane potential of each cell in the neural network. The temporal dynamics of each current consisted of two coupled linear differential equations (Destexhe et al., 1994):

$$\tau_{r,S} \frac{dI_S}{dt} = -I_S + R_S \quad (2)$$

$$\tau_{d,S} \frac{dR_S}{dt} = -R_S + \sum_{i \in k, S} W_{i,S} \delta(t - t_{k,S}), \quad (3)$$

where $\tau_{r,S}$ and $\tau_{d,S}$ are the rise and decay time constants, and R_S is an auxiliary variable. $w_{i,S}$ is the synaptic weight that determines the efficacy of synaptic transmission coming from neuron i. $t_{k,S}$ is the time neuron i emitted spike number k. We used two input driving currents corresponding to the paired-pulse facilitation and *GABAb*, and two recurrent network currents: *AMPA* and *GABAa*. The time constants were as follows: $\tau_{r,AMPA} = \tau_{r,GABAa} = 0.2ms$, $\tau_{d,AMPA} = \tau_{d,GABAa} = 0.7ms$, $\tau_{r,GABAb} = 21ms$.

Additionally, the values of $w_{i,AMPA}$ and $w_{i,GABAa}$ maintained the recurrent network response stable. These internal recurrent currents (AMPA and GABAa) had a smaller effect on interval processing, because the fast currents decayed before duration d_s. However, the balance of AMPA and GABAa was essential for maintaining the stability of the neural network (Brunel, 2000; Vogels et al., 2011). Thus, the GABAa weights ($w_{i,GABAa}$) were fourfold the AMPA weights ($w_{i,AMPA}$) because we used a 4:1 ratio between the number of excitatory and inhibitory neurons. In contrast,
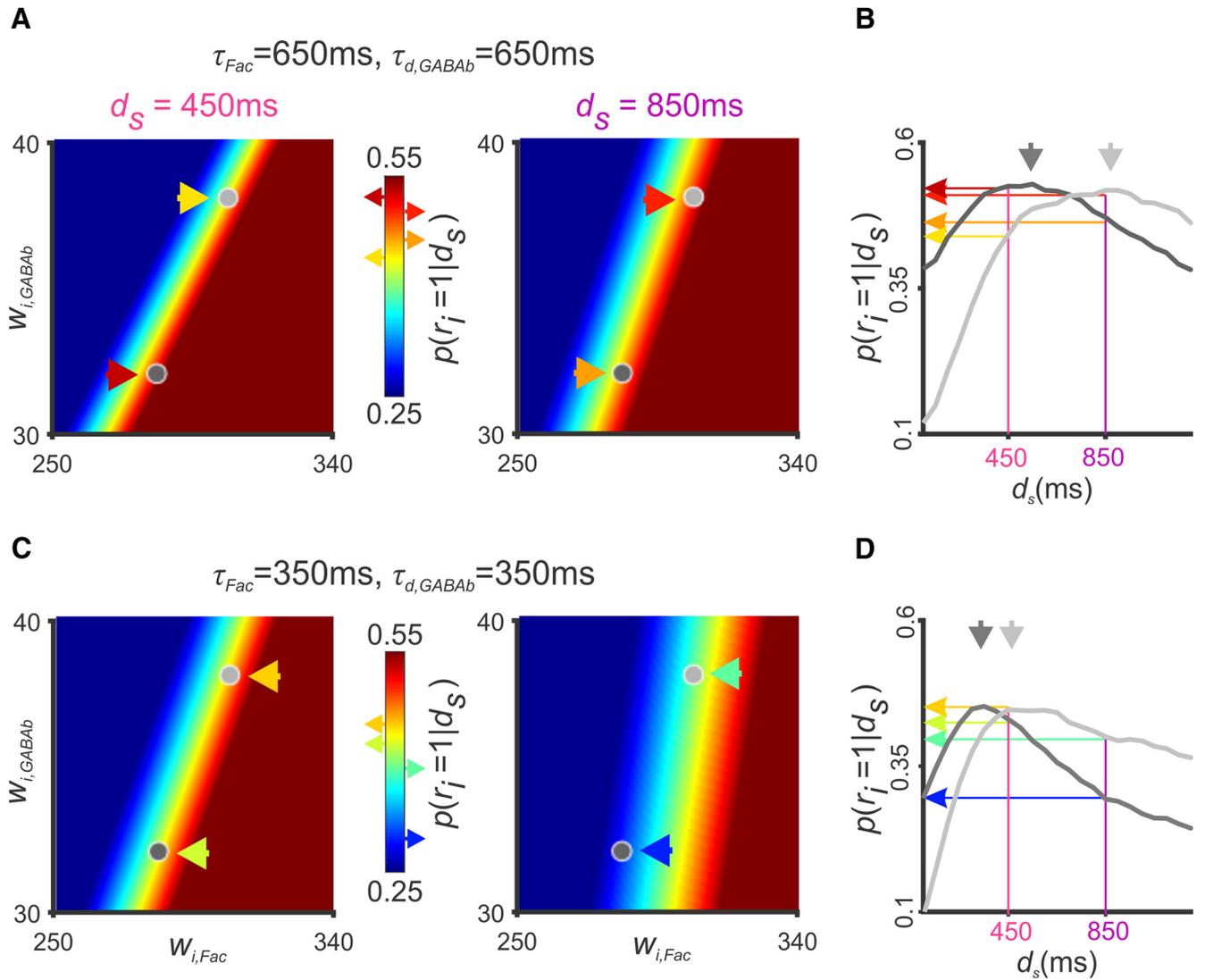
Figure 2. Modulation of the interval-selective responses of the network by the time constants of the synaptic currents. ***A***, The interval-selectivity depended on the balance of weights of the paired-pulse facilitation $w_{i,Fac}$ and the slow GABAb $w_{i,GABAb}$. The color gradient depicts the probability of a neuronal response to the d_s durations of 450 ms (left, magenta) and 850 ms (right, purple) during *So2* as a function of $w_{i,Fac}$ (abscissa) and $w_{i,GABAb}$ (ordinate), using the decay time constants of $\tau_{Fac} = 650$ ms and $\tau_{d,GABAb} = 650$ ms. ***B***, The interval-tuning functions of two neurons (dark and light gray functions) in the network that changed their response as a function of d_s. The magenta and purple vertical lines correspond to the durations of 450 and 850 ms, which are associated with a $p(r_i = 1 \mid d_s)$ shown as horizontal arrows with a color code depicted in the color bar in ***A***. Note the preferred duration of the neurons (543 ms dark gray, 871 ms light gray) depends on their balance of weights shown as dots with the corresponding color in the maps of ***A***. ***C***, The interval-selectivity also changed as a function of the time constants of the synaptic currents. At $\tau_{Fac} = 350$ ms $\tau_{d,GABAb} = 350$ ms and the response probability weight maps are shifted to the right with respect to ***A***. ***D***, The interval tuning functions of the two cells in ***B*** were shifted to the left, showing now a preferred duration of 337 ms (dark gray) and 464 ms (light gray).

the driving $w_{i,Fac}$ and $w_{i,GABAb}$ were randomly distributed (Fig. 2 *A*, *C*) and were responsible for the selectivity to interval durations and serial-order within the network activity.

Paired-pulse facilitation. The input current I_{Fac} was an AMPA current where the amplitude increased upon the repeated presentation of stimuli. Paired-pulse facilitation changed over time the synaptic efficacy and reflects the history of presynaptic inputs. Facilitation in actual neurons last hundreds of milliseconds and is caused by calcium influx inside the axon terminal after a presynaptic action potential. Thus, I_{Fac} had a progressive amplification for repetitive presynaptic activity, and it was modeled as an increase in transmitter release probability $p(t)$ (Tsodyks et al., 1998). Between presynaptic action potentials, this probability followed the equation:

$$\tau_{Fac} \frac{dp(t)}{dt} = P_0 - p(t), \qquad (4)$$

where τ_{Fac} was the time constant and $P_0 = 0.17$ was the stable release probability. In this model, $p(t)$ changed its value immediately after

the appearance of an input action potential following the rule $p \rightarrow p + f_P (1 - p)$ ($f_P = 0.62$). Therefore, the synaptic weight $w_{i,Fac}$ was multiplied by factor $\frac{p}{P_0}$ (Varela et al., 1997; Dayan and Abbott, 2001). Notably, $\tau_{d,GABAb}$ and τ_{Fac} were the key independent parameters of our analysis.

Neural encoding. Principal component analysis (PCA) was used to reduce the high dimensional activity of the recurrent network activity (Jolliffe, 2002; Figure 3*A*). Thus, instead of using 800 excitatory time-varying responses, we used the two principal components as the network responses $\mathbf{r}_{PCA} = (r_{PCA1}, r_{PCA2})$. The two principal components explained >40% of the variability across simulations. The PCA was performed on the covariance matrix (800 × 800) of the excitatory activity of the network using the 2900 observations that corresponded to 100 simulations for each of the 29 d_s. A bivariate normal distribution was fitted to $\mathbf{r}_{PCA}$ to calculate $p(\mathbf{r}_{PCA} \mid d_s)$ which is the probability of evoking response $\mathbf{r}_{PCA}$ given the stimulus d_s (Fig. 3*B*, ellipses). Figure 3*C* shows the gener-
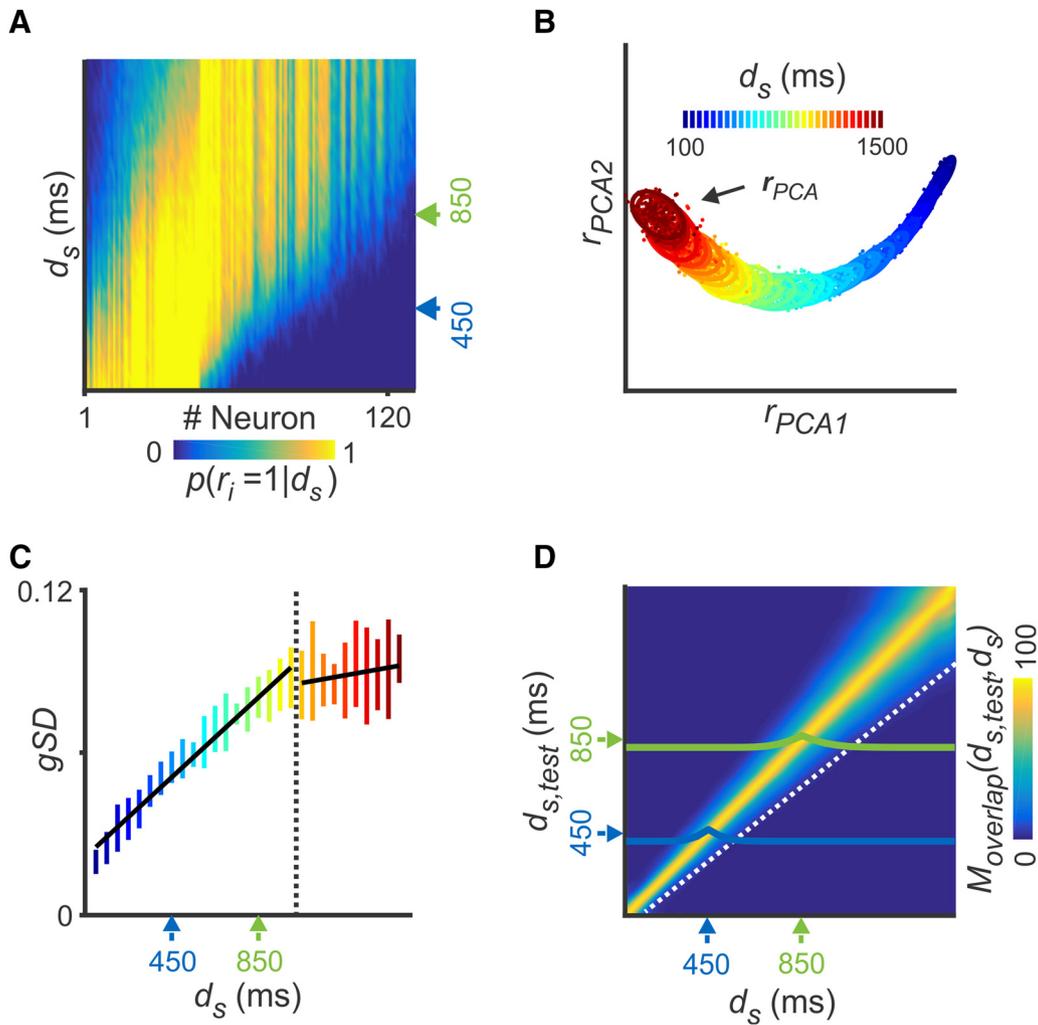
Figure 3. Encoding duration in the neural network. **A**, Excitatory neurons in the network that were selective to duration [minimum MI between neural response $r_i(d_s)$ and d_s was 0.1 bits, from a range between 0 and 1]. Each column is the probability of response induced by stimulus d_s in the second serial order, and the ordinate corresponds to the durations d_s, ranging from 100 to 1500 ms in steps of 50 ms. Neurons that showed interval-selectivity were sorted as a function of their average response to durations, covering a wide range of preferred durations in the population. **B**, PCA was used to determine the encoding of d_s in the network response, which corresponds to $\mathbf{r}_{PCA} = (\mathbf{r}_{PCA1}, \mathbf{r}_{PCA2})$. Each dot represents a simulation of the network given a stimulus duration d_s (color code). A bivariate normal distribution was fitted for each d_s to compute $p(\mathbf{r}_{PCA}|d_s)$ (ellipses include 90% of distributions). Note the strong interval-selectivity in network activity. **C**, Generalized SD (squared root of the determinant of the covariance matrix) was computed for each d_s. Error bars are the SD over 10 different network simulations. Network followed the scalar property <1000 ms and reached a plateau after this value. **D**, Matrix $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ of the percentage of overlap between the $p(\mathbf{r}_{PCA}|d_s)$ ellipses for every $d_{s,test}$ (ordinate) against d_s (abscissa). Note the increase in the width of the $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ as a function of $d_{s,test}$ (distance between the dashed white line and the diagonal).

alized SD [$gSD(d_s)$] for each d_s. The $gSD(d_s)$ is the square root of the determinant of the covariance matrix and is a measure that is directly associated with the area of the ellipse of $p(\mathbf{r}_{PCA}|d_s)$ of Figure 3B. The network activity showed an increase in $gSD(d_s)$ as a function of interval, until d_s reached ~1000 ms (Fig. 3C). After this duration, $gSD(d_s)$ was asymptotic. We also measured the percentage of overlap between the $p(\mathbf{r}_{PCA}|d_s)$ ellipses using the adjacent matrix $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ for every ellipse of $d_{s,test}$ against d_s. The diagonal of $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ corresponds to a perfect overlap (100%), because $d_{s,test}$ and d_s are the same (Fig. 3D). Interestingly, $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ showed an increase in width as a function of $d_{s,test}$ (Fig. 3D, dotted line). Based on previous neurophysiological studies we defined $d_{s,test}$ as 450, 550, 650, 750, or 850 ms (Merchant et al., 2011, 2013b, 2014).

Network decoding. Here, we assumed a normal prior distribution $p(d_s)$ with a mean of 650 ms and SD $\sigma_{prior} = 172$ ms (Fig. 4B). This prior created a clear bias property on the constant error that was difficult to obtain from a uniform prior. Indeed, the mean at 650 ms produced an indifference value close to 650 ms. Bayesian inference uses the posterior probability $p(d_s|\mathbf{r}_{PCA})$ for optimal decoding. This can be estimated using likelihood probability $p(\mathbf{r}_{PCA}|d_s)$ (Fig. 3B), prior distribution $p(d_s)$, and Bayes' theorem:

$$p(d_s|\mathbf{r}_{PCA}) = \frac{p(\mathbf{r}_{PCA}|d_s)p(d_s)}{p(\mathbf{r}_{PCA})}. \tag{5}$$

In this study, we calculated the Bayes least square decoded estimator (d_e), a scalar that was dependent on a particular $\mathbf{r}_{PCA}$. In fact, d_e was computed as the mean of the posterior probability $p(d_s|\mathbf{r}_{PCA})$.

Because d_e was a deterministic function of $\mathbf{r}_{PCA}$ [$d_e = f(\mathbf{r}_{PCA})$], then $p(d_e|\mathbf{r}_{PCA}) = \delta(d_e - f(\mathbf{r}_{PCA}))$, which is a Dirac function. Instead of using the Dirac distribution, which is difficult to calculate numerically, we used a normal distribution with mean $f(\mathbf{r}_{PCA})$ and small SD (5 ms), as follows:

$$p(d_e|\mathbf{r}_{PCA}) = N(d_e, f(\mathbf{r}_{PCA}), 5ms). \tag{6}$$

Finally, we calculated the conditional probability of evoking the duration d_e given that tested duration $d_{s,test}$ was presented (Fig. 4A). This probability depends on population response $\mathbf{r}_{PCA}$; nevertheless, we removed the dependence of the response by marginalization:

$$p(d_e|d_{s,test}) = \iint p(d_e|\mathbf{r}_{PCA})p(\mathbf{r}_{PCA}|d_{s,test})dr_{PCA1}dr_{PCA2}. \tag{7}$$

Then, we substituted Equation 6 on Equation 7 and obtained the relation:

A
Likelihood probability

B
Prior probability

C
Neural Decoding
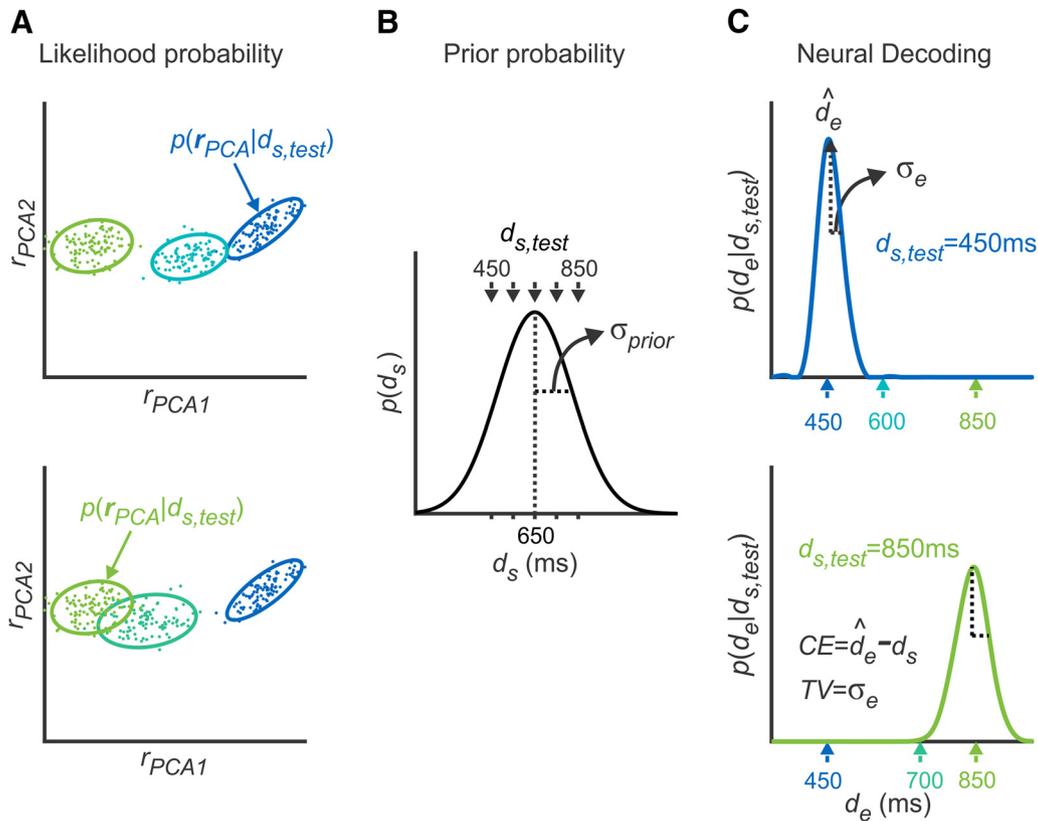
Figure 4. Decoding duration. **A**, Likelihood probabilities $p(r_{PCA}|d_s)$ showed an overlapping that depended on the difference between durations d_s [blue (450 ms) vs light blue (600 ms) or green (850 ms) on top, and green (850 ms) vs light green (700 ms) or blue (450 ms) at the bottom]. **B**, Five test intervals were analyzed ($d_{s,test}$ = 450, 550, 650, 750 and 850 ms; black arrows), thus, a normal prior distribution $p(d_s)$ (mean 650 ms and SDs σ_{prior} = 172 ms) was used to cover the range of tested durations and induce the bias property on the constant error. **C**, The marginal probability $p(d_e|d_{s,test})$ was computed from the encoding $p(r_{PCA}|d_{s,test})$ and decoding $p(d_e|r_{PCA})$. The $p(d_e|r_{PCA})$ was calculated using the Bayes theorem and the prior probability $p(d_e)$. The mean ($\hat{d}_e$) and SD (σ_e) of $p(d_e|d_{s,test})$ were used to calculate the key dependent parameters: constant error $CE = \hat{d}_e - d_{s,test}$ and temporal variability $TV = \sigma_e$. Note that the $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ increased as a function of $d_{s,test}$ (Figs. 4A, 3D); namely, $p(d_e|d_{s,test})$ showed a larger overlap between $d_{s,test}$ = 850 ms against d_s = 700 ms (**A**, bottom, green and light green curves, respectively) than the overlap between the $d_{s,test}$ = 450 ms against d_s = 600 ms (**A**, top, blue and light blue curves, respectively). Accordingly, the distribution $p(d_e|d_{s,test})$ showed a larger width for $d_{s,test}$ = 850 ms (green distribution) than $d_{s,test}$ = 450 ms (blue distribution).

$$p(d_e|d_{s,test}) = \int\int N(d_e, f(r_{PCA}), 5ms)\, p(r_{PCA}|d_{s,test})\, dr_{PCA1} dr_{PCA2}. \tag{8}$$

The decoded mean $\hat{d}_e$ and SD σ_e of the decoder density $p(d_e|d_{s,test})$ were used as the key dependent parameters in the present study. Indeed, the classical psychometric parameters of constant error and temporal variability correspond to the bias $CE(d_{s,test}) = \hat{d}_e - d_{s,test}$ and SD $TV(d_{s,test}) = \sigma_e$, respectively (Fig. 4C).

Mutual information. We computed the mutual information (MI) between the network decoding and the stimuli. It is a measure of the statistical dependency between the behavioral variable, in this case $d_{s,test}$, and a neural response parameter d_e (i.e., the output of the decoder). First, we calculated the joint probability $p(d_e, d_{s,test})$ and $p(d_e)$ using the decoder probability $p(d_e|d_{s,test})$. Then, we calculated MI as follows:

$$MI = \sum_{d_{s,test}} \int_{d_e} p(d_e, d_{s,test}) log_2 \left(\frac{p(d_e, d_{s,test})}{p(d_e)\, p(d_{s,test})} \right). \tag{9}$$

Results
Network model
We simulated a neural network of 800 excitatory and 200 inhibitory integrate-and-fire neurons that were sparsely and randomly connected (Fig. 1; see Materials and Methods). The neural response of each neuron was driven by two induced input currents: paired-pulse AMPA facilitation and slow GABAb inhibition.

Both input currents were time-dependent properties defined by their decay time constants ($\tau_{d,GABAb}$ and τ_{Fac}), which had postsynaptic effects on network neurons lasting hundreds of milliseconds (Fig. 1B). Within the recurrent network, the cells communicated through AMPA with no paired-pulse plasticity and fast GABAa currents (Fig. 1C). The induced input currents were modulated by four input pulses separated by a duration d_s (100 to 1500 ms). This produced an input stimulus sequence with three serial-order elements (So1, So2, So3; Fig. 1A). The neural selectivity to different d_s was determined by the history of input stimuli in the sequence and balance of weights of the paired-pulse facilitation and GABAb (Fig. 2). Hence, this is an unsupervised recurrent network whose time selective properties depend on the stimulus history and the equilibrium between induced slow excitatory and inhibitory currents. In contrast, the responses to the first stimulus were independent of d_s because they were not shaped by slow synaptic currents and paired-pulse synaptic plasticity. Consequently, we focused on cell responses triggered by the last three stimuli in the input sequence. The neural response $r_i(d_s)$ of each neuron i was 1 when at least one action potential was fired within the first 20 ms after each stimulus, and 0 otherwise (Fig. 1D).

Neurons in the network showed differences in their average response across the spectrum of d_s, because the balance between synaptic weights that undergo paired-pulse facilitation and GABAb generates temporal selectivity (Buonomano, 2000). In fact, the

response of excitatory neurons to temporal stimuli changed as a function of these synaptic weights. The color gradients shown in Figure 2A correspond to the probability of a neuronal response during *So2* as a function of $w_{i,Fac}$ (abscissa) and $w_{i,GABAb}$ (ordinate), using the decay time constants of $\tau_{Fac} = 650$ ms and $\tau_{d,GABAb} = 650$ ms. It is evident that the response map is shifted to the right and has a larger slope for a d_s of 850 (Fig. 2A, right, purple) than 450 ms (Fig. 2A, left, magenta). Consequently, the dark gray dots at the bottom of these response maps correspond to a neuron whose combination of weights generate a tuning function with a short-preferred interval, whereas the light gray dots at the top correspond to another neuron with a long-preferred interval (Fig. 2B). The neural network included neurons covering a wide range of $w_{i,Fac}$ and $w_{i,GABAb}$, and hence, showing an extensive range of preferred durations (Fig. 3A,B). These simulations suggest that artificial neural circuits were capable of encoding temporal information as recorded in the medial premotor cortex (Merchant et al., 2013b). In addition, the interval-selectivity also changed as a function of the time constants of the synaptic currents. At $\tau_{Fac} = 350$ ms and $\tau_{d,GABAb} = 350$ ms the response probability weight maps showed a shift to the right and an increase in slope (Fig. 2C). This shift is accompanied by a change in the preferred interval of the neurons in the network. For example, the tuning functions of the two cells in Figure 2B were shifted to the left when the time constants of the long-lasting currents are smaller (Fig. 2D) and in general, the width of the tuning functions of interval selective neurons had the tendency to increase as a function of preferred interval (Fig. 3A). However, this measure was highly variable across the neurons in the network (data not shown).

The wide range of preferred intervals in the network was a basic feature of the model. Nevertheless, our main goal was to compute an estimate of temporal variability and constant error from a decoder density function, which in turn was calculated from the population activity of the network (Fig. 3). Thus, the synaptic underpinnings of the bias and scalar properties were studied using the decoded estimates of temporal variability and constant error across durations.

Bayesian inference was used to compute the decoder density function. First, we characterized the network encoding of duration d_s using the response vector $r = (r_1, r_2, \ldots, r_N)$, where N is the number of excitatory neurons. It has been shown that different behavioral parameters can be encoded in the activity of cell populations (Ma et al., 2006; Naselaris et al., 2006) using a low dimensional representation (Stopfer et al., 2003; Churchland et al., 2012). Here, we found that some of neurons in the network were interval selective; and therefore, we used PCA on the covariance matrix of the excitatory activity of the recurrent network r across all durations (29 durations × 100 simulations; see Materials and Methods) to extract the time-related activity from the simulated population responses. In fact, the average response variability explained by the first two principal components $r_{PCA} = (r_{PCA1}, r_{PCA2})$ was ~40% (Fig. 3B). The encoding relationship between d_s and r_{PCA} was quantified by the conditional probability $p(r_{PCA}|d_s)$ that is the likelihood probability of observing the response r_{PCA} given that duration was d_s. In fact, $p(r_{PCA}|d_s)$ was computed from the bivariate normal distribution of r_{PCA} for each d_s (Figs. 3B, ellipses, 4A). The variability of neural network responses was determined using the generalized SD [$gSD(d_s)$], which is the square root of the determinant of the covariance matrix of the r_{PCA} and, therefore, is a measure that is directly associated with the $p(r_{PCA}|d_s)$ and area of the ellipses of Figure 3B. The network activity showed an increase in $gSD(d_s)$ as a func-

tion of interval, until d_s reached ~1000 ms. After this duration, the $gSD(d_s)$ was asymptotic (Fig. 3C). These results support the notion of different neural timing mechanisms for sub and suprasecond scale (Grondin, 2012, 2014; Méndez et al., 2017), and suggest an important role of paired-pulse facilitation and slow inhibition in the former time range. In addition, we computed the percentage of overlap between the $p(r_{PCA}|d_s)$ ellipses using the adjacent matrix $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ for all the ellipses of $d_{s,test}$ against d_s. The diagonal corresponds to a perfect overlap because $d_{s,test}$ and d_s are the same. Notably, the width of the overlap in $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ increased as a function of $d_{s,test}$ (Fig. 3D, dotted line), with a larger spread for a $d_{s,test} = 850$ ms against surrounding d_s (Fig. 3D, green curve) than for the $d_{s,test} = 450$ ms (Fig. 3D, blue curve).

Then, we performed Bayesian decoding from the population response r_{PCA} to estimate the duration d_e that evoked this response. Classical Bayesian inference depends on posterior probability $p(d_s|r_{PCA})$ (Eq. 5), which is the probability that the interval d_s was presented given that the population response r_{PCA} was recorded. This posterior probability was computed as the product of $p(r_{PCA}|d_s)$ (Figs. 3B, 4A) by $p(d_s)$, which was the prior probability of (d_s), divided by $p(r_{PCA})$, that, in turn, was the probability of the population responses across all d_s. The prior probability reflects the knowledge about the occurrence of previous durations, and several studies have successfully used a Gaussian prior distribution to model motor timing across different tasks in humans (Acerbi et al., 2012; Shi et al., 2013). Thus, we used as prior probability a Gaussian distribution with a mean of 650 ms and a SD of $\sigma_{prior} = 173$ ms (Fig. 4B; see the next section for the validation of this σ_{prior}). Given the Bayesian least-squares, d_e was a scalar that was dependent on a particular r_{PCA} and corresponded to the mean of the posterior probability $p(d_s|r_{PCA})$ (Fig. 4).

From the scalar decoded value d_e, we calculated the decoder density function $p(d_e|d_{s,test})$, which measures the probability of observing an estimated interval d_e given that the stimulus $d_{s,test}$ was presented (see Materials and Methods; Eq. 8; Fig. 4C). We computed the mean $\hat{d}_e$ and SD σ_e of the decoded density function $p(d_e|d_{s,test})$. Importantly, the decoded constant error was calculated as $CE(d_{s,test}) = \hat{d}_e - d_{s,test}$ and the temporal variability as $TV(d_{s,test}) = \sigma_e$ (Fig. 4C).

d_e was a decoded measure of the r_{PCA} linked to a specific $d_{s,test}$; however, some responses r_{PCA} were evoked by adjacent stimuli d_s, corresponding to the overlap of $p(r_{PCA}|d_s)$ between this specific $d_{s,test}$ and the tested d_s. Thus, the width of decoder density [$TV(d_{s,test})$] depended on the $\mathcal{M}_{overlap}(d_{s,test}, d_s)$, because a larger $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ produced a larger uncertainty that a particular r_{PCA} was linked to a $d_{s,test}$. For example, $\mathcal{M}_{overlap}(d_{s,test} = 850, d_s = 700)$ (Fig. 4A, bottom) was larger than $\mathcal{M}_{overlap}(d_{s,test} = 450, d_s = 600)$ (Fig. 4A, top). Therefore, the decoder density of $p(d_e|d_{s,test} = 850)$ (Fig. 4C, green density) was broader than $p(d_e|d_{s,test} = 450)$ (Fig. 4C, blue density).

We systematically measured the effects of the time constants of the slow inhibitory current ($\tau_{d,GABAb}$) and paired-pulse facilitation (τ_{Fac}) on the bias and scalar properties resulting from decoded behavior of the constant error and temporal variability across d_s, respectively.

Hallmark timing properties in the network

Figure 5A shows the decoder density function $p(d_e|d_{s,test})$ for the different $d_{s,test}$ for the serial order So_2. Three key features can be extracted from these data. First, the temporal estimator $\hat{d}_e$ was
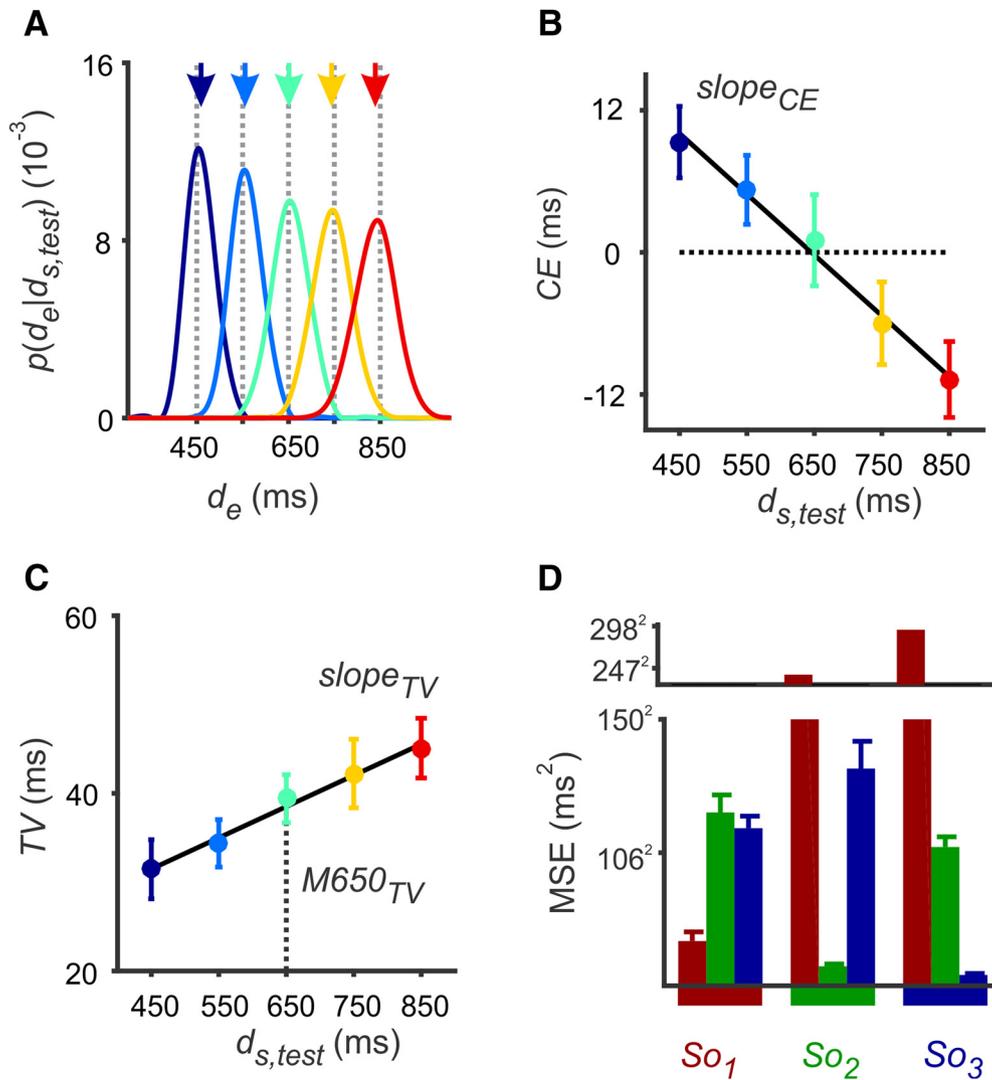
Figure 5. Decoding of time from the network's activity. ***A***, The decoder density function $p(d_e | d_{s,test})$ across $d_{s,test}$ and So_2. Each curve corresponds to the average of 10 different network simulations with a $\tau_{Fac} = 650$ ms and a $\tau_{d,GABAb} = 650$ ms. Note the increase in $p(d_e | d_{s,test})$ variability as a function of $d_{s,test}$. ***B***, The constant error $CE(d_{s,test})$ plotted as a function of duration for the second serial order in the stimulus sequence. The linear regression corresponds to $CE(d_{s,test}) = M650_{CE} + slope_{CE}(d_{s,test} - 650)$. ***C***, Temporal variability $TV(d_{s,test})$ for each $d_{s,test}$ and the linear regression corresponds to $TV(d_{s,test}) = M650_{TV} + slope_{TV}(d_{s,test} - 650)$. Error bars in ***B*** and ***C*** are the SD over 10 different network simulations. ***D***, The neural network also showed a modulation in decoding interval-selectivity as a function of the serial order of input stimuli. The MSE of the estimated duration d_e was computed for nine combinations of serial order responses (red, So_1; green, So_2; blue, So_3) for each r_{PCA} (the 3 bars for each So). The results show that the MSE was minimum when $p(r_{PCA} | d_s)$ and r_{PCA} corresponded to the actual So.

tuned to $d_{s,test}$ because of the balance of weights of the paired-pulsed facilitation and GABAb (Fig. 2), which in turn produced systematic changes in network responses for each stimulus d_s (Fig. 3B). Second, $\hat{d}_e$ was overestimated for the short durations and underestimated for longer intervals (Fig. 5A, arrows vs lines). Indeed, as seen in Figure 5B, there was a monotonic decrease in constant error (negative $slope_{CE}$) as a function of duration, and the indifference interval was observed at 650 ms (with intercept $M650_{CE}$ equal to 0 at 650 ms). Therefore, the network decoding followed the bias property observed in psychophysical studies. Last, the temporal variability followed the scalar property of interval timing, namely, there was an increase in variation as a function of $d_{s,test}$ (Fig. 5C), with a positive $slope_{TV}$ and a positive intercept $M650_{TV}$ at 650 ms.

The neural network showed also a response modulation as a function of the serial order of input stimuli. The mean squared error (MSE) of the decoded serial-order reached a minimum when the likelihood $p(r_{PCA} | d_s)$ and decoder $p(d_e | r_{PCA})$ probabil-

ities corresponded to the same serial order (e.g., the MSE of $So1$ was minimum when $p(r_{PCA} | d_s)$ and $p(d_e | r_{PCA})$ came from $So1$ [right, red] and increased for $So2$ [right, green], and $So3$ [right, blue]; Fig. 5D). Nevertheless, the bias and scalar properties were similar across So's. Hence, in the remaining of the paper we focused our analysis on the data from $So2$.

We determined the σ_{prior} of the Gaussian prior distribution that produced the largest mutual information of duration and that optimally generated the scalar and bias properties. Figure 6 shows that a $\sigma_{prior} = 172$ ms (green dotted line) was associated with the largest mutual information (Fig. 6D), and produced a decoded constant error and temporal variability across $d_{s,test}$ that robustly followed the bias and scalar properties (Fig. 6B,C, respectively). Indeed, from the linear fits of Figure 6B on the constant error as a function of $d_{s,test}$, we extracted the slope ($slope_{CE}$) and the intercept at 650 ms ($M650_{CE}$) across σ_{priors} to build Figure 6E, which defines four quadrants. This Figure demonstrates that an increase in σ_{prior} produced both an increase of $slope_{CE}$ and a
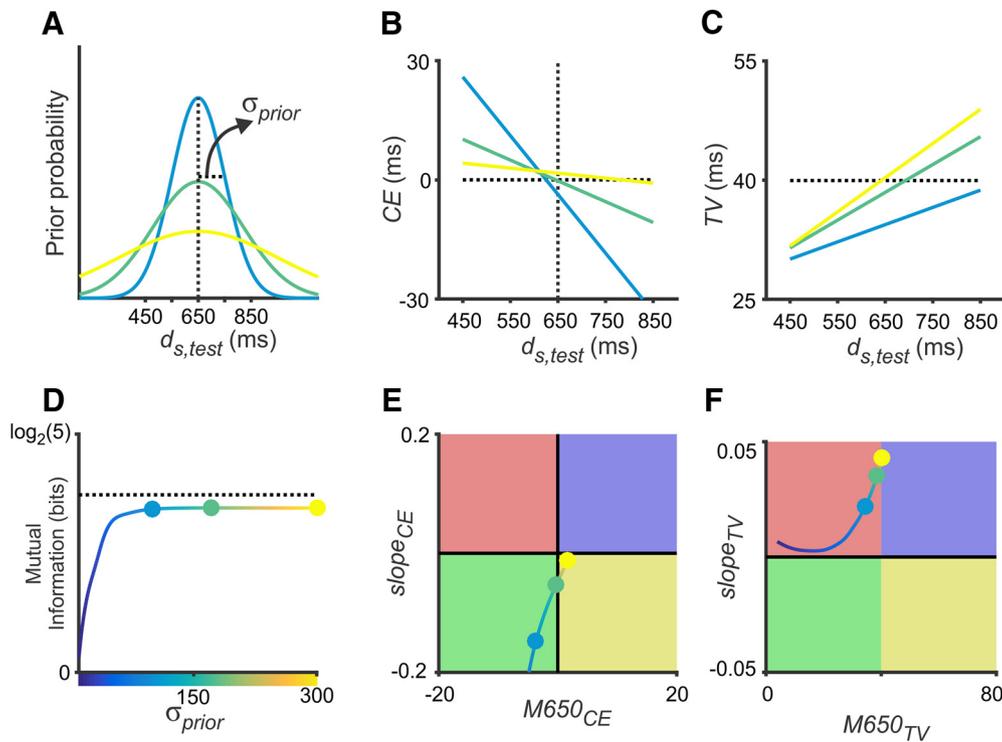
Figure 6. The effects of the Prior probability. **A**, The decoded interval depended on the prior distribution. $p(d_e)$ was normal distribution with a mean 650 ms and SDs σ_{prior} (**D**, bottom, color code). **B**, Constant error (CE) decreased and (**C**) temporal variability (TV) increased as a function of $d_{s,test}$, but the rate of decrease or increase depended on σ_{prior}. **D**, The mutual information between d_e and $d_{s,test}$ increased as a function of σ_{prior} and had a maximum at a value of 172 ms (green dot). Dashed line corresponds to mutual information between network response r_{PCA} and $d_{s,test}$. **E**, Slope ($slope_{CE}$) and constant ($M650_{CE}$) for the linear regression of CE as a function of d_s ($CE(d_{s,test}) = M650_{CE} + slope_{CE}(d_{s,test} - 650)$). Each color dot corresponds to different σ_{prior} value in **D**. Four sectors on the $slope_{CE}$-$M650_{CE}$ plane were defined using the combinations of positive or negative values for $slope_{CE}$ and $M650_{CE}$. **F**, Slope ($slope_{TV}$) and constant ($M650_{TV}$) for the linear regression of temporal variability [$TV(d_{s,test}) = M650_{TV} + slope_{TV}(d_{s,test} - 650)$]. The same four sectors in **E** were defined using $M650_{TV}$ and $slope_{TV}$.

drift of $M650_{CE}$ from negative to positive values, acquiring a value of zero (which is the indifference interval) close to $\sigma_{prior} = 172$ ms (green point). With this σ_{prior} the magnitude of the constant error and temporal variability as a function $d_{s,test}$ is similar to what we found empirically in previous studies (Zarco et al., 2009; Donnet et al., 2014). Note that a small σ_{prior} did not produce a perfect alignment of constant error $\sim$650 ms, suggesting that not only normal prior distribution but also the neural network population activity determined the bias alignment as a function of d_s. Similarly, Figure 6F shows the four quadrants resulting from plotting the $slope_{TV}$ against the intercept $M650_{TV}$ at 650 ms. In this case, the increase in σ_{prior} produced an increase in both the $slope_{TV}$ of the scalar property and the intercept $M650_{TV}$. It is important to notice that the σ_{prior} in the Bayesian decoding modified the slope of the scalar property (Fig. 6C).

Overall the scalar property depended on both the likelihood probability that was associated with the representation of duration in the activity of the neural network and the prior probability of the decoding. However, two features of the likelihood probability had a substantial impact on the decoded temporal variability $TV(d_{s,test})$: the generalized SD $gSD(d_{s,test})$ (Fig. 3C, Pearson correlation: TV vs gSD, $r = 0.23$, $p < 0.0001$) and, more importantly, the percentage of overlap between the $p(r_{PCA}|d_s)$ ellipses, corresponding to $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ [Figs. 3D, 4A,C; Pearson correlation: TV vs $\int_{d_s} \mathcal{M}_{overlap}(d_{s,test}, d_s)$, $r = 0.95$, $p < 0.0001$].

Synaptic effects

The properties of the decoder density function $p(d_e|d_{s,test})$ depend on the time-varying input synaptic currents. Hence, we changed the temporal dynamics of I_{Fac} and I_{GABAb} so that their decay time

constants ($\tau_{d,GABAb}$ and τ_{Fac}) varied from 50 to 1000 ms (Benardo, 1994; Markram et al., 1998). We investigated how the decay time constants affected time accuracy and precision for different durations. For this purpose, we characterized how the slopes and intercepts of constant error and temporal variability, with respect to $d_{s,test}$, varied as a function of network parameters. Figure 7A shows $M650_{CE}$ and $slope_{CE}$ values across time constants. The bias property corresponds to negative $slope_{CE}$ (Fig. 7A, inset, light green and yellow regions). There is a wide range of values of τ_{Fac} and $\tau_{d,GABAb}$ where the bias property was satisfied because the mean of the decoding prior distributions is 650 ms. However, there is a sector with low τ_{Fac} where the $slope_{CE}$ was positive and that did not follow the bias property (inset, red and blue regions). On the other hand, Figure 7B shows the $M650_{TV}$ and $slope_{TV}$ values across time constants. The blue and red regions located mainly along the diagonal of Figure 7B are associated with positive $slope_{TV}$; hence, they are the sectors where the scalar property was followed (Fig. 7B, inset, light red and blue regions). The gray shadow regions in Figure 7, A and B, show the combinations of decay constants associated with no linear relation between the constant error or the temporal variability as a function of $d_{s,test}$. Last, we used mutual information as a global measure of information about duration $d_{s,test}$ contained in the decoder. Because we used five different values for $d_{s,test}$, mutual information close to $log_2(5)$ indicated that d_e was strongly dependent on stimulus duration, whereas mutual information close to $log_2(1)$ implied random decoding with no interval-selectivity. Figure 5C shows mutual information as a function of $\tau_{d,GABAb}$ and τ_{Fac}. It increased monotonically as a function of the time constants, although it reached asymptotic values at large and similar $\tau_{d,GABAb}$ and τ_{Fac}.
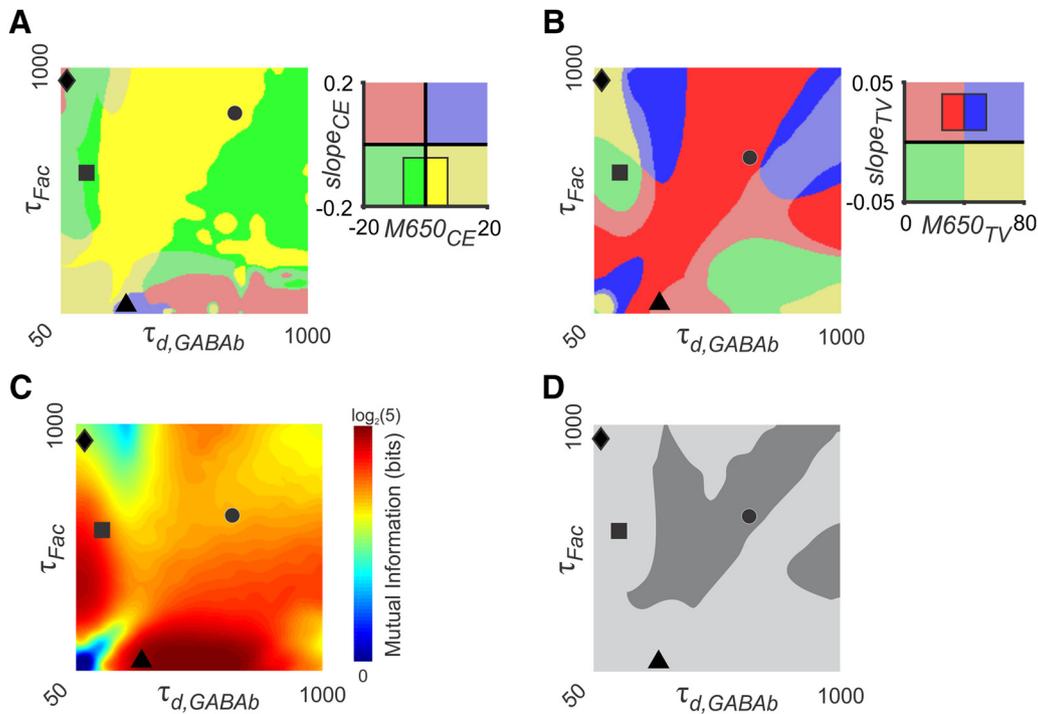
Figure 7. Synaptic effects. Effects of the time constants τ_{Fac} (ordinate) and $\tau_{d,GABAb}$ (abscissa) on the constant error (**A**), temporal variability (**B**), and mutual information (**C**). **A**, The color code corresponds to the four combinations of positive or negative values for $M650_{CE}$ and $slope_{CE}$, as depicted in the inset. The bright green and yellow regions define the combinations of time constants that produced a negative $slope_{CE}$ and, hence, followed the bias properly. **B**, Combinations of positive or negative $slope_{TV}$ and low or large $M650_{TV}$. Note that the bright red and blue regions follow the scalar property, namely, they show a positive $slope_{TV}$ (inset). The gray regions in **A** and **B** correspond to the time constant combinations that showed nonsignificant linear regressions between CE or TV and d_s, respectively. **C**, Mutual information between d_e and $d_{s,test}$, color scale corresponds to the bits of information varying to $log_2(1)$ to $log_2(5)$ for the five $d_{s,test}$ values. **D**, The area in dark gray indicates the values of τ_{Fac} and $\tau_{d,GABAb}$ where the bias and scalar properties were satisfied and the linear regression between CE or TV and d_s was significant. The black dot in each panel corresponds to the time constant values used in Figures 5 and 6, whereas the black diamond, square, and triangle correspond to the values used in Figure 8.

Notably, the largest values of mutual information were observed at low $\tau_{d,GABAb}$ and intermediate to high τ_{Fac}, as well as at intermediate to high $\tau_{d,GABAb}$ and low τ_{Fac}.

Together, these findings support the notion that interval-selectivity, as well as the bias and scalar properties showed complex relations when generated by a recurrent neural network with long-lasting input currents. Only specific ranges of the decay time constants produced networks with proper duration tuning that also followed the two hallmark properties of interval timing (Fig. 7 A,B, insets, bright areas; Fig. 7D, dark area). For example, the area around $\tau_{Fac} = 650$ ms and $\tau_{d,GABAb} = 650$ ms (Fig. 7, black circle in each panel) met these criteria and corresponds to physiological values of these decay constants.

There were combinations of τ_{Fac} and $\tau_{d,GABAb}$ where the bias and/or scalar properties were not met. For example, when $\tau_{Fac} = 100$ ms and $\tau_{d,GABAb} = 300$ ms the decoding followed the scalar but not the bias property (Figs. 7, black triangles, 8C,D, black triangles), whereas with a $\tau_{Fac} = 600$ ms and $\tau_{d,GABAb} = 150$ ms the decoding results followed the bias but not the scalar property (Fig. 7, black squares, 8C,D, black squares). Finally, there is a region around $\tau_{Fac} = 950$ ms and $\tau_{d,GABAb} = 50$ ms where both bias and scalar property were not satisfied (Fig. 7, black diamonds, 8C,D, black diamonds). Two important notions emerge from these findings. First, the prior distribution centered ~650 ms induced the bias property in specific ranges of the decay time constants (Fig. 7C). However, there were regions where the constant error did not follow the bias property, suggesting that not only prior distribution but also the activity of the neural network (Fig. 8A) was able to shape the constant error. Second, the increase in width of $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ is necessary to ob-

serve an increase in temporal variability as a function of d_s (Fig. 8B).

Noise effects

We also determined the effect of noise within the recurrent network on the decoding of duration. The purpose was to determine how the background activity of the network, in combination with the long-lasting synaptic input, modulates the constant error and temporal variability across intervals. We changed the membrane potential V of each neuron with white noise of magnitude σ_N (measured as percentage of threshold V_T; see Materials and Methods). As expected, larger noise produced a decrease in accuracy, precision, and mutual information (Fig. 9). However, the effects of noise on these parameters had an interaction with the decay time constants of the input currents. Network noise produced an increase in the negative $slope_{CE}$, indicating that the decoding accuracy decreased with noise, although the bias property was held across noise levels (Fig. 9A). Figure 9B shows that the temporal variability increased as a function of noise, and that the slope of the scalar property ($slope_{TV}$) showed an increase as a function of noise. Notably, at large noise levels the temporal variability remained similar across d_s, and therefore, the scalar property was lost. Finally, mutual information decreased asymptotically as a function of noise (Fig. 9C).

Discussion

The present study supports three conclusions. First, recurrent neural networks with time-dependent properties show time-selective responses that can follow the bias and scalar properties of interval timing in the range of hundreds of milliseconds when
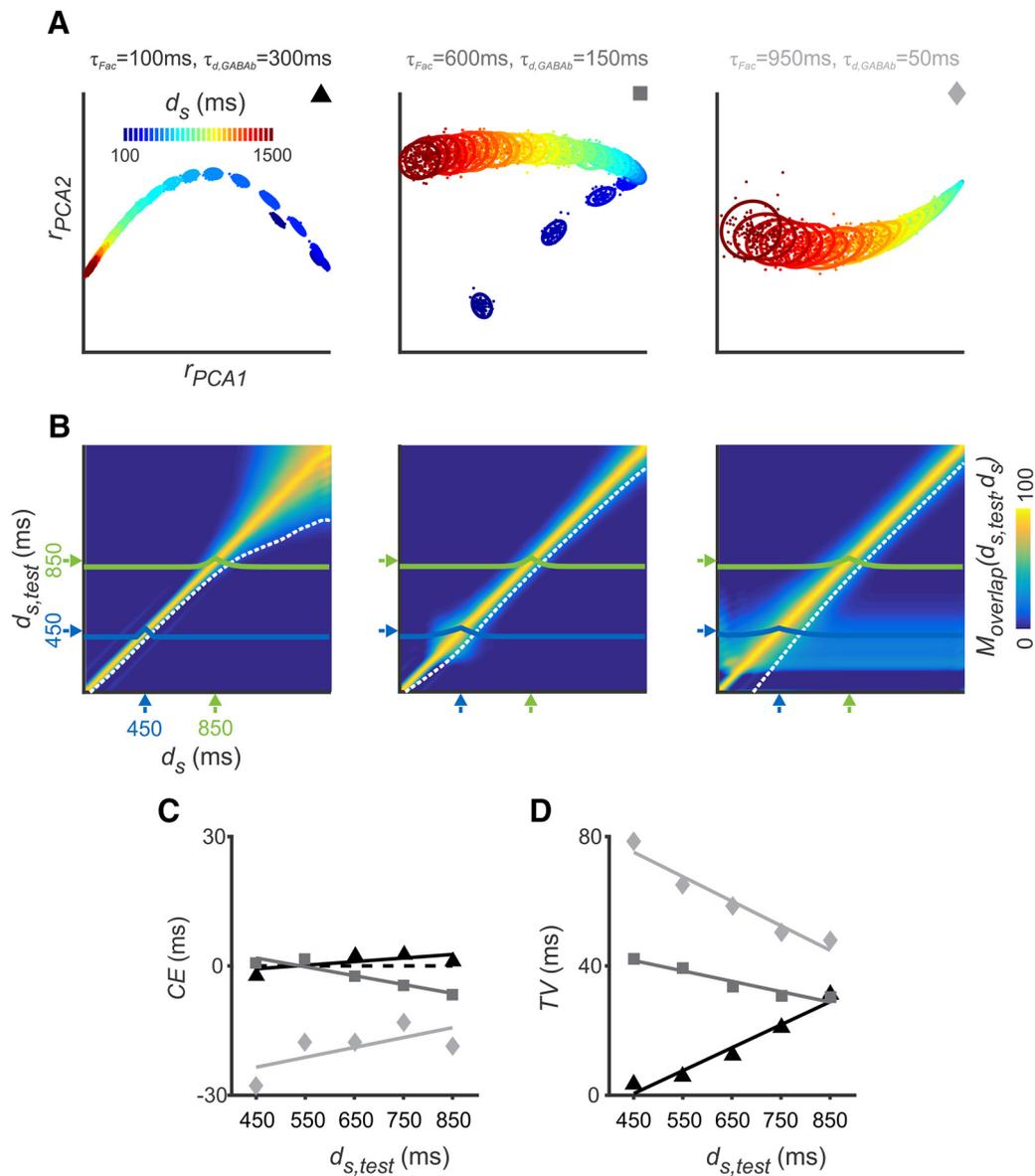
Figure 8. The bias and/or scalar properties were not followed at specific synaptic time constants. **A**, Likelihood probabilities $p(r_{PCA}|d_s)$ across d_s when the scalar but not the bias property was observed (left, $\tau_{Fac} = 100$ ms and a $\tau_{d,GABAb} = 300$ ms, black triangle; **C**, **D**), when the bias but not the scalar feature was obeyed (center, $\tau_{Fac} = 600$ ms and a $\tau_{d,GABAb} = 150$ m, gray square), and when both the scalar and bias properties were disrupted (right, $\tau_{Fac} = 950$ ms and a $\tau_{d,GABAb} = 50$ ms, light gray diamond). **B**, $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ matrix for the three conditions in **A**. A $\mathcal{M}_{overlap}(d_{s,test}, d_s)$ between adjacent durations that showed an increase as function of $d_{s,test}$ (dotted line, same notation as in Fig. 3D) produced a temporal variability that followed the scalar property. **C**, Constant error and (**D**) temporal variability as a function of $d_{s,test}$ for the three time-constant combinations in **A** and **B** (markers depicted in **A**).

read by an optimal Bayesian decoder. Second, interval-selectivity, accuracy, and precision showed complex behaviors as a function of the decay time constants of the modeled synaptic properties, suggesting that the constant error and the temporal variability, present in numerous species and timing tasks, depend on specific combinations of time constants for paired-pulse facilitation and GABAb. Finally, the increase in the variability of cell membrane potential (i.e., noise) of the recurrent network produced a generalized decrease in mutual information, accuracy, and precision of temporal processing. At large noise levels the scalar property is lost.

Several neurocomputational models have been implemented to describe the main operations behind temporal processing in the subsecond scale. These models include ramping activity (Durstewitz, 2003; Miller et al., 2003; Méndez et al., 2014), drift diffusion (Simen et al., 2011; Merchant and Averbeck, 2017), synfire chains (Bienenstock, 1995; Hass et al., 2008), neural oscil-

lators (Treisman, 1963; Medina et al., 2000; Matell and Meck, 2004; Merchant and Bartolo, 2018), coincidence detectors (Karmarkar and Buonomano, 2002; Matell and Meck, 2004), and state-dependent networks (Buonomano, 2000; Karmarkar and Buonomano, 2007; Hardy and Buonomano, 2018). Most of them implicitly or explicitly produce time selective responses and can replicate the scalar property of interval timing. However, none of the previous models were designed to reproduce, as a whole, the neural selectivity to durations, as well as the bias and the scalar properties (Hass and Durstewitz, 2014; Addyman et al., 2016). We accomplished this by using recurrent neural networks that simulated time-varying synaptic properties and membrane potential noise variations and then using Bayesian inference to optimally decode the activity of the network.

The scalar property is a hallmark of temporal processing and had an enormous impact on the study of the psychophysical and

A **B** **C**
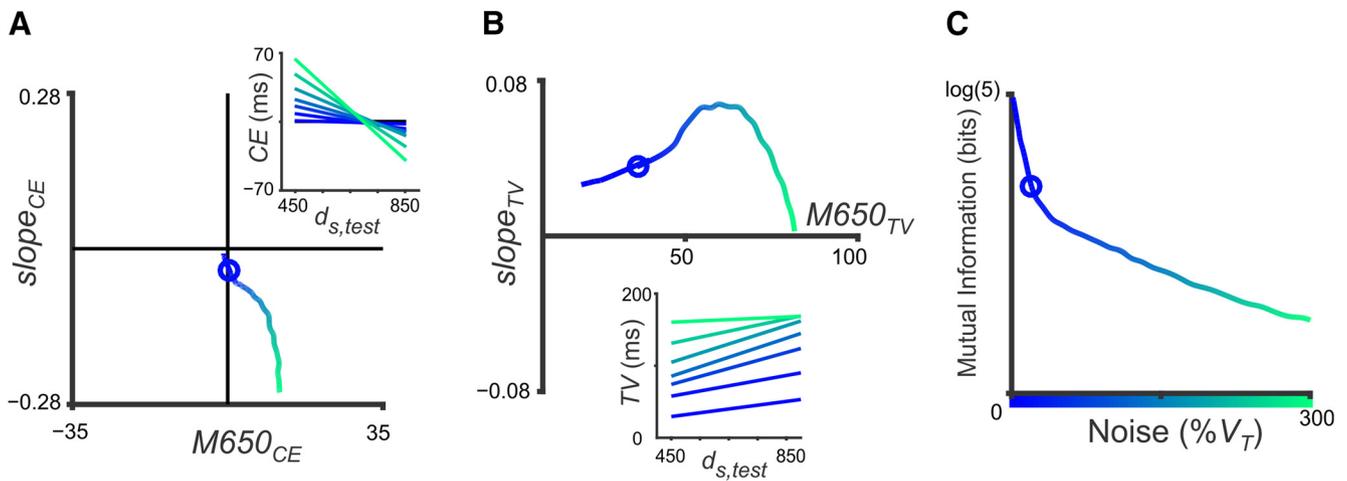
Figure 9. Noise effects. Constant error (**A**), SD (**B**), and mutual information (**C**) as a function of noise level (Eq. 1, N) in a recurrent network with a $\tau_{Fac} = 650$ ms and a $\tau_{d,GABAb} = 650$ ms. Note that at large levels of noise there was an increase in the constant error, especially at low durations (**A**, inset). In addition, as the noise increased there was a systematic increment in temporal variability accompanied by an increase in $slope_{TV}$ (**B**, inset). However, at the largest levels of network noise the scalar property was lost (**B**, green lines), and the mutual information for $d_{s,test}$ was low and no interval-selectivity was observed in the network responses (**C**). The color scale in **C** corresponds to the amount of network noise measured as percentage of V_T (normalized membrane potential; see Materials and Methods). The open blue circle corresponds to the level of noise used in Figures 4–8.

neural basis of interval timing (Gibbon et al., 1997; Grondin, 2001; Merchant et al., 2008b,c; Merchant and Averbeck, 2017). In addition, bias property, also known as the Vierordt's law, has been described since the end of the 19th century (Vierordt, 1868; Woodrow, 1934) and has had a key role in the study of the neurobiology of temporal processing. Thus, the preferred internal interval (Fraisse, 1963), the temporal memory (McAuley and Jones, 2003; Gu and Meck, 2011; Taatgen and van Rijn, 2011), the distribution of the intervals used in an experiment (i.e., global temporal context; Jones and McAuley, 2005; Jazayeri and Shadlen, 2010), and the timing abilities of the subjects (e,g, musicians vs non-musicians; Cicchini et al., 2012) have an important influence on the indifference interval and the slope of the bias property. Both the scalar and bias properties are commonly seen in the temporal performance across tasks and species, and both can be explained using the classical scalar timing model (Treisman, 1963; Gibbon et al., 1997; Grondin, 2001; Shi et al., 2013). From a Bayesian perspective, these properties emerge from the integration of two independent functions: likelihood and prior distribution. A recent study has shown a Bayesian encoding system in V1, where the activity evoked by complex visual stimuli corresponds to the likelihood response, and the spontaneous activity that changes with the animal visual experience represents the prior distribution (Berkes et al., 2011). Using this probabilistic framework, the temporal performance can be explained properly (Jazayeri and Shadlen, 2010; Shi et al., 2013; Gu et al., 2016), although it has been shown that the bias feature is not optimal when a high accuracy is needed, such as in the case of professional drummers reproducing a beat (Cicchini et al., 2012). A previous study successfully used the Bayesian scheme to model time reproduction in humans, although it assumed sensory and motor likelihood functions that were Gaussian and that explicitly followed the scalar property (Jazayeri and Shadlen, 2010; Shi et al., 2013; Roach et al., 2017). In contrast, the present study obtained the likelihood probabilities from the synaptic properties of cells in recurrent neural networks without further assumptions. Indeed, the neural population activity in the network follows the scalar property when the proper long-lasting synaptic constants are used (Figs. 4, 7). However, it is important to mention that we had to include a normal prior distribution to reproduce the change

in timing accuracy as a function of duration, consistent with previous studies on time processing (Acerbi et al., 2012; Shi et al., 2013; Roach et al., 2017). A uniform prior did not generate a robust bias property. Furthermore, the width of the Gaussian prior (σ_{prior}) influenced the slope of the scalar feature. Hence, an optimal reader that has access to this sensory timing information needs to include prior knowledge about the task conditions to reproduce the bias property and modulate the slope of the scalar feature (Ivry and Hazeltine, 1995; Merchant et al., 2008c, 2013a).

Interval tuning occurs early in the processing of sensory stimuli, as in the case of the midbrain for auditory stimuli in the range of tens of milliseconds (10–100 ms; Sayegh et al., 2011). In contrast, the primary auditory cortex (A1) shows cells that are tuned to a wider range of durations (10–500 ms) than the inferior colliculus, suggesting that duration selectivity in A1 results from integration along the time domain across the auditory ascending pathway (He et al., 1997). Conversely, the first node of the visual pathway that shows duration tuned cells is the primary visual cortex or V1, with cells tuned to a range of 50 to 400 ms (Duysens et al., 1996). In these neurophysiological studies, the width of the tuning function does not seem to scale as a function of the interval (A1: He et al., 1997; Fig. 5; V1: Duysens et al., 1996; Figure 4), although no specific analysis were performed to determine the scalar property of neurons. Accordingly, the present study predicts that using an Integrate and Fire recurrent neural network it is difficult to see the scalar property from the tuning curves of individual neurons.

However, at the population level the generalized SD $gSD(d_{s,test})$ (Fig. 3C) and, more importantly, the percentage of overlap between the ellipses $[\mathcal{M}_{overlap}(d_{s,test}, d_s)]$, had a large correlation with the decoded scalar property. On the other hand, it has been shown recently that single striatal cells follow the scalar property with a spread in the discharge rate that increases as a function of the onset response latency (Mello et al., 2015). The medial premotor areas also show temporal scaling in their response profile, namely, the responses are stretched or compressed in accordance with the produced interval (Crowe et al., 2014; Wang et al., 2018). Altogether these studies support the idea that temporal processing depends on the top-down reading of sensory timing information and the generation of predictive signals in the

motor system that obey the scalar property (Merchant and de Lafuente, 2014; Merchant and Yarrow, 2016; Petter and Merchant, 2016).

With the model reported here, the two key timing properties showed a complex behavior and could be reproduced simultaneously for a limited set of neural noise magnitudes and decay time constants, which correspond to physiological values. The magnitude of the dynamic synaptic facilitation (Varela et al., 1997; Markram et al., 1998), as well as the slow inhibition due to the activation of GABAb receptors (Benardo, 1994; Tamás et al., 2003; Wang et al., 2010) in cortical and subcortical areas show time constants that affect the processing of intervals in the scale of hundreds of milliseconds. In addition, the dynamics of both time-dependent properties can be modulated by behavioral experience (Schulz et al., 1994; Misgeld et al., 1995). As far as we know, this is the first modeling study describing how the input synaptic properties of a recurrent network can modulate the accuracy and precision of time processing. Thus, this paper shows that the ability of cortical networks to intrinsically process temporal information and follow the scalar and bias properties depends on reaching decay time constants ~650 ms for paired-pulse facilitation and 650 ms for GABAb metabotropic receptors, both of which are physiologically meaningful values.

Large neural noise in the recurrent network, which can be linked to the spontaneous activity of cells in the mammalian brain (Shadlen and Newsome, 1994; Amit and Brunel, 1997b; Vogels et al., 2011), produced an asymptotic decrease in mutual information and an increase of constant error and temporal variability. Crucially, at large noise levels the scalar property was lost because the temporal variability was similar across d_s. These findings are consistent with the observation that a subgroup of Parkinson's disease (PD) patients showed an increase in variability and $slope_{TV}$ during a time production task, whereas others lost the scalar property of interval timing (Merchant et al., 2008a; Gu et al., 2016). Hence, the heterogeneity of temporal performance in PD patients can be associated with different magnitudes of damage in the dopaminergic input, which in turn can generate different neuronal noise levels in the corticobasal ganglia circuit, as demonstrated previously (Levy et al., 2000; Timmermann et al., 2003). To end, it is important to notice that we limited the effect of fast recurrent synapses (AMPA and GABAa) on interval processing on our network. Thus, the reported results were associated to synaptic noise and not to the network recurrence. In further studies we will determine the role of the recurrent connections within the neural network on temporal processing, because it is well known that asynchronous and irregular dynamics within a neural network, due to a complex balance between fast recurrent synapses, is important to induce broad tuning activity in response to the transient input stimuli (Vogels et al., 2011; Hennequin et al., 2017).

In conclusion, learning a timing task may be associated with the appearance of interval tuning, the reaching of particular levels of spontaneous activity, as well as the adoption of specific time constants on paired-pulse synaptic plasticity and slow synaptic currents in cortical or subcortical recurrent circuits. Although the experiments needed to test these hypotheses are technically challenging, our laboratory has already demonstrated interval tuning in cells of the primate medial premotor areas and the putamen during interval production tasks (Merchant et al., 2013b, 2015; Perez et al., 2013; Crowe et al., 2014).

References

Acerbi L, Wolpert DM, Vijayakumar S (2012) Internal representations of temporal statistics and feedback calibrate motor-sensory interval timing. PLoS Comput Biol 8:e1002771. CrossRef Medline

Addyman C, French RM, Thomas E (2016) Computational models of interval timing. Curr Opin Behav Sci 8:140–146. CrossRef

Amit DJ, Brunel N (1997a) Dynamics of a recurrent network of spiking neurons before and following learning. Netw Comput Neural Syst 8:373–404. CrossRef

Amit DJ, Brunel N (1997b) Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cereb Cortex 7:237–252. CrossRef Medline

Bartolo R, Prado L, Merchant H (2014) Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. J Neurosci 34:3910–3923. CrossRef Medline

Benardo LS (1994) Separate activation of fast and slow inhibitory postsynaptic potentials in rat neocortex *in vitro*. J Physiol 476:203–215. CrossRef Medline

Berkes P, Orbán G, Lengyel M, Fiser J (2011) Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. Science 331:83–87. CrossRef Medline

Bienenstock E (1995) A model of neocortex. Netw Comput Neural Syst 6:179–224. CrossRef

Brunel N (2000) Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. J Comput Neurosci 8:183–208. CrossRef Medline

Brunel N, van Rossum MC (2007) Lapicque's 1907 paper: from frogs to integrate-and-fire. Biol Cybern 97:337–339. CrossRef Medline

Buonomano DV (2000) Decoding temporal information: a model based on short-term synaptic plasticity. J Neurosci 20:1129–1141. Medline

Buonomano DV, Merzenich MM (1998) Net interaction between different forms of short-term synaptic plasticity and slow-IPSPs in the hippocampus and auditory cortex. J Neurophysiol 80:1765–1774. CrossRef Medline

Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, Shenoy KV (2012) Neural population dynamics during reaching. Nature 487:51–56. CrossRef Medline

Cicchini GM, Arrighi R, Cecchetti L, Giusti M, Burr DC (2012) Optimal encoding of interval timing in expert percussionists. J Neurosci 32:1056–1060. CrossRef Medline

Crowe DA, Zarco W, Bartolo R, Merchant H (2014) Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. J Neurosci 34:11972–11983. CrossRef Medline

Dayan P, Abbott LF (2001) Theoretical neuroscience. Cambridge, MA: MIT.

Destexhe A, Mainen ZF, Sejnowski TJ (1994) Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. J Comput Neurosci 1:195–230. CrossRef Medline

Donnet S, Bartolo R, Fernandes JM, Cunha JP, Prado L, Merchant H (2014) Monkeys time their pauses of movement and not their movement-kinematics during a synchronization-continuation rhythmic task. J Neurophysiol 111:2138–2149. CrossRef Medline

Durstewitz D (2003) Self-organizing neural integrator predicts interval times through climbing activity. J Neurosci 23:5342–5353. Medline

Duysens J, Schaafsma SJ, Orban GA (1996) Cortical off response tuning for stimulus duration. Vision Res 36:3243–3251. CrossRef Medline

Fraisse P (1963) The psychology of time. New York, NY: Harper & Row.

García-Garibay O, Cadena-Valencia J, Merchant H, de Lafuente V (2016) Monkeys share the human ability to internally maintain a temporal rhythm. Front Psychol 7:1971. CrossRef Medline

Gibbon J, Malapani C, Dale CL, Gallistel C (1997) Toward a neurobiology of temporal cognition: advances and challenges. Curr Opin Neurobiol 7:170–184. CrossRef Medline

Grondin S (2001) From physical time to the first and second moments of psychological time. Psychol Bull 127:22–44. CrossRef Medline

Grondin S (2012) Violation of the scalar property for time perception between 1 and 2 seconds: evidence from interval discrimination, reproduction, and categorization. J Exp Psychol Hum Percept Perform 38:880–890. CrossRef Medline

Grondin S (2014) About the (non) scalar property for time perception. Adv Exp Med Biol 829:17–32. CrossRef Medline

Gu BM, Meck WH (2011) New perspectives on vierordt's law: memory-mixing in ordinal temporal comparison tasks. In: Multidisciplinary aspects of time and time perception (Vatakis A, Esposito A, Giagkou M, Cummins F, Papadelis G, eds), pp. 67–78. Berlin, Heidelberg: Springer.

Gu BM, Cheng RK, Yin B, Meck WH (2011) Quinpirole-induced sensitization to noisy/sparse periodic input: temporal synchronization as a component of obsessive-compulsive disorder. Neuroscience 179:143–150. CrossRef Medline

Gu BM, Jurkowski AJ, Shi Z, Meck WH (2016) Bayesian optimization of interval timing and biases in temporal memory as a function of temporal context, feedback, and dopamine levels in young, aged and Parkinson's disease patients. Timing Time Percept 4:315–342. CrossRef

Hardy NF, Buonomano DV (2018) Encoding time in feedforward trajectories of a recurrent neural network model. Neural Comput 30:378–396. CrossRef Medline

Hass J, Durstewitz D (2014) Neurocomputational models of time perception. Adv Exp Med Biol 829:49–71. CrossRef Medline

Hass J, Blaschke S, Rammsayer T, Herrmann JM (2008) A neurocomputational model for optimal temporal processing. J Comput Neurosci 25:449–464. CrossRef Medline

He J, Hashikawa T, Ojima H, Kinouchi Y (1997) Temporal integration and duration tuning in the dorsal zone of cat auditory cortex. J Neurosci 17:2615–2625. Medline

Hennequin G, Agnes EJ, Vogels TP (2017) Inhibitory plasticity: balance, control, and codependence. Annu Rev Neurosci 40:557–579. CrossRef Medline

Ivry RB, Hazeltine RE (1995) Perception and production of temporal intervals across a range of durations: evidence for a common timing mechanism. J Exp Psychol Hum Percept Perform 21:3–18. CrossRef Medline

Jazayeri M, Shadlen MN (2010) Temporal context calibrates interval timing. Nat Neurosci 13:1020–1026. CrossRef Medline

Jolliffe I (2002) Principal component analysis. New York: Springer.

Jones MR, McAuley JD (2005) Time judgments in global temporal contexts. Percept Psychophys 67:398–417. CrossRef Medline

Karmarkar UR, Buonomano DV (2002) A model of spike-timing dependent plasticity: one or two coincidence detectors? J Neurophysiol 88:507–513. CrossRef Medline

Karmarkar UR, Buonomano DV (2007) Timing in the absence of clocks: encoding time in neural network states. Neuron 53:427–438. CrossRef Medline

Laje R, Buonomano DV (2013) Robust timing and motor patterns by taming chaos in recurrent neural networks. Nat Neurosci 16:925–933. CrossRef Medline

Levy R, Hutchison WD, Lozano AM, Dostrovsky JO (2000) High-frequency synchronization of neuronal activity in the subthalamic nucleus of parkinsonian patients with limb tremor. J Neurosci 20:7766–7775. Medline

Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. Nat Neurosci 9:1432–1438. CrossRef Medline

Markram H, Wang Y, Tsodyks M (1998) Differential signaling via the same axon of neocortical pyramidal neurons. Proc Natl Acad Sci U S A 95:5323–5328. CrossRef Medline

Matell MS, Meck WH (2004) Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. Brain Res Cogn Brain Res 21:139–170. CrossRef Medline

McAuley JD, Jones MR (2003) Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing. J Exp Psychol Hum Percept Perform 29:1102–1125. CrossRef Medline

Medina JF, Garcia KS, Nores WL, Taylor NM, Mauk MD (2000) Timing mechanisms in the cerebellum: testing predictions of a large-scale computer simulation. J Neurosci 20:5516–5525. Medline

Mello GB, Soares S, Paton JJ (2015) A scalable population code for time in the striatum. Curr Biol 25:1113–1122. CrossRef Medline

Mendez JC, Prado L, Mendoza G, Merchant H (2011) Temporal and spatial categorization in human and non-human primates. Front Integr Neurosci 5:50. CrossRef Medline

Méndez JC, Pérez O, Prado L, Merchant H (2014) Linking perception, cognition, and action: psychophysical observations and neural network modelling. PLoS One 9:e102553. CrossRef Medline

Méndez JC, Rocchi L, Jahanshahi M, Rothwell J, Merchant H (2017) Probing the timing network: a continuous theta burst stimulation study of temporal categorization. Neuroscience 356:167–175. CrossRef Medline

Merchant H, Averbeck BB (2017) The computational and neural basis of rhythmic timing in medial premotor cortex. J Neurosci 37:4552–4564. CrossRef Medline

Merchant H, Bartolo R (2018) Primate beta oscillations and rhythmic behaviors. J Neural Transm 125:461–470. CrossRef Medline

Merchant H, de Lafuente V (2014) Introduction to the neurobiology of interval timing. Adv Exp Med Biol 829:1–13. CrossRef Medline

Merchant H, Yarrow K (2016) How the motor system both encodes and influences our sense of time. Curr Opin Behav Sci 8:22–27. CrossRef

Merchant H, Luciana M, Hooper C, Majestic S, Tuite P (2008a) Interval timing and Parkinson's disease: heterogeneity in temporal performance. Exp Brain Res 184:233–248. CrossRef Medline

Merchant H, Zarco W, Bartolo R, Prado L (2008b) The context of temporal processing is represented in the multidimensional relationships between timing tasks. PLoS One 3:e3169. CrossRef Medline

Merchant H, Zarco W, Prado L (2008c) Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. J Neurophysiol 99:939–949. CrossRef Medline

Merchant H, Zarco W, Pérez O, Prado L, Bartolo R (2011) Measuring time with different neural chronometers during a synchronization-continuation task. Proc Natl Acad Sci U S A 108:19784–19789. Ava CrossRef Medline

Merchant H, Harrington DL, Meck WH (2013a) Neural basis of the perception and estimation of time. Annu Rev Neurosci 36:313–336. CrossRef Medline

Merchant H, Pérez O, Zarco W, Gámez J (2013b) Interval tuning in the primate medial premotor cortex as a general timing mechanism. J Neurosci 33:9082–9096. CrossRef Medline

Merchant H, Bartolo R, Pérez O, Méndez JC, Mendoza G, Gámez J, Yc K, Prado L (2014) Neurophysiology of timing in the hundreds of milliseconds: multiple layers of neuronal clocks in the medial premotor areas. Adv Exp Med Biol 829:143–154. CrossRef Medline

Merchant H, Pérez O, Bartolo R, Méndez JC, Mendoza G, Gámez J, Yc K, Prado L (2015) Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. Eur J Neurosci 41:586–602. CrossRef Medline

Miller P, Brody CD, Romo R, Wang XJ (2003) A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. Cereb Cortex 13:1208–1218. CrossRef Medline

Misgeld U, Bijak M, Jarolimek W (1995) A physiological role for GABA B receptors and the effects of baclofen in the mammalian central nervous system. Prog Neurobiol 46:423–462. CrossRef Medline

Naselaris T, Merchant H, Amirikian B, Georgopoulos AP (2006) Large-scale organization of preferred directions in the motor cortex: I. motor cortical hyperacuity for forward reaching. J Neurophysiol 96:3231–3236. CrossRef Medline

Perez O, Kass RE, Merchant H (2013) Trial time warping to discriminate stimulus-related from movement-related neural activity. J Neurosci Methods 212:203–210. CrossRef Medline

Petter EA, Merchant H (2016) Temporal processing by intrinsic neural network dynamics. Timing Time Percept 4:399–410. CrossRef

Roach NW, McGraw PV, Whitaker DJ, Heron J (2017) Generalization of prior information for rapid Bayesian time estimation. Proc Natl Acad Sci U S A 114:412–417. CrossRef Medline

Sayegh R, Aubie B, Faure PA (2011) Duration tuning in the auditory midbrain of echolocating and non-echolocating vertebrates. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 197:571–583. CrossRef Medline

Schulz PE, Cook EP, Johnston D (1994) Changes in paired-pulse facilitation suggest presynaptic involvement in long-term potentiation. J Neurosci 14:5325–5337. Medline

Shadlen MN, Newsome WT (1994) Noise, neural codes and cortical organization. Curr Opin Neurobiol 4:569–579. CrossRef Medline

Shi Z, Church RM, Meck WH (2013) Bayesian optimization of time perception. Trends Cogn Sci 17:556–564. CrossRef Medline

Simen P, Balci F, de Souza L, Cohen JD, Holmes P (2011) A model of interval timing by neural integration. J Neurosci 31:9238–9253. CrossRef Medline

Stopfer M, Jayaraman V, Laurent G (2003) Intensity versus identity coding in an olfactory system. Neuron 39:991–1004. CrossRef Medline

Taatgen N, van Rijn H (2011) Traces of times past: representations of temporal intervals in memory. Mem Cognit 39:1546–1560. CrossRef Medline

Tamás G, Lőrincz A, Simon A, Szabadics J (2003) Identified sources and targets of slow inhibition in the neocortex. Science 299:1902–1905. CrossRef Medline

Timmermann L, Gross J, Dirks M, Volkmann J, Freund HJ, Schnitzler A (2003) The cerebral oscillatory network of parkinsonian resting tremor. Brain 126:199–212. CrossRef Medline

Treisman M (1963) Temporal discrimination and the indifference interval: implications for a model of the "internal clock". Psychol Monogr 77:1–31. CrossRef Medline

Tsodyks M, Pawelzik K, Markram H (1998) Neural networks with dynamic synapses. Neural Comput 10:821–835. CrossRef Medline

Varela JA, Sen K, Gibson J, Fost J, Abbott LF, Nelson SB (1997) A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. J Neurosci 17:7926–7940. Medline

Vierordt K (1868) Der zeitsinn nach versuchen. Tubingen, Germany: H. Laupp.

Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W (2011) Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. Science 334:1569–1573. CrossRef Medline

Wang J, Narain D, Hosseini EA, Jazayeri M (2018) Flexible timing by temporal scaling of cortical responses. Nat Neurosci 21:102–110. CrossRef Medline

Wang Y, Neubauer FB, Lüscher HR, Thurley K (2010) GABAB receptor-dependent modulation of network activity in the rat prefrontal cortex *in vitro*. Eur J Neurosci 31:1582–1594. CrossRef Medline

Woodrow H (1934) The temporal indifference interval determined by the method of mean error. J Exp Psychol 17:167–188. CrossRef

Zarco W, Merchant H, Prado L, Mendez JC (2009) Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. J Neurophysiol 102:3191–3202. CrossRef Medline