# A neuromechanistic model for rhythmic beat generation

Amitabha Bose[1], Áine Byrne[2*], John Rinzel[2,3]

**1** Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, NJ 07102 USA
**2** Center for Neural Science, New York University, New York, NY 10003 USA
**3** Courant Institute of Mathematical Sciences, New York University, New York, NY 10012 USA

All authors contributed equally to this work.
* aine.byrne@nyu.edu

## Abstract

When listening to music, humans can easily identify and move to the beat. Numerous experimental studies have identified brain regions that may be involved with beat perception and representation. Several theoretical and algorithmic approaches have been proposed to account for this ability. Related to, but different from the issue of how we perceive a beat, is the question of how we learn to generate and hold a beat. In this paper, we introduce a neuronal framework for a beat generator that is capable of learning isochronous rhythms over a range of frequencies that are relevant to music and speech. Our approach combines ideas from error-correction and entrainment models to investigate the dynamics of how a biophysically-based neuronal network model synchronizes its period and phase to match that of an external stimulus. The model makes novel use of on-going faster gamma rhythms to form a set of discrete clocks that provide estimates, but not exact information, of how well the beat generator spike times match those of a stimulus sequence. The beat generator is endowed with plasticity allowing it to quickly learn and thereby adjust its spike times to achieve synchronization. Our model makes generalizable predictions about the existence of asymmetries in the synchronization process, as well as specific predictions about resynchronization times after changes in stimulus tempo or phase. Analysis of the model demonstrates that accurate rhythmic time keeping can be achieved over a range of frequencies relevant to music, in a manner that is robust to changes in parameters and to the presence of noise.

## Author summary

Music is integral to human experience and is appreciated across a wide range of cultures. Although many features distinguish different musical traditions, rhythm is central to nearly all. Most humans can detect and move along to the beat through finger or foot tapping, hand clapping or other bodily movements. But many people have a hard time "keeping a beat", or say they have "no sense of rhythm". There appears to be a disconnect between our ability to perceive a beat versus our ability to produce a beat, as a drummer would do as part of a musical group. Producing a beat requires beat generation, the process by which we learn how to keep track of the specific time intervals between beats, as well as executing the motor movement needed to produce the sound associated with a beat. In this paper, we begin to explore neural mechanisms that may be responsible for our ability to generate and keep a beat. We develop a

computational model that includes different neurons and shows how they cooperate to learn a beat and keep it, even after the stimulus is removed, across a range of frequencies relevant to music. Our dynamical systems model leads to predictions for how the brain may react when learning a beat. Our findings and techniques should be widely applicable to those interested in understanding how the brain processes time, particularly in the context of music.

## Introduction

Humans have the ability to estimate and keep track of time over a variety of timescales in a host of different contexts ranging from sub-seconds to tens of seconds or more [1, 2]. On the millisecond to second time scale, for example, numerous studies have shown that humans can accurately discriminate shorter intervals from longer intervals [3, 4]. On a longer timescale, we utilize a form of time estimation that can span hours, days or years [5]. Many such examples involve the brain making a calculation over a single event, so-called "interval timing" [6, 7]. Humans can also track timing that involves multiple or repeated events. For example, we instinctively move to the beat of a piece of music through a form of sensorimotor synchronization, so-called beat-based timing [8–11]. Doing so involves identifying an underlying beat within a piece of music and coordinating the frequency and timing of one's movements to match this beat.

Understanding how humans perceive a beat has been an active area of research for quite some time. Beat perception refers to our ability to extract a periodic time structure from a piece of music. It is a psychological process in which beats can be perceived at specific frequencies, even when the musical stimulus does not specifically contain that frequency [12]. In a recent study by Nozaradan *et al.* [13], brain activity was found to entrain to the beat frequency of a musical rhythm. Additionally, participants with strong neural entrainment exhibited the best performance when asked to tap to the rhythm [13]. Various parts of the brain have been identified as being active during beat perception. Grahn and Brett reported that basal ganglia and the supplementary motor area showed increased activity for beat-based tasks, and as such, postulated that these areas mediate beat perception [14]. Interestingly, fMRI studies of participants asked to lie still with no movement while listening to music revealed that the putamen, supplementary motor area, and premotor cortex are active [15]. Thus although no external movement may be occurring, various motor areas are nevertheless active when the brain is representing a passage of time. From the theoretical perspective, error-correction [16–22], entrainment [12, 23, 24], and Bayesian [25–27] models have been proposed to account for the ability to perceive a beat.

Many beat perception studies have involved finger tapping while listening to a piece of music or a metronome [13, 28–32]. However, humans can also mentally conjure a beat in the absence of motor movement and external stimuli. These observations, in part, lead us to ask what neural mechanisms might be responsible for detecting, learning and generating a beat. We define beat generation as the brain's process of construction and maintenance of a clock-like mechanism that can produce repeatable, essentially constant, time intervals that demarcate a beat. In this formalism, the brain may be monitoring the firing times of neurons involved in beat generation to match event times of an external source. Beyond such reactive synchronization, we suggest that beat generation can occur as a strictly internally created and self-driven phenomenon in the absence of external stimuli.

Models of beat generation should take into consideration several empirically observed features taken from human finger tapping studies in the presence of isochronous tones (evenly spaced in time). First, the model's output should rapidly synchronize with the external tone sequence. Second, a model should mimic the human

ability to continue tapping even after the stimulus is removed [33], a property known as synchronization-continuation. Third, a model should quickly resynchronize after a tempo change or perturbation (deviant or phase shift) to the tone sequence. Fourth, ideally a model should be capable of addressing the phenomenon of negative mean asynchrony (NMA), the reported tendency of humans to tap on average prior to tone onsets. Although the cause of, extent of, and even the existence of NMA are still in dispute new models can potentially provide insight into the phenomenon.

Two primary modeling frameworks for addressing the above mentioned properties have been proposed: entrainment models and error-correction models. Entrainment models rely on principles of dynamical systems. Mathematically, entrainment refers to an external forcing that sychronizes a set of oscillators to a specific frequency. In the context of beat perception, entrainment models posit the existence of oscillators that resonate and entrain to the underlying periodicity creating an oscillation whose spectral profile matches that of the sound sequence. These models have been used to explain various beat-related phenomena including the emergence of pulse and meter [24] and the missing pulse percept [12]. These oscillator models are typically abstract mathematical formulations and, although generic in structure, presuppose a formulation in which the system is poised near to oscillatory-destabilization of a steady state, a Hopf bifurcation [24]. Error-correction models, on the other hand, are formulated at an algorithmic level to understand how a motor movement, such as a finger tap, can be synchronized to an isochronous tone sequence [18–22]. Errors between the current tap and tone times and between the current intertap time and stimulus period are used to adjust the timing of the next finger tap. Error-correction models provide different algorithmic ways in which to make an adjustment (see [29] for a review), but typically do not propose mechanisms for how a set of neurons would estimate and correct for the error.

In this paper, we introduce a neuromechanistic framework that can be used to construct neuronal network models that are capable of learning and retaining isochronous rhythms. In its simplest form, the network consists of a single, biophysically-based, beat generator neuron ($BG$), a periodic brief stimulus and a time-interval computation mechanism based on counting cycles of a gamma oscillation. The $BG$ does not directly receive input from the external stimulus and is thus not being entrained by it. Instead, the $BG$ learns (within a few cycles) the frequency of the stimulus thereby allowing the $BG$ to continue oscillating at this frequency, even in the absence of the stimulus. Our approach combines ideas from entrainment based, information processing based and interval timing based models. In part, it extends the heuristic two-process model [18–21] to a neural setting, by developing a neural system that learns the period and phase of the periodic stimulus, in order to bring its spikes into alignment with the stimulus tones.

A central feature of our model is the concept of a gamma counter. Gamma rhythms (30-90 Hz) are ubiquitous throughout the human nervous system [33, 34]. Here we utilize roughly 40 Hz gamma oscillations to form two discrete-time clocks that count the number of gamma cycles between specific events. This idea is similar in spirit to pacemaker-accumulator models, which also use counting mechanisms [6, 7, 35]. In our case, one clock counts gamma cycles between successive onsets of a stimulus, sometimes called the interonset interval in behavioral studies. Another neuronal clock counts gamma cycles between successive spikes of the $BG$, the interbeat interval. A comparison is made via a putative gamma count comparator (GCC) between these two counts and this information is sent to the $BG$, consistent with recent studies of neuronal circuits that can count discrete events and compare counts with stored information [36–38]. Our $BG$ neuron possesses plasticity and uses the difference in count to adjust an intrinsic parameter so that it learns the interonset interval of the stimulus. The gamma counters

also provide information to the $BG$ about its firing phase relative to stimulus onset times, thereby allowing for the possibility of synchronization with zero-phase difference of the $BG$ spikes with the stimulus. The coupling of the stimulus and $BG$ counts through the GCC is both nonlinear and non-periodic. This contrasts with coupling between stimulus and oscillator in many entrainment and information processing models where either the phase or period is directly targeted for change. Further, in such models either the coupling is periodic or the update rules are linear or vice versa [39]. Our model updates are neither periodic or linear. We note that the neuronal clocks that count cycles need not operate exclusively in the range of 40 Hz. The comparison mechanism that we describe will work for any sufficiently fast frequency oscillator.

In this paper, we will show how the $BG$ model learns and holds an isochronous beat over a wide range of frequencies that includes the band of0.5 Hz to 8 Hz which is relevant for beat generation and perception. Using mathematical analysis and a continuous time clock, we explain first how the $BG$ learns to period match. We then show how this extends to both period and phase for the discrete time clock counters. As will be seen, the discrete time clocks give rise to a natural bounded variability of spike times of the BG even when it is holding a beat. We will test our model using standard paradigms which parallel behavioral studies that are designed to study specific aspects of beat perception and production, such as NMA [13, 40], synchronization-continuation [41] and fast resynchronization times [42, 43]. The BG's mean firing time exhibits negative mean asynchrony and we describe how this NMA can be manipulated. The model resynchronizes rapidly in response to perturbations. For an abrupt change in tempo we predict and explain with the model an asymmetry in the resynchronization time for increases versus decreases in tempo. As with tempo changes, our model predicts an asymmetry in the resynchronization time due to a phase advance versus delay of the stimulus sequence. An asymmetry also arises for a transient perturbation, a single deviant onset timing in the stimulus sequence. We explain these effects by understanding how our model incorporates linear, but discrete step, error-correction to invoke non-linear changes in frequency of the BG. In turn, we develop a set of testable predictions for human behavior that help to contrast our proposed model framework from existing ones.

# Materials and methods

The main components of our model consist of a periodic stimulus with an associated neuron $S$ whose spikes mark each stimulus onset, a neuronal model for the beat generator, $BG$, and a gamma count comparator, GCC, which acts as a type of neural integrator as well as error detector. These components are linked together as shown in Fig. 1A. The output from the spiking neuron $S$ and of the $BG$ are sent to the GCC. There a comparison is made which is then sent via a period learning rule, $LR_T$, and a phase learning rule, $LR_\phi$, to adjust $I_{\text{bias}}$ which subsequently changes the instantaneous frequency of the $BG$. The term $I_{\text{bias}}$ is taken here to represent the drive to the $BG$ that governs its frequency. It could be considered as a parameter internal to the $BG$, or can be more generally associated with summed synaptic input that drives the $BG$. In either case, it is a term that regulates the $BG$'s excitability.

## The beat generator and stimulus

The $BG$ in our model can be described using biophysical, conductance based equations and is required to have only two specific properties. The first is that it possesses a sharp (voltage) threshold. A spike of the $BG$ occurs when the voltage increases through this threshold. The second requirement is that the $BG$ has a slow process that governs

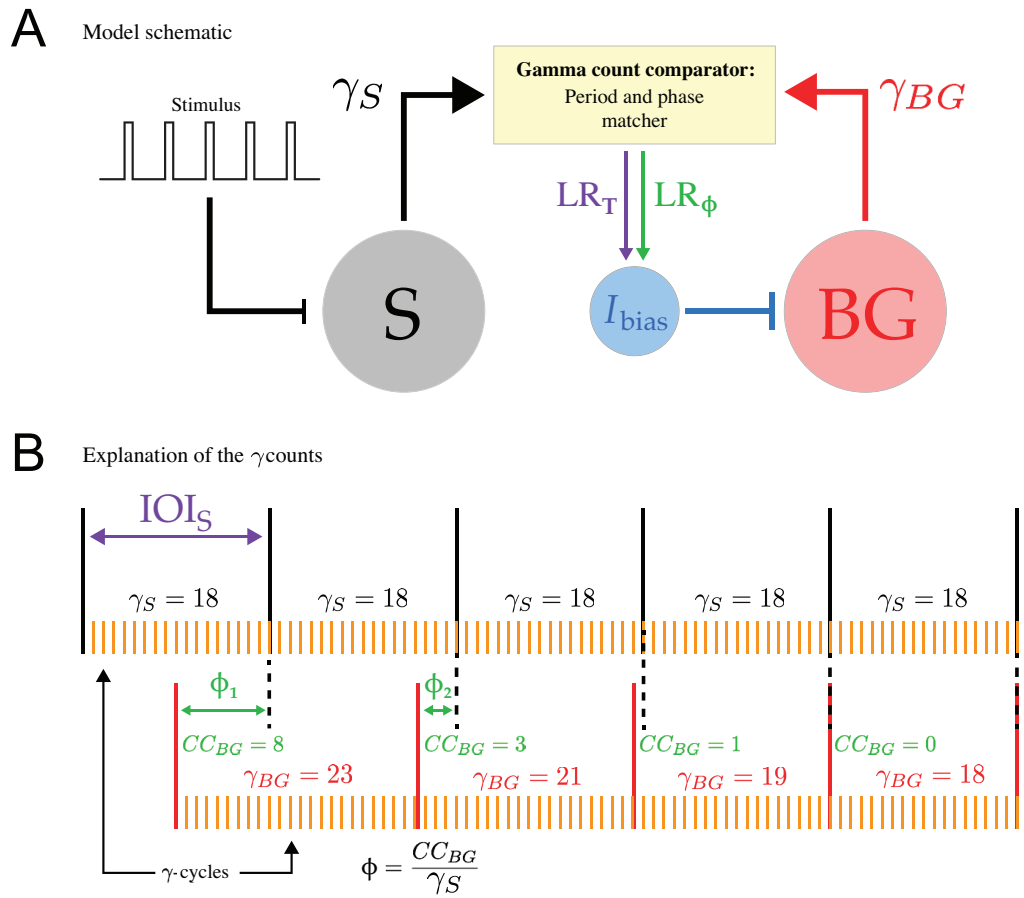



**Fig 1. Model Schematic A.** The basic components of the model include a periodic stimulus, a neuron $S$ whose firing demarcates the interonset interval, a beat generator neuron $BG$ and a gamma-count comparator, GCC. Output from the GCC is sent via the period and phase learning rules $LR_T$ and $LR_\phi$ to adjust $I_{\text{bias}}$, which controls the frequency of the $BG$. **B.** The black vertical lines indicate periodic spike times of the $S$ neuron which mark the stimulus onset. In this schematic, the interonset interval $IOI_S$ is subdivided into 18 gamma cycles ($\gamma_S = 18$) as indicated. The red vertical lines indicate $BG$ firing times with the gamma counts $\gamma_{BG}$ as indicated. As the $BG$ spikes align to the stimulus, both $\gamma_{BG}$ and the phase $\phi = CC_{BG}/\gamma_S$ change, until $\gamma_{BG} = \gamma_S$ and $\phi = 0$.

the time between its spikes. A simple model that possesses these basic properties is the leaky integrate and fire ($LIF$) model which we first use to describe our analytic findings. Our simulation studies utilize a biophysical model motivated by models of delta waves in sleep, as we require a similar frequency range for the $BG$. To that end, we chose voltage-gated currents similar to those from an idealized model for sleep spindle rhythms of thalamo-cortical relay cells, namely for the slow wave of relay cells in burst mode [44]. The slow wave can be generated by the interplay of a $T$-type calcium current, and $I_h$ current and a leak current. Here, the $BG$ has a persistent sodium $I_{NaP}$, $T$-type calcium $I_{CaT}$, sag $I_h$ and leak $I_L$ currents. In the text, we refer to this as the $I_{NaP}$ model; for a full description and the equations see Appendix. The analytic results also hold for the $I_{NaP}$ model, but as the analysis is more complicated, showing so is outside the scope of this paper. For either the $LIF$ or $I_{NaP}$ models, parameters are chosen that allow for a wide range of intrinsic frequencies of the $BG$ up through 8 Hz as

$I_{\text{bias}}$ is varied as quantified by a neuron model's $I_{\text{bias}}$ versus frequency relationship which is presented in the Results. This range of frequencies is appropriate for speech and music. The time between successive spikes of the $BG$ is called the interbeat interval and is denoted by $IBI_{BG}$.

The voltage in the $LIF$ model evolves according to

$$v' = \frac{I_{\text{bias}} - v}{\tau} \tag{1}$$

where $v$ is a dimensionless variable representing voltage, $I_{\text{bias}}$ is the drive to the neuron and $\tau$ is the membrane time constant. The $LIF$ model has a spike and reset condition which makes it discontinuous. When the voltage reaches one at $t = t_s$, it is instantaneously reset to the value 0; if $v(t_s^-) = 1$, then $v(t_s^+) = 0$. When $I_{\text{bias}} > 1$ oscillations exist. In this case, the $LIF$ model is rhythmic with period given by

$$T = \tau \log_e \left( \frac{I_{\text{bias}}}{I_{\text{bias}} - 1} \right). \tag{2}$$

The period of $BG$ given by equation (2) can be adjusted to any positive value by appropriately adjusting $I_{\text{bias}}$.

For both the $LIF$ and $I_{NaP}$ models, the specific nature of the stimulus is not modeled, only the onset is of interest here.We limit our simulations to a range between 1 and 6 Hz, which corresponds to an interstimulus interval ranging from1 s down to 166 ms. There is no theoretical or practical problem to extend the model outside of this range, as further addressed in the Discussion. We utilize a neuron $S$ to faithfully transform the stimulus sequence into spikes. The interonset interval, $IOI_S$, is then defined as the time between successive $S$ spikes. The model for $S$ is not important provided that it is set to be an excitable neuron that fires quickly in response to input; see the Appendix for equations.

## The gamma oscillation counters and learning rules

The gamma count comparator, GCC, in our model utilizes two generic oscillators with frequency sufficiently larger than that of both the stimulus and the $BG$. Here it is taken to lie in the gamma range at roughly 40 Hz (Fig. 1B). We choose the oscillators to be identical, though this is not a requirement of the model. To avoid integer values, both have a frequency of 36.06 Hz (period 27.73 ms); see the Appendix for details. We let $\gamma_{BG}$ be a variable that counts the number of gamma cycles between consecutive $BG$ spikes and $\gamma_S$ be a variable that counts the number of gamma cycles between consecutive spikes of $S$. At each spike of $BG$ or $S$, the appropriate counter is reset to zero. We stress that $\gamma_{BG}$ and $\gamma_S$ are integers, but, in general, the periods that they are estimating are not integer multiples of a gamma cycle (27.73 ms). Hence, although the stimulus period may be constant the gamma counts may vary from cycle to cycle. The difference $\gamma_{BG} - \gamma_S$ provides an estimate of how different or aligned the frequencies of $BG$ and $S$ are. For example, if $S$ oscillates at 5 Hz and the $BG$ is initially oscillating at 3 Hz, then the GCC would count roughly 8 cycles between $S$ spikes and 13 cycles between $BG$ spikes. In this case, the GCC determines that the $BG$ is oscillating too slowly and sends a speed up signal to the $BG$. Alternatively if the $BG$ were initially oscillating at 6 Hz, then the GCC counts roughly 6 cycles and sends a slow down signal. In general, speeding up or slowing down of the $BG$ is achieved by changing $I_{\text{bias}}$. At each spike of the $BG$, the period learning rule adjusts $I_{\text{bias}}$ by a fraction of the difference between the gamma oscillator count of cycles between $S$ and $BG$. This learning rule ($LR_T$) assigns at each $BG$ spike

$$\mathbf{LR_T}: \quad I_{\text{bias}} \to I_{\text{bias}} + \delta_T(\gamma_{BG} - \gamma_S), \tag{3}$$

where the parameter $\delta_T$ is independent of period. This simple rule is enough to align the frequencies of $S$ and $BG$, the details of which will be explored through the derivation and analysis of a one-dimensional map. However, this frequency matching rule does not provide the beat generator with any information about its firing phase relative to stimulus onset.

To align the phase of the $BG$ to the stimulus onset, we formulate a second learning rule. We define the current count of the $BG$, $CC_{BG}$, as the number of gamma cycles from the last $BG$ spike to the current $S$ spike; see Fig. 1B. Then at each $S$ spike define $\phi = CC_{BG}/\gamma_S$ to be the phase of $BG$ firing. We use the phase to determine if the $BG$ fires "before" or "after" $S$ at each cycle. In a rhythmically active network, the concept of whether $BG$ fired before $S$ is somewhat ambiguous. We define the $BG$ to be "before" the stimulus if it fires in the second half of the stimulus period $\phi \in (0, 0.5)$. In this case we say that the $BG$ is too fast and needs to slow down. Conversely, if $\phi \in (0.5, 1)$, the $BG$ is said to fire "after" $S$ and needs to be sped up. At each $S$ spike, we update $I_{\text{bias}}$ with the second part of the learning rule ($LR_\phi$)

$$\mathbf{LR}_\phi : \quad I_{\text{bias}} \to I_{\text{bias}} + \delta_\phi q(\phi)\phi|1-\phi| \tag{4}$$

where $\delta_\phi$ is independent of period and phase and $q(\phi) = sgn(\phi - 0.5)$, with $q(0.5) < 0$. Thus if $\phi = 0$ (or 1), there is no change to $I_{\text{bias}}$. But if the $BG$ fires before $S$ ($\phi \in (0, 0.5)$), then $q(\phi) < 0$ and $I_{\text{bias}}$ is decreased to slow down the $BG$. The opposite occurs if the $BG$ fires after $S$. The absolute value keeps the last term positive as $\phi$ can become larger than 1. For example, during transitions from high to low frequency, $CC_{BG}$ can exceed $\gamma_S$. The quadratic nature of $LR_\phi$ is chosen so that the maximum change occurs for phases near 0.5. With this two-part learning rule, the $BG$ learns both the period and phase of $S$. Both parts of the rule are implemented concurrently so that the process of period and phase alignment occurs simultaneously. The two rules $LR_T$ and $LR_\phi$ target the value $I_{\text{bias}}$. In the Results section, we describe how these changes to $I_{\text{bias}}$, in turn, affect the frequency of the $BG$ which then affects the period and phase of oscillations.

## Synchronizing to the beat, stationary behavior and natural variability of spike times

Given the discreteness of our gamma counters, the $BG$ learns to fire a spike within a suitably short window of time of the stimulus onset, an interval equal to plus or minus one gamma cycle. We define this concept as *one gamma cycle accuracy*. For the earlier described choice of parameters, this amounts to $\pm 27.73$ ms from stimulus onset. We address two important and related concepts: synchronization to the beat and holding a beat. In our model, synchronization to the beat refers to the process by which the $BG$ brings its spike times within one gamma cycle accuracy of a specific stimulus frequency. Holding a beat refers to the ability of the $BG$ to maintain synchronized firing at a specific frequency over a specified stretch of time. We will say that $BG$ has synchronized to the stimulus if three consecutive $BG$ spikes each fall within one gamma cycle accuracy in time of a stimulus onset. The $BG$ is said to be holding a beat for as long as it continues to remain synchronized with the stimulus onset.

In the presence of an isochronous stimulus, the $BG$ displays what we shall call stationary behavior. This refers to the pattern of spike times of the $BG$ in response to a fixed frequency stimulus. Despite there being no source of noise in our model, the discrete nature of the gamma count comparator allows the $BG$'s spike times during stationary behavior to naturally display variability. Thus, during stationary behavior, while the $BG$'s spike times typically land within one gamma cycle accuracy of stimulus onset they can also fall outside this window. The variability of the $BG$'s spikes arises

because the gamma counters and learning rules adjust $I_{\text{bias}}$ in discrete steps whenever $\gamma_S \neq \gamma_{BG}$ or $\phi \neq 0$. What this means is that during stationary behavior, the $BG$ does not converge to a limit cycle oscillation (periodic orbit). The variables that govern the dynamics of the $BG$ do not periodically return to the same values, but instead can vary by small amounts from cycle-to-cycle. In practice, these small differences affect the exact spike times of the $BG$, creating the variability. Furthermore, as the gamma counts are not exact representations of the period, they may be equal even when $IOI_S$ and $IBI_{BG}$ are unequal. This amounts to additional variability in the $BG$'s spike times relative to the spike times of $S$.

We will determine the time that it takes for the $BG$ to resynchronize its spikes to stimulus onset after a change to the stimulus. Resynchronization is declared similarly to synchronization in that the $BG$ is required to fire three consecutive spikes each of which must lie within one gamma cycle accuracy of a stimulus onset. The resynchronization time is then taken as the time of the first synchronized spike. In all studies, we begin with the $BG$ displaying stationary behavior at a specific frequency. Because of the variability present in stationary behavior, the resynchronization times will depend on the initial conditions at the moment that the change to the stimulus profile is enacted. We will compute mean resynchronization times and standard deviations over 50 realizations, each of which differs by a small change in the initial condition of the $BG$ at the time that the stimulus profile is changed. All simulations are carried out in MATLAB with a standard Euler solver (Euler-Maruyama when noise is introduced).

# Results

We provide a short outline of the results that follow. We start with a demonstration of how the $BG$ learns to synchronize to an isochronous stimulus sequence. We then describe how the $BG$ learns a period by first utilizing a continuous time version of the gamma counters to derive a one-dimensional map. The discrete gamma counters are then used to describe period and phase matching. Next, we present the basic behaviors of the $BG$ describing its response under both stationary (isochronous stimuli) and transient (tempo changes, phase shifts and deviants) conditions. The section concludes with a brief description of the effects of parameter changes and noise.

## $I_{\text{bias}}$ determines the $BG$'s frequency

An oscillatory neuronal model spikes with a period that is quantifiable by its frequency versus $I_{\text{bias}}$ relation (*f-I*). This relationship is obtained from the reciprocal of (2) for the $LIF$ model and computed numerically for the $I_{NaP}$ models (Fig. 2A). The blue (red) curve depicts the *f-I* curve for the $LIF$ ($I_{NaP}$) model. In the $LIF$ model, the interspike interval is governed by the difference between $I_{\text{bias}}$ and the spiking threshold, as well as the parameter $\tau$. In the $I_{NaP}$ model, the interspike interval is determined by an interplay of the various non-linear currents (Fig. 2B). In particular, the $I_{CaT}$ and $I_L$ currents provide basic excitability to the model, the $I_{NaP}$ current allows for spikes once a voltage threshold is crossed and the $I_h$ current provides a slow depolarization of the membrane allowing the neuron's voltage to gradually reach spiking threshold. Thus the primary determinant of the interspike interval is the time constant of the $I_h$ current. An important point regarding the *f-I* relations is that they are both strictly increasing. Hence, there is exactly one value of $I_{\text{bias}}$ that yields a specific frequency. The learning rules we use make discrete changes to $I_{\text{bias}}$. Thus, there is little chance of adjusting $I_{\text{bias}}$ to the exactly correct value. Instead, the learning rules adjust $I_{\text{bias}}$ so that it stays within a small window of the correct one. The frequency relations increase steeply from frequency equal to zero. Therefore, at low frequencies, larger changes in frequency can

result from small changes in $I_{\text{bias}}$. The same is also true for the $I_{NaP}$ model at frequencies in the 3 to 8 Hz range. It is important to note that any implementation of a $BG$ model with a monotone increasing $f$-$I$ relation will produce the qualitative results described below. However, the quantitative details will certainly depend on the slope and non-linearities of these relations that are produced by different ionic currents and parameters. For example, changes to $I_{\text{bias}}$ in the LIF at frequencies above 1 Hz lead to linear changes in $BG$ frequency. Alternatively, for the $I_{NaP}$ model a change of $I_{\text{bias}}$ from say 5 to 10 produces a smaller change in $BG$ frequency than changes from 10 to 15.

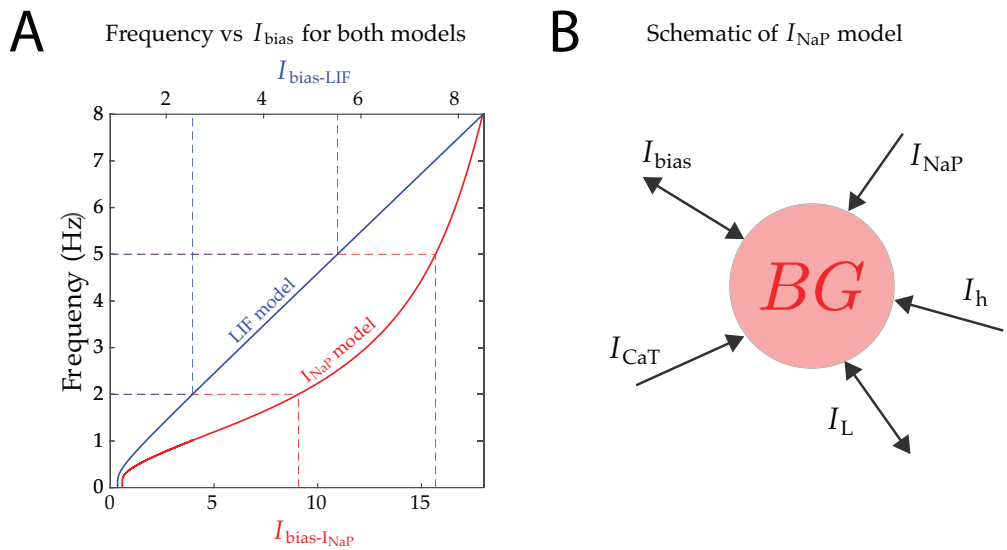



**Fig 2. Biophysical models A.** Frequency vs. $I_{\text{bias}}$ curves for $LIF$ (blue) and $I_{NaP}$ (red) models. Observe the different scales for the upper and lower horizontal axes. Dashed lines show the corresponding $I_{\text{bias}}$ values for two different frequencies. Both curves are increasing, starting from zero frequency, which yields qualitatively similar behavior from either model. While the $LIF$ relation is almost linear for frequencies larger than 1 Hz, the $I_{NaP}$ relation is more power-like, with a steeper gradient from 3 to 8 Hz, this yields differences in the quantitative behavior of the models. **B.** Schematic of $I_{NaP}$ neuron with the different ionic currents that contribute to its excitability and spiking behavior.

The $BG$ learns to oscillate at a frequency by adjusting its bias current through the set of plasticity rules $LR_T$ and $LR_\phi$ (Fig. 3). The $BG$ is initially set to oscillate at 2 Hz with $I_{\text{bias}} = 9.06$ Hz. At $t = 0$ ms, we adjust the stimulus frequency to 4.65 Hz and activate the period learning rule $LR_T$ (Fig. 3A). Notice how the cycle period of the $BG$ increases on a cycle-by-cycle basis until it matches the stimulus period. This results from the value of $I_{\text{bias}}$ iteratively increasing over the transition, based on the difference $\gamma_{BG} - \gamma_S$. The first change to $I_{\text{bias}}$ does not occur until $t = 500$ ms, which is the time the $BG$ naturally would fire when oscillating at 2 Hz, since $LR_T$ updates (purple curve in the bottom panel of Fig. 3A) are only made at spikes of the $BG$. At around $t = 2.25$ s, the value of $I_{\text{bias}}$ falls within one gamma cycle accuracy as depicted by the blue band but continues to adjust. Note that $I_{\text{bias}}$ does not settle down to a constant value. Instead it changes by $\pm\delta_T$ whenever $|\gamma_{BG} - \gamma_S| = 1$. Additionally, since $LR_T$ contains no phase information, the spikes of the $BG$ are not synchronized in phase with those of $S$. At $t = 4.2$ s (20 cycles of the stimulus), the stimulus is completely removed, and the $BG$ continues to oscillate at roughly 4.65 Hz. This shows that the $BG$ has learned the new frequency and does not require periodic input to make it spike. There are still

adjustments to $I_{\text{bias}}$ because the $BG$ continues to compare $\gamma_{BG}$ with the last stored value of $\gamma_S$. This example demonstrates how the $BG$ oscillates at a learned frequency rather than through entrainment to external input.

When both learning rules operate together, the $BG$ learns both the correct period and phase. Starting with the same initial conditions as in Fig. 3A, now at $t = 0$ ms both $LR_\phi$ and $LR_T$ are turned on (Fig. 3B). Note the very rapid synchronization of the $BG$'s spikes with the stimulus onset times. The middle panel shows how $I_{\text{bias}}$ grows much more quickly when both rules are applied. The first update is due to the phase learning rule at the third stimulus spike, at $t = 433.5$ ms, which is earlier than in the previous example. This causes enough of an increase in $I_{\text{bias}}$ for the $BG$ to immediately fire, which causes an update due the period learning rule. The lower panel shows how the two rules $LR_T$ and $LR_\phi$ contribute to the change in $I_{\text{bias}}$. Note that the learning is not sequential with period learning preceding phase learning or vice versa. Rather, period and phase learning occurs concurrently. Below we shall describe in more detail each of these learning rules and their role in synchronizing the $BG$ with the stimulus.

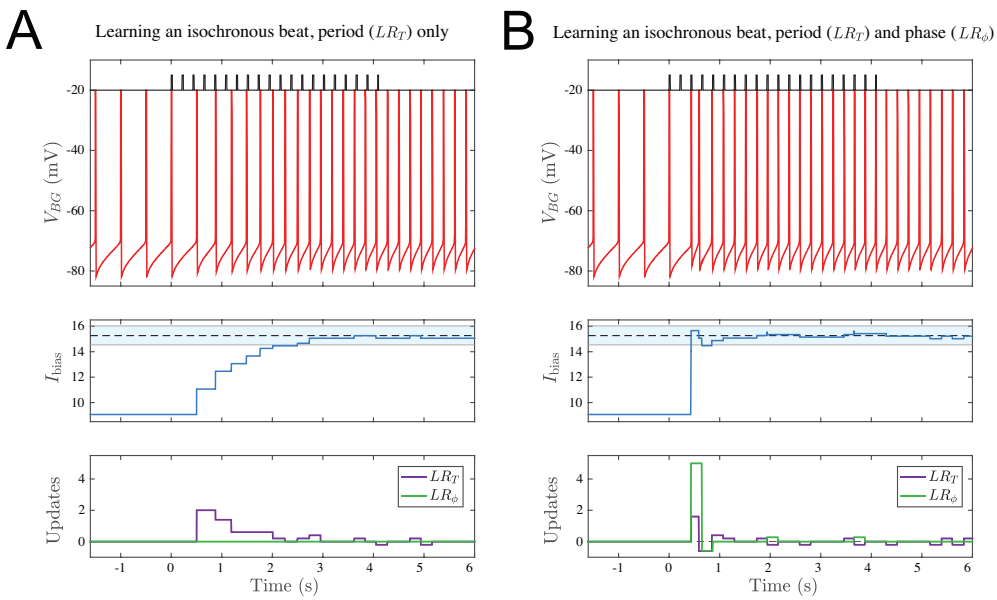



**Fig 3. Effect of learning rules A.** Period Matching with $LR_T$. The voltage time course (top) of the $BG$ is shown under three different conditions; i) Endogenously oscillating at 2 Hz with no stimulus and no learning rule from $t = -1.5$ to 0 s; ii) Learning the 4.65 Hz isochronous period in the presence of the stimulus from $t = 0$ to 4.2 s; iii) Performing synchronization continuation once the stimulus is removed at $t = 4.2$ s. $I_{\text{bias}}$ changes on a cycle-by-cycle basis (middle panel). Around $t = 2.25$ s, $I_{\text{bias}}$ enters the shaded blue regime $(14.54, 16.03)$ that represents the one gamma cycle accuracy range from the 15.27 value that produces exactly 4.65 Hz. After $t = 2.75$ s , $I_{\text{bias}}$ lies very close to the correct value. Updates to $I_{\text{bias}}$ rely solely on $LR_T$, the period learning rule (lower panel). **B.** Period and phase matching with both $LR_T$ and $LR_\phi$. Time courses are the counterparts of those in Panel A. Now the convergence to the correct frequency and phase occurs very quickly by about $t = 1.2$ s. Note the early large change in $I_{\text{bias}}$ due to $LR_\phi$. In both panels after $t = 4$ s, there are still changes to $I_{\text{bias}}$ as there are minor adjustments to $I_{\text{bias}}$ though the frequency of the $BG$ stays within the equivalent of one gamma cycle accuracy of 4.65 Hz ($IBI_{BG} = 215 \pm 27$ ms). Note that once the stimulus is switched off, only $LR_T$ continues to update $I_{\text{bias}}$. Here, and in Figs. 5–8, we used the $I_{NaP}$ model and set $\delta_T = 0.2$, $\delta_\phi = 2.5$.

## $LR_T$: The learning rule for period matching

The dynamics of how the period learning rule $LR_T$ matches the interbeat interval of the $BG$, $IBI_{BG}$, with the interonset interval of $S$, $IOI_S$, can be explained in terms of an event-based map. Each spike of the $BG$ is treated as an event and we define a map that updates the bias current $I_{\text{bias}}$ on a cycle-by-cycle basis. To derive the map, we first use exact time differences, in effect, equivalent to a continuous time-keeping mechanism. This will allow the map to possess a parameter-dependent asymptotically stable fixed point. For simplicity of presentation, we use the $LIF$ model to derive the specifics of the map. We will then discuss how those findings inform simulations of the $I_{NaP}$ model for the discrete gamma count case.

Assume that the stimulus sequence occurs with a fixed period $T^*$ corresponding to a specified $IOI_S$, and that the $BG$ is initially oscillating with an interbeat interval of $T_0$. This $IBI_{BG}$ corresponds to a specific value $I_0$ of $I_{\text{bias}}$ given by solving (2). $I_{\text{bias}}$ is then updated to $I_1$ by comparing $T_0$ to $T^*$. In turn, this produces a new cycle period $T_1$ and so on. In general, the continuous time version of $LR_T$ updates $I_{\text{bias}}$ at each firing of the $BG$ as follows:

$$\begin{aligned} I_{n+1} &= I_n + \delta_T(T_n - T^*) \\ &= I_n + \delta_T\left(\tau \log_e\left(\frac{I_n}{I_n - 1}\right) - T^*\right), \end{aligned} \tag{5}$$

where the second line is obtained by substituting equation (2) evaluated at $I_n$ for $T_n$. Error-correction models also take the form of an iteration scheme, but typically target the next cycle period for adjustment, i.e. $T_n$ and $T_{n+1}$ would replace $I_n$ and $I_{n+1}$, respectively, in the first equation of (5). In contrast, the adjustment in our model is made to the biophysical parameter $I_{\text{bias}}$ ($I_n$) which then has a subsequent effect on the cycle period ($T_n$).

Equation (5) defines a one-dimensional map which can be expressed as $I_{n+1} = f(I_n)$, where $f(I)$ denotes the right-hand side. A fixed point of the map satisfies $I^* = f(I^*)$ whose stability can be determined by checking the condition $|f'(I^*)| < 1$. A fixed point of the map corresponds to a case where the $IBI_{BG}$ of the $BG$ is equal to the $IOI_S$ of $S$. Stability of the fixed point implies that the learning rule is convergent. Note that for any $T^*$, there is a unique fixed point of the map which satisfies $I^* = 1/(1 - \exp(-T^*/\tau))$. This means that any stimulus period can in practice be learned by the $BG$, provided that the fixed point is stable. A simple calculation shows that $|f'(I^*)| < 1$ provided $0 < \delta_T < 2I^*(I^* - 1)/\tau$. For fixed $\delta_T$, as the stimulus frequency gets smaller, $I^*$ converges to 1, and as a result the term $2I^*(I^* - 1)/\tau$ goes to zero. This expression provides the insight that convergence for lower stimulus frequencies requires taking smaller increments in the learning rule. This finding carries over to any *f-I* relation that is steeply sloped at low frequencies.

Parameter dependence and the ensuing dynamics of the map are readily illustrated graphically (Fig. 4). The one-dimensional map has a vertical asymptote at $I = 1$, a local minima at $I = (1 + \sqrt{1 + 4\delta_T\tau})/2$ and a slant asymptote of $I - \delta_T/T^*$. The graph intersects the diagonal at exactly one point, and the slope of the intersection determines the stability as calculated above. For increasing stimulus frequency, with $\delta_T$ and $\tau$ fixed, the map's graph shifts upward (Fig 4A) and the fixed point moves to larger values of $I_{\text{bias}}$. Note, for low stimulus frequency (here, 1 Hz) the fixed point is unstable. The update parameter $\delta_T$ does not change the value of the fixed point $I^*$, but affects the stability (Fig. 4B). As $\delta_T$ increases, the slope at the intersection decreases through 0 , then through -1, at which point stability is lost. If the stimulus frequency changes (eg, 2 Hz to 5 Hz), $I_{\text{bias}}$ changes dynamically as the $BG$ learns the new rhythm. The learning trajectory corresponds to the cobweb diagram on the map (Fig. 4C, black dashed lines

and arrows). Each adjustment of $I_{\text{bias}}$ occurs at a spike of $BG$ and allows it to speed up for the next cycle. In this example, it takes only a few cycles for the $BG$ to learn the new rhythm. The transition from 5 to 2 Hz (Fig. 4C, red dashed lines and arrows) demonstrates the asymmetry in convergence for similar sized changes of opposite directions. Here the convergence for the decrease in frequency occurs over fewer cycles because the value of $\delta_T$ chosen yields a fixed point at 2 Hz with near zero slope. The smaller in magnitude the slope, the less the number of cycles needed to converge. This result suggests that certain preferred frequencies can exist for specific choices of parameters.

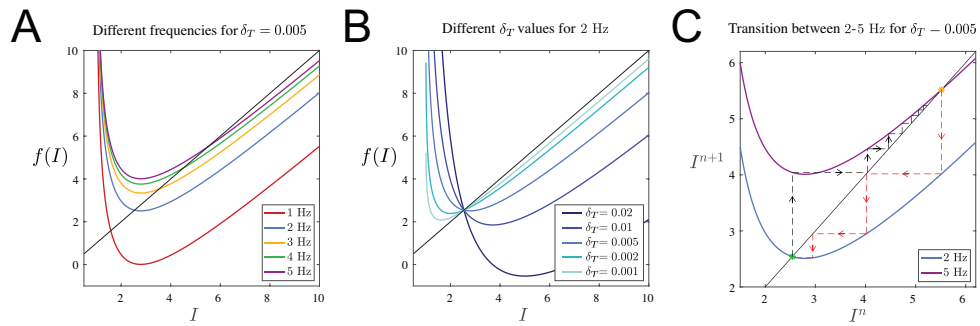



**Fig 4. One-dimensional map for period matching illustrated using the $LIF$ model A.** Plots of the map for different frequency stimulus are shown. Each curve crosses the diagonal at exactly one point, corresponding to a fixed point of the map. The slope at this intersection determines stability. As frequency increases, the fixed point moves up and gains stability. **B.** For the case of 2 Hz, as $\delta_T$ increases, stability is lost. Similar results hold for any stimulus frequency. **C.** A cobweb diagram of the convergence of a trajectory is shown. Initially, the black trajectory starts at a value corresponding to a 2 Hz oscillation. Then $I_{\text{bias}}$ is shifted to a value corresponding to a 5 Hz oscillation. The trajectory cobwebs over a number of cycles until it converges to the new stimulus frequency. The opposite transition from 5 Hz to 2 Hz is shown in red and occurs over less cycles.

In contrast to the idealized continuous-time learning rule, the gamma count-based case does not lead to updates that converge to zero. An interesting illustration is seen after an $IOI_S$ has been learned and the stimulus is turned off. Small updating persists (e.g., Fig. 3A, bottom panel). Just after the turn-off, the $IBI_{BG}$ is less than the last stored $IOI_S$ ($\gamma_{BG} < \gamma_S$). So the $BG$ is too fast, and at the next $BG$ spike the period rule $LR_T$ activates and decreases $I_{\text{bias}}$ by $\delta_T$ producing a new, longer $IBI_{BG}$. Not immediately, but after a while (just after $t = 5s$) a difference in gamma counts again arises. This time $LR_T$ increases $I_{\text{bias}}$, shortening $IBI_{BG}$ and so on. These changes are all due to $LR_T$ as the phase learning rule $LR_\phi$ can never be invoked since the stimulus is no longer present.

## $LR_\phi$: The learning rule for phase matching and synchronization with the stimulus onset

The phase learning rule $LR_\phi$ considers the current $BG$ gamma count, $CC_{BG}$, at each firing of $S$. As a result, the $BG$ has information about its phase at each stimulus onset. We use a learning rule function $\phi|1-\phi|$ that has maximal effect at $\phi = 0.5$ and no effect at $\phi = 0$ and 1. This is similar to a logistic function that attracts dynamics towards $\phi = 1$; see also [45] for a similar mathematical rule used in a different biological context. In our case $\phi = 0$ is equivalent $\phi = 1$, so our learning rule $LR_\phi$ utilizes a sign changing

function $q(\phi) = sgn(\phi - 0.5)$, $q(0.5) = -1$ to stabilize $\phi = 0$ as well. This will allow convergence via either phase increase or decrease towards synchrony. At each $S$ spike-time, the $BG$ is sped up (if $\phi \in (0.5, 1)$) or slowed down (if $\phi \in (0, 0.5)$) by adjusting $I_{\text{bias}}$ until the phase reaches a neighborhood of 0 or 1. This, in conjunction with $LR_T$ which equalizes the $IOI_S$ and $IBI_{BG}$, brings about synchronization. Note that when the $BG$ fires within one gamma cycle accuracy of $S$, $\phi = 0$ or 1. In that case, there is no update to $I_{\text{bias}}$. Thus as with $LR_T$, because of the discreteness of the learning rule updates, the value of $I_{\text{bias}}$ is brought into close proximity of the value of $I_{\text{bias}}$ that produces a specified rhythm but need not become exact. The rapid synchronization results shown earlier in Fig. 3 hold for a large range of stimulus frequencies. Under certain conditions, it is possible to derive a two dimensional map that tracks how $I_{\text{bias}}$ and $\phi$ change on a cycle-by-cycle basis. Though it is outside the scope of this paper, an analysis of the map shows that stable period and phase matching can be achieved for each $IOI_S$, if $\delta_T$ and $\delta_\phi$ are not too large. In practice these two parameters should be chosen so that the changes due to $LR_T$, $\delta_T(\gamma_{BG} - \gamma_S)$, and $LR_\phi$, $\delta_\phi \phi |1 - \phi|$, are of the same order of magnitude.

If the $IBI_{BG}$ is at least one gamma cycle longer than the $IOI_S$, then $\phi > 1$. We could have restricted the phase to be less than one by periodically extended $LR_\phi$ beyond the unit interval, but this would allow for stable fixed points at $\phi = 2, 3, 4$,etc.. Instead, the learning rule utilizes an absolute value around the term $1 - \phi$ to keep it positive. This introduces an asymmetry in the resets of $I_{\text{bias}}$. For example if $\phi = 0.1$ then $\phi(1 - \phi) = 0.09$ but if $\phi = 1.1$, then $\phi|1 - \phi| = 0.11$. Thus when the $BG$ fires after the $S$ spike after a long $IBI_{BG}$ (e.g. $\phi = 1.1$) then $I_{\text{bias}}$ is increased more than it is decreased if it fires before the $S$ spike (e.g. $\phi = 0.1$). As a result, the learning rule favors the $BG$ firing *before* the stimulus onset, as if in anticipation. This is more pronounced at lower frequencies where the slope of the *f-I* curve is much steeper than linear. This issue is explored in more detail in the following sections.

The phase learning rule $LR_\phi$ adjusts $I_{\text{bias}}$ as opposed to directly affecting the phase of the $BG$, for example, via a perturbation and reset due to the phase response curve (PRC). Using a PRC to adjust phase would lead to a situation of entrainment rather than learning. Indeed, with a PRC, bringing the value of $I_{\text{bias}}$ to within one gamma cycle accuracy to achieve a specific frequency is rarely reached. Thus if the stimulus were to be removed, a $BG$ with a PRC-based phase rule would fail to continue spiking at the correct target frequency, i.e. it would fail in a synchronization-continuation task.

## Stationary behavior and the dynamics of holding a beat

To hold a beat, the $BG$ must fire spikes within a time window of one gamma cycle accuracy of stimulus onset. As discussed earlier, the discreteness of the gamma counters and comparator causes the $BG$ spike times to naturally display variability. Thus the $BG$ must at each firing compare its period and phase relative to stimulus onset times and make necessary corrections. Holding a beat is an example of stationary behavior of the $BG$ in response to a constant frequency stimulus (Fig. 5). In this typical example, here shown at 2 Hz, each spike of the $BG$ is aligned to the closest spike of $S$ and then a timing error equal to the $BG$ spike time minus $S$ spike time is computed. The value of $I_{\text{bias}}$ hovers around the dashed black line $I_{\text{bias}} = 9.06$ which is the value that produces exactly a 2 Hz oscillation (Fig. 5A, upper). The spike times of the $BG$ jitter around those of $S$, and thus, the timing error is poised around zero (Fig. 5A, lower). While holding a beat, these differences fall within a single gamma cycle time window (dashed gray lines at $\pm 27.73$ ms). During some time windows (pink shaded region in Fig. 5B) no updating of $I_{\text{bias}}$ occurs (lower time course), but the timing differences progressively decrease and become more negative. The $BG$ spike times are drifting relative to onset times because the $IBI_{BG}$ is slightly smaller than $IOI_S$, but close enough that

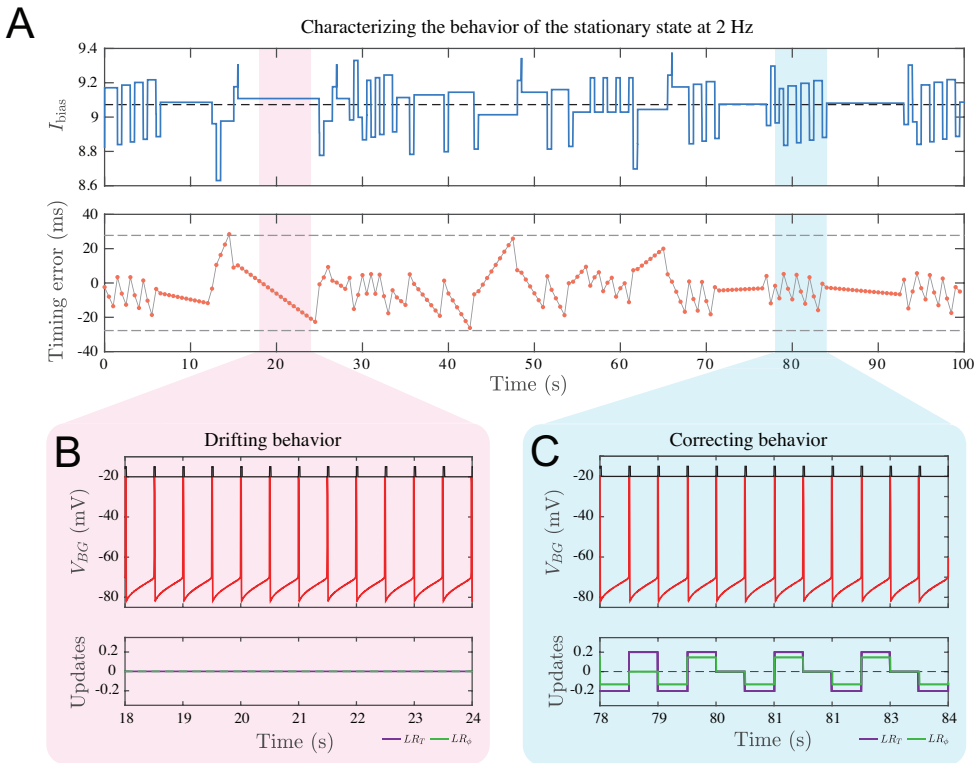



**Fig 5. Characteristics of stationary behavior:** The $BG$ alternates between drifting and corrective behavior while maintaining a beat within one gamma cycle accuracy for a 2 Hz stimulus. **A.** The upper time course of $I_{\text{bias}}$ shows intervals of rapid change interspersed with intervals of no change. In both cases, $I_{\text{bias}}$ is centered near $I_{\text{bias}} = 9.06$ (dashed line) representing the value that produces exactly 2 Hz oscillations. The lower time course shows the timing error differences, $S$ spike time subtracted from $BG$ spike times. Intervals of rapid change intermingle with intervals of slower constant change, corresponding to the intervals of rapid changes and no change, respectively, in the upper panel. Timing errors never exceed a time interval of one gamma cycle, $\pm 27.73$ ms shown by the dashed grey lines. **B.** Drifting behavior in a 6 s interval shows no updates due to the learning rules during this time. The $BG$ spikes systematically advance relative to stimulus onset times, consistent with the negative slope seen in pink in panel A (lower). **C.** Corrective behavior in a different 6 s interval shows how the learning rules $LR_T$ and $LR_\phi$ help maintain the 2 Hz oscillation. These rules are invoked whenever the counts $\gamma_{BG}$ or $CC_{BG}$ do not match with $\gamma_S$.

$\gamma_{BG} = \gamma_S$. The drift represents the fact that during this interval, the $BG$ is not entrained to the stimulus. The slope of the timing error in this interval is shallower when $I_{\text{bias}}$ is closer to 9.06, as shown for example in the time window between $t = 72$ and 76 $s$. During some intervals (e.g. shaded blue region), the update rules are actively trying to keep the period and phase of the $BG$ aligned with $S$. Although the timing errors are not large in this case, the counts $\gamma_{BG}$ and $CC_{BG}$ differ from $\gamma_S$ thus causing $LR_T$ and $LR_\phi$ to be invoked. During either drifting or corrective behavior, the $BG$ spike times occur closely in time with the stimulus onsets. Note that drifting and

corrective behaviour continue to occur after the stimulus is turned off, but only the period rule will be active (see Fig. 3). Many studies have pointed towards drifting dynamics in human tapping experiments [46–48]. Dunlap first noted this behavior, stating that the errors tend to get progressively more negative/positive, until a correction occurs and causes a change in direction [46].

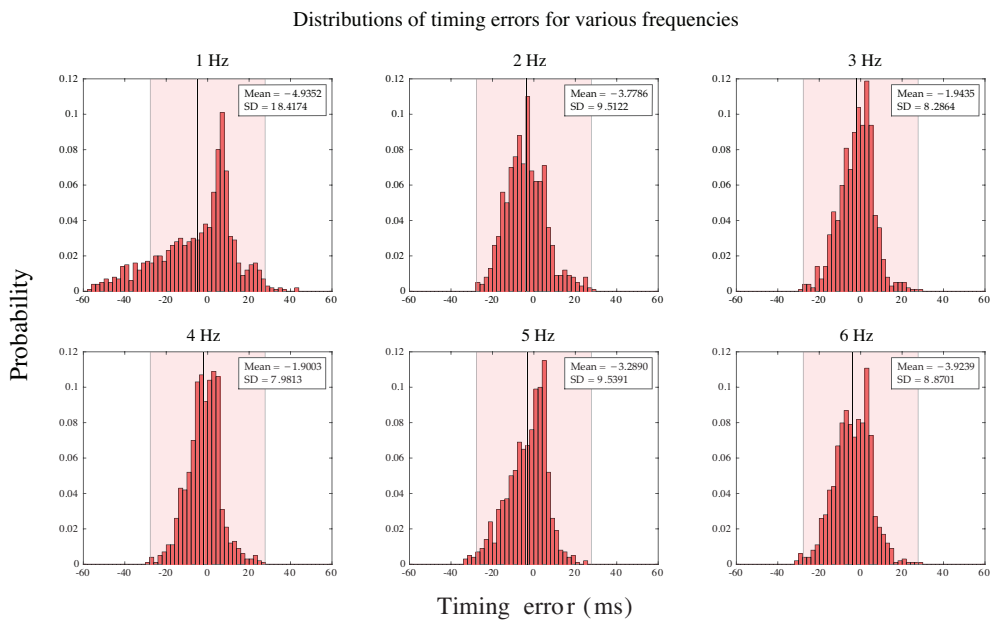



**Fig 6. Timing error distributions of beat times** Histograms of timing errors (2 ms bins) based on simulations which ran for 1000 stimulus cycles during stationary behavior at six different stimulus frequencies. Negative mean asynchrony was seen in all cases (black vertical lines) and the standard deviation at the slowest frequency is largest. Pink shading represents a region of one gamma cycle accuracy.

Although the dynamics of the $BG$ are deterministic, they are sensitive in quantitative detail to changes in initial conditions. This is because the learning rules $LR_T$ and $LR_\phi$ ignore timing differences less than one gamma cycle. To get a more general sense of the fluctuations in $BG$ firing times, we ran a simulation for 1000 stimulus cycles and calculated error distribution plots (spike time of $BG$ minus spike time of $S$). This was performed at six different stimulus frequencies in steps of 1 Hz (Fig. 6). There are several points to note. First, at all frequencies, the error distribution shows negative mean asynchrony [49, 50]. In other words, the actual time of the beat generators firing, on average, preceded the time of the stimulus onset. Second, the variance in the error distribution shows some frequency dependence, particularly with the standard deviation increasing at slower frequencies. Further, the standard deviation increases as the frequency decreases down to 0.5 Hz (simulations not shown). We also found that the standard deviation increases with frequency in the 6-8 Hz range. Accurately tapping at rates above $\sim$ 4 Hz is extremely difficult, hence, no tapping studies exist for this frequency range to either corroborate or contradict our result. However, Drake *et al.* [51] found a U-shaped dependence of the subject's tempo discrimination ability on frequency in the range 1-10 Hz, consistent with our result. We additionally note that increasing $\delta_T$ and $\delta_\phi$ leads to increased variability and larger NMA at all frequencies.

In our model, the interplay of the $BG$ neuron's *f-I* curve and the learning rules, $LR_T$ and $LR_\phi$, are responsible for the frequency dependent results. In particular, at both low

($\sim$ 0.5Hz) and high ($\sim$ 5-8Hz) frequencies the *f-I* curve is steeper than in the intermediate range. Hence, equidistant changes of $I_{\text{bias}}$ around a mean value result in different changes in instantaneous frequency. The negative mean asynchrony arises from the non-linearity in the *f-I* curve and the asymmetry in the phase learning rule $LR_\phi$. As states earlier, $LR_\phi$ pushes the $BG$ to fire before the stimulus.

## Resynchronization time: Responses to frequency changes, phase shifts and temporal deviants

As demonstrated in Fig. 3, the $BG$ is able to quickly learn a new frequency. This learning can be quantified as a resynchronization of the $BG$'s spike times with the new stimulus onset times. As previously stated, we declare the $BG$ to be resynchronized if three consecutive spikes each fall within one gamma cycle accuracy of an $S$ spike. We computed the resynchronization times as a function of several parameters including initial and final stimulus frequency (Fig. 7 shows one example). From a fixed initial stimulus frequency, we changed the stimulus frequency to different values within the range 1 to 6 Hz and computed resynchronization times. In one such case, the stimulus frequency is decreased from 3 to 2 Hz (Fig. 7A). The change is applied at $t = 0$ s (gold star) and the $BG$ takes about four seconds (eight stimulus cycles) to synchronize to the new frequency (depicted by the shaded region). During the transient, the learning rules $LR_T$ and $LR_\phi$ drive $I_{\text{bias}}$ down in order to slow the $BG$ down (lower panels of Fig. 7A). Adjustments due to $LR_T$ occur whenever the $BG$ spikes. For the first second after the change in frequency, $BG$ spikes at roughly its initial 3 Hz rate. The $S$ neuron spikes at $t = 0.5$ s, which resets $\gamma_S$. But the $BG$ is not aware of this new larger value of $\gamma_S$ until it fires at around $t = 0.66$ s. At this point, $\gamma_S > \gamma_{BG}$ and the period learning rule $LR_T$ decreases $I_{\text{bias}}$. Adjustments due to $LR_\phi$ occur whenever the stimulus neuron $S$ spikes, which now occur at the slower 2 Hz rate. These adjustments depend on the current phase of the $BG$ and are seen at times to increase $I_{\text{bias}}$, but at other times to decrease it. Within two seconds, both rules have succeeded in bringing $I_{\text{bias}}$ within one gamma cycle accuracy of the 2 Hz target value (dashed black line inside blue band in middle panel). Aligning the spike times then takes a few more seconds. In contrast, an increase in stimulus frequency can lead to much shorter resynchronization times (Fig. 7B). In the transition from a 3 to 4 Hz stimulus frequency, the $BG$ only takes about one and a half seconds (six stimulus cycles) to synchronize. The phase learning rule $LR_\phi$ plays a more prominent role as it is invoked more often due to the increase in stimulus frequency. These examples illustrate two important properties of the resynchronization process. First, the two learning rules act concurrently to adjust $I_{\text{bias}}$, but are asynchronous in that $LR_T$ adjustments occur at different times than those of $LR_\phi$ (see the lower panels of either Fig. 7). Second, adjustments to $I_{\text{bias}}$ are not periodically applied, they occur at a $BG$ or $S$ spike, and only the $S$ spikes occur periodically.

Resynchronization times increase with decreasing frequency, but are nearly constant and mostly flat for increasing frequency (Fig. 7C). Decrements from initial to final frequency lead to slower convergence than equally-sized increments. This follows from the slope of the *f-I* curve being steeper while increasing from 3 Hz than when decreasing. For the slowest stimulus frequency (1 Hz) oscillation, we have included a broader measure of synchrony, defined by $BG$ spike times falling within 5% of the interonset times, i.e. $\pm$ 50 ms around $S$ spike times. This definition is consistent with the stationary behavior shown in Fig. 6 where many of the $BG$ spikes fall outside of one gamma cycle accuracy. With this broader measure of resynchronization, the average number of cycles and standard deviation of the resynchronization to 1 Hz rhythm are reduced. Although resynchronization times are longer for frequencies decrements, the number of stimulus cycles for resynchronization do not show major differences for

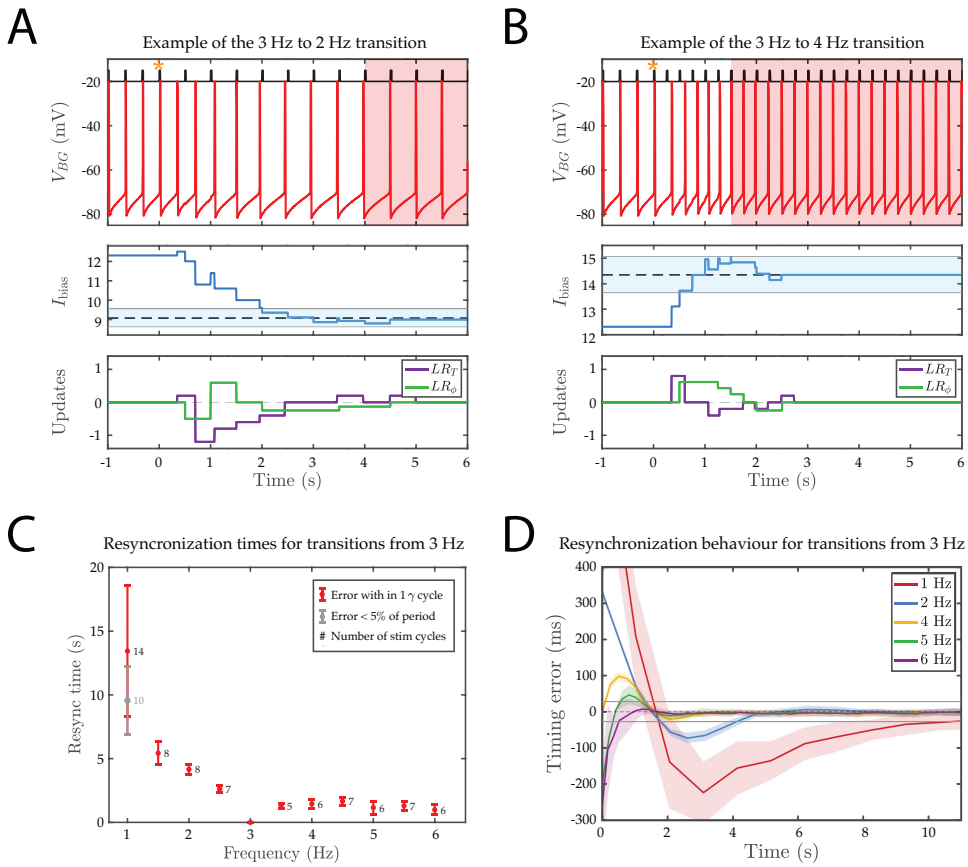



**Fig 7. Dynamics of learning a new $IOI_S$: A.** and **B.** Top time courses show examples of the resynchronization process of the $BG$ spikes with those of $S$. The change in frequency is invoked at $t = 0$ (gold star) and the $BG$ resynchronizes within a few seconds (shaded region). Middle and bottom panels show the changes in $I_{\text{bias}}$ due to $LR_T$ and $LR_\phi$ indicating that the two rules sometimes change $I_{\text{bias}}$ in the same direction, but often counteract the effect of the other. $I_{\text{bias}}$ enters the region of one gamma cycle accuracy(blue shaded band) within a handful of cycles. **C.** Mean and standard deviation for resynchronization times, together with the mean numbers of stimulus cycles needed to achieve the resynchronization are shown. A broader measure of synchrony for the 1 Hz oscillation is also included. The numbers beside the data points indicate the mean resynchronization time in terms of beats. **D.** Mean timing error transition curves for resynchronization to different terminal frequencies averaged over 50 realizations are depicted. For each time course, we aligned the last spike of the $BG$ with the last stimulus spike and then subtracted the vector of $S$ spike times from the $BG$ spike time vector. Shaded bands represent the standard deviation. Grey lines centered about zero depict timing errors within one gamma cycle accuracy. Resynchronization to lower frequencies is longer than to higher ones. Not visible on this scale, the time course for the 1 Hz (red) curve intercepts the y-axis at $\sim$ 850 ms.

increments and decrements, except for the 1 Hz case (the mean number of cycles for resynchronization are reported beside each data point).

The resynchronization process occurs stereotypically depending on whether there is an increase or decrease in frequency (Fig. 7D). We calculated the average cycle-by-cycle time differences for 50 realizations of the resynchronization process from 3 Hz to the

target frequency with the standard deviation shown in the shaded region. Decreases (increases) in frequency show initial time errors that are positive (negative). This is due to our spike alignment process (see Fig. 7 caption). Each curve is non-monotonic and, except for the 6 Hz curve, has an under- or over-shoot that transiently takes the curve outside the band of one gamma cycle accuracy (horizontal grey line). The average resynchronization times in Panel C are shown as the time at which the curve reenters this band, i.e. the time at which the timing errors become consistently less than one gamma period. Consistent with our prior results, the standard deviation bands are largest for the 1 Hz curve and relatively similar and small for the other curves.

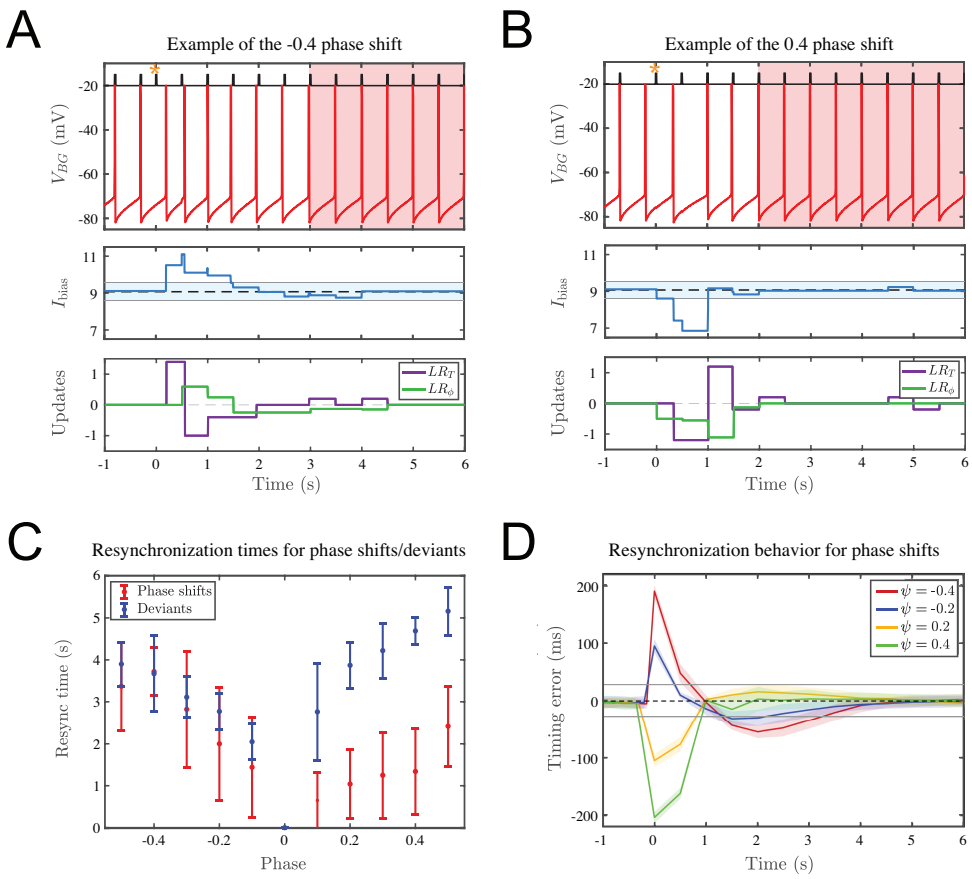



**Fig 8. Response to phase shifts and deviants:** The description of generic features of each panel is the same as in Fig. 7. **A.** and **B.** Two specific examples show that resynchronization to a positive phase shift is faster than to a negative one. Negative phase shifts cause a transient increase to $I_{\text{bias}}$ as shown in the middle panel; positive phase shifts have the opposite effect. After the transient, the return to baseline is faster for the positive phase shift because the $BG$ interprets this as equivalent to resynchronization to a higher frequency. **C.** Mean resynchronization times and standard deviations for baseline 2 Hz oscillations are shown for phase shifts (red) and deviants (blue). Notice the asymmetry for both situations, and that resynchronization to a phase shift is generally faster than to a deviant. **D.** Characteristic time courses of average timing differences (standard deviation shaded) for phase shifts give further evidence for the asymmetry in resynchronization times. Note that for the negative phase shifts the model produces an overshoot, where the timing errors temporally fall outside the region of one gamma cycle accuracy (gray line) before returning to it later for eventual resynchronization.

Resynchronization also occurs when a phase shift of the stimulus sequence occurs. Now consider the 2 Hz case for which the $IOI_S$ is 500 ms. A phase advance will occur if we shorten one $IOI_S$ to be less than 500 ms and then return the remainder of sequence to the original $IOI_S$ of 500 ms. A phase delay is the opposite, where a single $IOI_S$ is elongated. We define the phase $\psi$ of the shift to lie within $(-0.5, 0.5)$ where negative values represent advances and positive values represent delays. If the phase shift falls within one gamma cycle of the normal onset time, the $BG$ is likely to initially ignore it since no change in the gamma counts will occur. But for larger valued phase shifts, resynchronization will need to occur (Fig. 8). As an example, resynchronization for a positive phase shift at $\psi = 0.4$ (Fig. 8A) is much quicker than the corresponding negative phase shift $\psi = -0.4$ (Fig. 8B). The reason for this is how $LR_T$ changes $I_{\text{bias}}$ in either case. A negative phase shift causes the $BG$ to increase its frequency in response to the temporarily shorter $IOI_S$, followed by a return to a lower frequency. A positive phase shift causes the opposite, a transient decrease in the $BG$ frequency followed by an increase. As we have shown earlier, resynchronization times are shorter when the target frequency is larger (Fig. 7). Hence, the model predicts that resynchronization times should be shorter for positive phase shifts (Fig. 8C red). The mean timing errors (standard deviation shaded) for different phase shifts (Fig. 8D) are stereotypical in much the same way as the timing errors for tempo changes. In the current context, the timing errors start out large and then systematically reduce until they fall within one gamma cycle accuracy. The graph clearly shows that the resynchronization after positive phase shifts is faster than after negative shifts, as negative phase shifts exhibit an overshoot.

Another case where we see resynchronization is the introduction of a temporal deviant where a single $S$ spike occurs at an unexpected early or late time. Unlike phase shifts in which a single $IOI_S$ changes, a single deviant causes both the $IOI_S$ before and the $IOI_S$ after the deviant to change. An early (late) deviant causes a shorter (longer) interonset interval, followed by a longer (shorter) than normal one, followed by a return to the standard $IOI_S$. The model's response is different for early versus late deviants. For an early deviant, the phase learning rule $LR_\phi$ is invoked at the time of the early deviant. This is then followed by the period learning rule $LR_T$ at the next $BG$ spike. Both of these signal the $BG$ to speed up. For a late deviant, however, the $BG$ spikes when it normally would have. At that time, it has no new information about its phase or about the value of $\gamma_S$. Therefore there is no change to $I_{\text{bias}}$. When the late deviant arrives, it now causes $LR_\phi$ to send a slow down signal to $BG$. But any potential changes due to $LR_T$ have to wait one full $BG$ cycle to be invoked. Thus, the model reacts quicker to an early deviant than to a late deviant. Thus we predict shorter resynchronization times for earlier deviants (Fig. 8C blue). Additionally, because of the need to adjust to two different $IOI_S$s, we predict that resynchronization times due to deviants will be longer than those for comparably sized phase shifts where only a single $IOI_S$ is changed.

## Effects of parameter changes and noise

The model results are robust to perturbations and can operate over a range of parameters. To assess this we ran several simulations (not shown), where we varied intrinsic parameters of the $BG$ and the gamma counter speeds. For example, the maximal conductance for $I_L$, $I_{NaP}$ and $I_{CaT}$ was varied by up to 10 % and we measured the subsequent performance across a range of periods. This did not affect the ability of the $BG$ to learn the correct period and phase, because the *f-I* relation remained qualitatively unchanged. At a quantitative level though, the range of $I_{\text{bias}}$ values which yield gamma cycle accuracy will differ and the $BG$ neuron may have a different preferred frequency. These results indicate that the $BG$ does not require fine tuning of parameters to learn a rhythm.

Next we allowed the speed of the gamma counters for $S$ and $BG$ to be different. We kept the gamma counter for $IOI_S$ at 36.06 Hz while we varied the gamma counter frequency for the $BG$ counter by up to 10%. In all cases, across a range of frequencies there was little qualitative difference in the $BG$'s ability to learn the correct frequency. This is not surprising as the discreteness of the gamma counts allows for similar values of the counts despite there being differences in counting speed. Note that a faster gamma counter for the $BG$ tends to lead to earlier firing times relative to stimulus onset times for the parameters we have chosen. In this case, $LR_T$ tends to increase $I_{\text{bias}}$ since $\gamma_{BG}$ is larger than $\gamma_S$. On the other hand $LR_\phi$ decreases the same quantity and it is the parameter-dependent balance between the two rules that determines how much earlier on average the $BG$ fires. A slower $BG$ gamma counter has the opposite effect. These results imply that the extent of NMA can be manipulated by changing the counter frequency of the $BG$ and can even be transformed to a positive mean asynchrony if the $BG$ gamma counter is too slow.

To assess the effects of noise, we introduced stochasticity into the gamma counters (see Appendix for details). This acts to jitter the gamma periods, but for modest noise this will only cause the gamma count to discretely change by at most plus/minus 1. Since the $BG$ is monitoring its period and phase at each spike and stimulus event, it quickly adjusts to counteract these potential changes. We also see an increase in the standard deviation of the timing error, across all frequencies, during stationary behavior, as well as an increase in NMA. While this widened the distributions (as seen in Fig. 6), approximately 90% of the timing errors remain within one gamma cycle accuracy, apart from at 1 Hz where only 60% of the distribution lies within one gamma cycle accuracy (80% lie within 5% accuracy). Finally, although not explicitly modeled here, one could introduce intrinsic noise in the $BG$, for example a noisy spike threshold or ionic conductance. This small amount of noise would not change the $IBI_{BG}$ by more that a single gamma cycle and, as above, should not change the $BG$'s ability to synchronize to the external rhythm. In general, noise makes the model fit better with tapping data, exhibiting more variability and larger NMA. Given that human tapping data contains both motor and time keeper noise that our model does not attempt to disentangle (our model does not have an explicit movement component), we did not address this further.

# Discussion

We presented a modeling framework that begins to address how a neuronal system may learn an isochronous rhythm across a range of frequencies relevant to speech and music. We showed how a biophysical conductance-based model neuron, the beat generator ($BG$), adjusts its spiking frequency and aligns its spike times to an external, metronome-like stimulus. Our model employs two gamma frequency oscillators to estimate the number of oscillatory cycles between certain salient events. We posit a mechanism that compares the states of these independent counts to inform the $BG$ to either increase or decrease its instantaneous frequency and adjust its relative phase. With this idealized paradigm, we showed that the $BG$ quickly learns to hold a beat over a range of frequencies that includes, but is not limited to, 1 to 6 Hz. Further, we showed how the $BG$ reacts within a few cycles to changes in tempo, phase shifts (permanent realignment of the stimulus sequence) and the introduction of deviants (temporary misalignment of a single stimulus event). Of particular note, the $BG$ displays an asymmetry in reacting to changes to the rhythm. It adapts more quickly when the tempo is increased as opposed to decreased; correspondingly, it reacts faster to phase delays than phase advances, but slower to late deviants than early deviants. Importantly in our model formulation no direct input from the stimulus to the $BG$ is provided. This implies that the $BG$ is *learning* the correct period and phase rather than

being *entrained* to them. Secondly, no explicit or exact time intervals are required to be calculated, implying that the $BG$ does not need specific mechanisms to exactly track time. Instead, in order to tune the $BG$, one needs only to know, in some rough sense, whether the $BG$'s spikes are happening too fast or slow relative to stimulus frequency and too early or late relative to the stimulus onset. Finally, because of the discrete nature of the gamma counters the $BG$ dynamics are robust to modest parameter changes and noise.

## Beat generation differs from beat perception

Beat perception as described in many previous studies [52, 53] refers to the ability of an individual to discern and identify a basic periodic structure within a piece of music. Beat perception involves listening to an external sound source as a precursor to trying to discern and synchronize with the beat. Alternatively, we might ask how do we (humans) learn and then later reproduce a beat in the absence of any external cues. Such issues and questions lead us to consider what neuronal mechanisms might be responsible for producing an internal representation of the beat. At its most basic level, we refer to this as beat generation, and a neuronal system that does so we call a beat generator. Different than beat perception, beat generation is envisioned to be able to occur in the absence of an external cue. A $BG$ is a neural realization of an internal clock that can be used as a metronomic standard by other internally driven processes that depend on time measurements. While demonstration of a beat involves a motor action (tapping, clapping, vocalizing, head bobbing), the $BG$ could include a general representation of a motor rhythmicity but the specific motor expression (say, foot tapping) may not be an integral part of the $BG$.

## Learning a beat differs from entraining to a beat

Our formulation proposes that time measurement for beat perception and the beat generator model are oscillator-based. In this view a beat can be learned and stored as a neuronal oscillator (cell or circuit). The frequency range of interest, 1-6 Hz, is relatively low compared to many other neuronal rhythms, but similar to those seen in sleep. We rely on faster (gamma-like) oscillators to provide clocklike ticks and we assume two counters and a comparator circuit can be used for adjusting the $BG$ period and phase to match with the stimulus. Conceptually, counting and comparing with a target period are essential features of the algorithmic (or sometimes called, information processing timekeeper) approach falling into the class of error-correction strategies; see [16, 17, 20–22, 54] for examples of two-process models. These models suggest mechanisms used by humans to bring their movements into alignment with a rhythmic stimulus. They do not, however, provide a biological framework for these mechanisms. We provided a neuronal implementation of the $BG$ in the form of an oscillator with a tunable biophysical knob and two learning rules; the $BG$ is a continuous-time dynamical system, a realizable neuronal oscillator. It does not require a separate reset mechanism. The implementation also does not require a separate knob for phase correction; the two learning rules both make adjustments/corrections to the same parameter, $I_{\text{bias}}$, and they are ongoing whenever a stimulus is present. We propose this $BG$ as the internal clock — an oscillator that learns a beat and keeps it.

A different class of oscillator models for beat perception relies on large networks of neuronal units [12, 24, 28]. The units' intrinsic frequencies span the range of those that are relevant in speech and music. In the neural entrainment models of Large and collaborators, different units within the ensemble respond by phase-locking to the periodic stimulus. Units with intrinsic frequencies near that of the metronome will entrain 1:1 while those with higher intrinsic frequencies entrain with different patterns,

such as 2:1. Dominant responses are found at harmonics and sub-harmonics of the external input. Amplitude, but not precise timing relative to stimulus features (say, stimulus onset times), are described in these models. The framework is general although the identities of neuronal mechanisms (synaptic coupling or spike generation) are not apparent as the description is local, based on small amplitude perturbation schemes around a steady state and the coupling is assumed to be weak. The approach is nonlinear and provides interpretations beyond those of linear models, e.g. it identifies a beat for complex input patterns even if the beat/pulse is not explicitly a component of the stimulus [12].

Our model cannot be described as entrainment in the classical sense. Entrainment occurs when an intrinsically oscillatory system is periodically forced by an external stimulus to oscillate and, in the present context, to phase lock at the forcing frequency (or some subharmonic) that may differ from its endogenous frequency. Our $BG$ neuron is not entrained by the stimulus but rather it learns the frequency of the stimulus. The $BG$'s frequency is adapted indirectly through the control parameter in order to match with the stimulus. The influence of the stimulus on the $BG$ diminishes as learning proceeds. In fact, in the continuous time version when the frequency and phase are eventually learned, the $BG$ no longer requires the stimulus; it will oscillate autonomously at the learned frequency if the stimulus is removed or until the stimulus properties change. In the discrete time version, even after the stimulus and $BG$ periods and phase agree (to within a gamma period accuracy) modest adjustments are ongoing to maintain the rhythm. In contrast, for an entrainment model, the oscillator's parameters are fixed. The stimulus does not lead to a change in the oscillator's intrinsic properties. For a transient perturbation, the dynamics of resynchronization are according to an entrainment unit's phase response curve, which instantaneously changes the current phase of the oscillator. In contrast, the $BG$'s response to transient inputs impacts the parameter $I_{\text{bias}}$ invoked by adjustments according to either or both of the period and phase learning rules. Our model is further distinguished from entrainment models in that the $BG$ strives for zero phase difference but in an entrainment setting there is typically a phase difference between the stimulus and the units. Finally, for an entrainment model the coupling from stimulus to oscillator is periodic. In our model the influence of a periodic stimulus is delivered both periodically (via $LR_\phi$) and aperiodically (via $LR_T$).

## Implications of using gamma counters

Although humans can learn to accurately estimate time intervals [1], little is known about the neural mechanisms used to generate these estimates. For beat generation we are positing an ability to estimate time intervals (e.g., between stimulus onset events) in real time in an ongoing and flexible manner. We introduced the idea of gamma counters to perform such measurements. These counters provide a rough estimate of elapsed time that can be used to compare the internal representation of an interval with that of an external cue. The model then produces a finer representation of the interval by adjusting the $BG$'s spike time and period. There is growing evidence for the existence of counting mechanisms within neuronal systems. For example, Rose and collaborators have demonstrated that neurons in the auditory mid-brain of anurans (frogs and toads) count sound pulses in order to make mating decisions [36, 55]. These neurons have been called 'interval counting neurons' because they respond only after a threshold number of pulses have been counted provide that those pulses are spaced in time intervals of specific lengths [37]. In a very different context, it has been recently demonstrated that mossy fiber terminals in rat hippocampus have the ability to count action potentials, an ability cited as improving the reliabilty and accuracy of information transfer [38].

The discreteness of the gamma counter, used in our model, leads to variability in the

$BG$ spike times, allowing the model to exhibit negative mean asynchrony (NMA). This is consistent with finger tapping experiments which show that humans display variability in their tap times relative to an isochronous stimulus and tend to, on average, tap before the stimulus [13]. As discussed earlier, the NMA as shown in Fig. 6 is rather modest, however, changes in parameters will lead to larger NMA. In contrast, replacing the gamma frequency oscillators with continuous time clocks, which exactly determine time intervals, leads to perfect phase alignment, $\phi = 0$ (no NMA). Thus, our work posits the existence of discrete time clocks as a potential source of intertap time variability.

The gamma counters also provide an upper bound on the stimulus frequency which can be reliably learned by the $BG$ neuron. For the 36 Hz clocks used here, this limit is roughly 9 Hz. After this point the phase rule overcorrects, transiently increasing $I_{\text{bias}}$ to a value corresponding to a much larger frequency. We stress that this upper bound is dependent on the specific gamma frequency, and faster clocks may be used to keep track of shorter intervals. An interesting experimental study would be to look at the EEG power spectrum while subjects listen to periodic stimuli and monitor whether the gamma band activity changes with stimulus presentation rate.

## Relation to interval timing and other models for beat production

Many interval timing models involve accumulation (continuous time or counting of pacemaker cycles) with adjustment of threshold or ramp speed [6,7] to match the desired time interval. Applications to periodic beat phenomena, say the metronome case, would include instantaneous resetting and some form of phase adjustment/correction [56,57]. Algorithmic models may not specifically identify the accumulator as such, but instead refer to counters or elapsed time. Our $BG$ model shares some features with interval models for beat production (as described in [9] and [58]), as the $BG$ relies on counters and accumulators. Additionally, as described earlier, it shares features with entrainment models, as the $BG$ is a nonlinear oscillator. In short, the $BG$ is a hybrid.

Interval- and oscillator-based models are related. Even if not explicitly stated as such, in an interval model, the accumulator and its reset are equivalent to highly idealized models for neuronal integration, the so-called integrate-and-fire ($IF$) class of models [59]. For steady input, the state variable rises toward a target value (that is above the event threshold), rising linearly for a non-leaky $IF$ model and with a decreasing slope for a leaky $IF$ model ($LIF$), and is reset once the state variable exceeds the threshold. These $IF/LIF$ models are dynamical system oscillators, and are also nonlinear by way of the reset mechanism. However, the time constant/integration rate required for beat applications is much longer/slower than in typical applications of $IF$ models for neuronal computations where timescales of 10-30 ms are more common. These models have entrainment and phase-locking properties [60,61] and they typically show a phase difference from the stimulus. Extended in this way, periodic in time, such an interval model can be recast as an entrainment model (see also [39]). As noted by Loehr *et al.* [39], differences between such interval and continuous oscillator models do appear in some circumstances. Adding a plasticity mechanism, say for the threshold or input drive, then allows learning of a period. We described how one may analyze the dynamics of such an $LIF$ oscillator-like interval model in terms of a map (Fig 4). One could additionally add a phase correction mechanism as in two-process models in order to achieve zero-phase difference. This can be achieved in a $LIF$ model, for example, by adjusting the reset condition after reaching threshold or by utilizing phase response curves. Our mechanism for phase correction differs from these approaches in that we target the excitability parameter $I_{\text{bias}}$ for adjustment. This has the advantage that the $BG$ learns the correct phase and period allowing it to continue to hold a beat after the

stimulus is removed, similar to other two-process interval models.

The effects of noise on time estimation/production have been studied with interval models, cast as first passage time problems for accumulator models (drift-diffusion models) [62–64]. In that context, the issue of scalar timing is of significance [5, 63, 65, 66], however the time intervals of interest are typically longer than what one would find in a musical context. Wing and Kristofferson [16, 17] considered effects of noise and contrasted sensory noise with motor sources of noise, concluding that timekeeper noise was frequency dependent but motor noise was not. Whether or not scalar timing holds for short rhythmic intervals is unsettled. A number of tempo discrimination studies have failed to produce any evidence for frequency dependent errors for periods below 1000 ms [51, 67]. However, Collyer *et al.* [68] reports scalar timing in the distribution of tap times when tapping to an isochronous rhythm.

A distinctive interval model was developed by Matell and Meck [69, 70] – the striatal beat frequency (SBF) model. In this neuromechanistic description, the basic units are neuronal oscillators with different fixed frequencies. All oscillators are reset at $t = 0$; differences in frequencies of convergent units will eventually lead to collective near-coincidence (so-called beating phenomenon of non-identical oscillators) at a time that through learned choices (synapses onto coincidence detector units) can match the desired interval. It may be extended to the periodic case and considered for beat generation as discussed in [56, 57] although the brain regions involved may be different for explicit time estimation than for rhythmic prediction/reproduction [71, 72].

## Limitations of this $BG$ model

We consider here only the case of isochronous inputs. A natural next step is to consider more complex, non-isochronous stimulus sequences. Additionally, we have side-stepped questions of perception in order to focus solely on timing. Our $BG$ model does not recognize variations in pitch or sound level. For example, if stimulus events were alternating in, say, sound level, (as in [73] ) our model, as is, would not capture the effects. An extension of our model involving pairs of stimulus and beat generator clocks for each sound level could conceivably address this shortcoming.

We have chosen a particular biophysical instantiation for the $BG$. The capabilities of learning and holding a beat over a range of frequencies depends only on the monotonic frequency dependence of the control ("learnable") parameter and would not be compromised by variation of biophysical parameters. Some features of the $BG$ dynamics (say, the degree and signatures of asymmetries in resynchronization for speeding up or slowing down) can be expected to depend on the specifics of, say, the relationship between $I_{\text{bias}}$ and the intrinsic frequency, but we have not explored this in detail.

The learning rules $LR_T$ and $LR_\phi$ utilized in our study both target the excitability parameter $I_{\text{bias}}$ with a simple goal to either speed up or slow down the $BG$ so that it synchronizes with the stimulus. Alternatively, the drive could be provided as the summed synaptic input from a population of neurons afferent to the $BG$. The synaptic weights onto the $BG$ and/or internal to the afferent population could be plastic and affected by our learning rules which in spirit are similar to spike time dependent plasticity rules [74]. Our model assumes significant increments of drive at each learning step, leading to fast learning. This may be relatable at a population scale to balanced network models, where fast learning can be achieved with smaller step changes due to the large number of synapses [75]. Currently, during synchronization continuation, our $BG$ model retains its estimate of the most recent stimulus period, $\gamma_S$. We have not yet included a slow decay of this memory or a slow degradation of the $BG$ rhythm. It is plausible that the addition of noise could lead to this slow drift after the stimulus is

removed since, as we showed, noise does introduce additional variablity during stationary behavior.

We have not ascribed a location for the $BG$ within a specific brain region. As a result, we have not addressed issues of sensorimotor synchronization (SMS) where sensory processing of a beat must be coordinated with the motor action that demonstrates the beat (e.g. finger tapping). Several models for SMS in the context of beat perception already exist, for example the two-layer error-correction model of Vorberg and Wing [76] and the entrainment model of Large *et al.* [12] described earlier. Van der Steen and Keller have developed the Adaptation and Anticipation Model (ADAM) [22], a type of algorithmic error-correction SMS model, and they noted a need for an extended ADAM that would incorporate dynamical systems principles. Our model could certainly be a starting point for such an endeavor. Patel and Iversen [77] proposed the Action Simulation for Auditory Prediction (ASAP) hypothesis. In their conceptual model, the motor system primes the auditory system to be able to process auditory input. In particular, ASAP proposes that the motor system is required for beat perception. Generally, these studies raise questions about whether the causal roles of sensory and motor systems can be disambiguated in the context of beat perception and beat generation. Addressing such questions from a modeling perspective is a natural next step.

## Predictions based on the $BG$ model

Our model framework allows us to make several predictions, which are summarized here. First, the $BG$ model succeeds at synchronization continuation [78]; it can hold a beat after the sound stimulus terminates. The $BG$ will continue to oscillate with fine adjustments of its period as needed, according to $LR_T$, to match that of the most recently stored $IOI_S$ of the stimulus (as in Fig. 3). Error-correction models would also continue oscillating at the last stored frequency [20, 21]. In contrast, for an entrainment model without a plasticity mechanism, the oscillators are likely to return to their original intrinsic frequencies after the stimulus is removed. However, Large *et al.* [12] have illustrated using their two-layer model, that the network can hold a beat if the units within the motor layer, have bistable properties, i.e. a unit may have a steady state (damped resonator) coexistent with an oscillator state. The main difference between our approach and others is that period error correction ($LR_T$) occurs at the $BG$ spike times rather than at stimulus onsets. As such, even after stimulus removal, comparison between the $BG$ period and last stored stimulus period continues. Second, the time course of adjusting to a sudden tempo change occurs over seconds and has more or less monotonic phases of slowing down or speeding up (Fig. 7). If the new sound stimulus is stopped during this transition, we predict from the model that the $BG$ will still learn the new beat frequency. However, the phase of the $BG$ will differ depending on when during the transition the stimulus is removed. This could be detected using EEG or perhaps a finger-tapping demonstration. This prediction differs from those made by traditional error-correction models which will cease making updates after the stimulus is turned off. Hence, these models may not reach the new frequency, and instead settle at an intermediate frequency. An entrainment model also relies on the external input to match its frequency. Thus if the sound stimulus is stopped during the transition, an entrainment model is likely to return to its basal frequency, and not learn the new one. Third, resynchronization should be faster after a phase shift of the rhythmic stimulus than after a single timing-deviant sound event. Lastly, the model predicts an asymmetry in the resynchronization process after phase shifts (advance versus delay), deviants (early versus late) and tempo changes in the stimulus sequence. We hesitate to assign whether the resynchronization time will be shorter or longer in these cases, as this property depends on the intrinsic dynamics of the $BG$ as defined by

its $f$-$I_{\text{bias}}$ relationship. Similarly, the specific responses of an entrainment model for these cases depends on the phase response properties of the underlying oscillator model. Typically these are one time adjustments to the phase of the oscillator, followed by a transient return to the entrained solution. Thus rather than elucidating concrete differences in predictions of our $BG$ model versus others, instead consideration of different models may help to narrow the choice of biophysical currents that produce $I_{\text{bias}}$-$f$ or phase response properties that allow neuron spike times to match empirical data.

### Future directions

There are several questions that we plan to address in our future modeling and behavioral studies. How sensitively do timing errors depend on variability of the gamma counters and, say, on stimulus frequency? To what degree can the $BG$ model track modulations of the beat frequency? Different candidate beat generators that possess different ionic currents will have *f-I* relationships that are different than the ones presented here. How sensitively do the quantitative results described here depend on the shape of these curves? When considering an ensemble of beat generating neurons, how does coupling between these neurons shape the dynamics of learning? How could the model be enhanced to become predictive, to not just track modulation but to predict dynamic trends? Going beyond isochronous timing only, we plan to consider more complex rhythms. For example, suppose we consider the effect of shifting identically the timing of alternate stimulus tones. Eventually, after a sequence of modest shifts, the beat frequency would be halved although the number of stimulus events would be maintained but with a different temporal pattern. How is the transition of frequency halving executed dynamically? Perhaps there is a regime of shift values where beat determination is ambiguous, a possible regime of bistability. A different manipulation toward a complex stimulus could involve parametrically changing the sound intensity or pitch of alternate tones. Such cases will bring us toward questions of perception and auditory streaming together with beat perception.

### Conclusion

The questions surrounding how we perceive and keep a beat are easy to pose but developing models for beat perception and generation present challenges. Our model is a first-pass attempt at formulating and analyzing a neuromechanistic model that can learn a beat. Our approach stems from a neurobiological and dynamical systems perspective to develop neuronal system-based models for beat learning and generation. The essential features involve neuro-based elapsed timekeepers, time difference comparators and a neural oscillator (cellular or circuit level) with some plasticity and learning rules. Looking ahead one hopes for development of more general beat and rhythm pattern generators (for complex rhythmic sounds, music pieces) that can be stored in a silent mode and are both recallable and replayable.

## Supporting information

**S1 Appendix.** The model for the $BG$ consists of a set of biophysical conductance-based equations which we call the $I_{NaP}$ model that incorporate a low-threshold calcium current $I_{CaT}$, a sag current, $I_h$, a persistent sodium current $I_{NaP}$

and a leak current, $I_L$. The current balance equations are given by

$$\begin{aligned} C\frac{dV}{dt} &= I_{\text{bias}} + I_{int} - g_L[V - E_L] - g_{CaT}m_\infty(V)h[V - E_{Ca}] \\ &\qquad - g_h r[V - E_h] - g_{NaP}a_\infty(v)[v - E_{Na}] \quad (6) \\ \frac{dh}{dt} &= \frac{h_\infty(V) - h}{\tau_h(V)} \quad (7) \\ \frac{dr}{dt} &= \frac{r_\infty(V) - r}{\tau_r(V)} \quad (8) \end{aligned}$$

The term $I_{\text{bias}}$ refers to a drive whose value determines whether the isolated $BG$ can oscillate and if so, at which frequency. We set $I_{int} = -33\mu\text{A/cm}^2$ so that if $I_{\text{bias}} = 0\mu\text{A/cm}^2$ there are no oscillations. The parameters $g_{CaT}, g_h, g_{NaP}, g_L$ and $E_{Ca}, E_h, E_{Na}, E_L$ refer to conductances and reversal potentials for the calcium, sag, sodium and leak currents, respectively. The functions $x_\infty(V) = 1/(1 + \exp(-(V - v_x)/k_x))$ for $x = m, a, r, h$ are each sigmoidal functions with half-activation voltages $v_x$ and accompanying slopes $k_x$. The time constants are $\tau_h(V) = \tau_L/(1 + \exp((V - v_h)/k_h)) + \tau_R(1 + \exp(-(V - v_h)/k_h))$ and $\tau_r(V) = \tau_{r_{max}}/\cosh((V - v_{r_\tau})/(2k_{r_\tau}))$. The $T$-current is considered to have instantaneous activation, modeled by $m_\infty(V)$ and a slow inactivation, governed by the $h$ variable. The sag current simply has a slow activation variable $r$. The persistent sodium current has just instantaneous activation given by $a_\infty(V)$. The parameter values are as follows: $C = 1$, $g_{CaT} = 11$, $g_h = 1$, $g_{NaP} = 0.1$, $g_L = 1.6$, $E_{Ca} = 50$, $E_h = -30$, $E_{Na} = 50$, $E_L = -70$, $v_m = -40$, $k_m = 6.5$, $v_a = -67$, $k_a = 1$, $v_r = -70$, $k_r = 12$, $v_h = -60$, $k_h = 6$, $v_{r_\tau} = -75$, $k_{r_\tau} = 8$, $\tau_L = 30$, $\tau_R = 5$, $\tau_{r_{max}} = 850$ (where capacitance is in units of $\mu\text{F/cm}^2$, conductances are in $\text{mS/cm}^2$, time constants are in ms and all others parameter units are mV):.

The equations for the $S$ neuron are

$$\begin{aligned} C_S\frac{dV_S}{dt} &= I^S_{bias} + g_{stim}I_{stim}(t) - g_{L_S}[V_S - E_L] - g_{CaT_S}m_\infty(V_S)h_S[V_S - E_{Ca}] \quad (9) \\ \frac{dh_S}{dt} &= \frac{h_\infty(V_S) - h_S}{\tau_h(V_S)} \quad (10) \\ & \quad (11) \end{aligned}$$

The term $I_{stim}(t)$ is the periodic current provided from the stimulus. During each cycle, it is positive for 25 ms and 0 otherwise. It is taken to be large enough to ensure that $S$ fires within 5 ms of sound onset times. The $m_\infty$, $h_\infty$ and $\tau_h$ functions are as above, the parameter values are the same unless otherwise stated: $I^S_{bias} = -14$, $g_{stim} = 6$, $g^S_{CaT} = 10$ (units as above).

The $\gamma$ counters are constructed as follows. Solve $x' = -x/\tau_x$ with $x(0) = 2$ until it reaches 1 at $t = t_g$ and is reset to 2; $x(t_g^-) = 1$ reset to $x(t_g^+) = 2$. The counters keep track of the number of resets. We chose $\tau_x = 40$ ms which yields an inter-spike interval of 27.73 ms (frequency of 36.06 Hz). Heterogeneity between the $BG$ and $S$ counters of roughly 10% was introduced to the $IOI_S$ by varying $\tau_x$. In the case of stochasticity, Ornstein-Uhlenbeck noise was added to the $x$ variable, with a timescale of 5 ms and Gaussian white noise with mean 0 and standard deviation 0.005. All numerical simulations were carried out in MATLAB.

# Acknowledgments

A. Bose thanks the Courant Institute of Mathematical Sciences at New York University for their support during his sabbatical. A. Byrne was supported by the Swartz Foundation.

## References


1. Merchant H, Harrington DL, Meck WH. Neural basis of the perception and estimation of time. Annual Review of Neuroscience. 2013;36(1):313–336.
2. Hardy NF, Buonomano DV. Neurocomputational models of interval and pattern timing. Current Opinion in Behavioral Sciences. 2016;8:250–257.
3. Grondin S, Killeen PR. Effects of singing and counting during successive interval productions. NeuroQuantology. 2009;7(1):77–84.
4. Mioni G, Stablum F, Grondin S. Interval discrimination across different duration ranges with a look at spatial compatibility and context effects. Frontiers in Psychology. 2014;5(717):1–11.
5. Buhusi CV, Meck WH. What makes us tick? Functional and neural mechanisms of interval timing. Nature Reviews Neuroscience. 2005;6(10):755–765.
6. Gibbon J, Church RM, Meck WH. Scalar timing in memory. Annals of the New York Academy of Sciences. 1984;423(1):52–77.
7. Treisman M. Temporal discrimination and the indifference interval. Implications for a model of the "internal clock". Psychological Monographs. 1963;77(13):1–31.
8. Tranchant P, Vuvan DT, Peretz I. Keeping the beat: A large sample study of bouncing and clapping to music. PloS ONE. 2016;11(7):e0160178.
9. Grahn JA. Neural mechanisms of rhythm perception: Current findings and future perspectives. Topics in Cognitive Science. 2012;4(4):585–606.
10. Iversen JR. In the beginning was the beat: Evolutionary origins of musical rhythm in humans. In: Hartenberger R, editor. The Cambridge Companion to Percussion. Cambridge University Press; 2016. p. 281–295.
11. Iversen JR, Balasubramaniam R. Synchronization and temporal processing. Current Opinion in Behavioral Sciences. 2016;8:175–180.
12. Large EW, Herrera JA, Velasco MJ. Neural networks for beat perception in musical rhythm. Frontiers in Systems Neuroscience. 2015;9(159):1–14.
13. Nozaradan S, Peretz I, Keller PE. Individual differences in rhythmic cortical entrainment correlate with predictive behavior in sensorimotor synchronization. Scientific Reports. 2016;6:1–12.
14. Grahn JA, Brett M. Rhythm perception in motor areas of the brain. Journal of Cognitive Neuroscience. 2007;19(5):893–906.
15. Bengtsson SL, Ullén F, Henrik Ehrsson H, Hashimoto T, Kito T, Naito E, et al. Listening to rhythms activates motor and premotor cortices. Cortex. 2009;45(1):62–71.
16. Wing AM, Kristofferson AB. The timing of interresponse intervals. Perception & Psychophysics. 1973;13(3):455–460.
17. Wing AM, Kristofferson AB. Response delays and the timing of discrete motor responses. Perception & Psychophysics. 1973;14(1):5–12.
18. Hary D, Moore GP. Synchronizing human movement with an external clock source. Biological Cybernetics. 1987;56(5-6):305–311.

19. Hary D, Moore GP. On the performance and stability of human metronome-synchronization strategies. British Journal of Mathematical and Statistical Psychology. 1987;40:109–124.

20. Mates J. A model of synchronization of motor acts to a stimulus response. I. Timing and error corrections. Biological Cybernetics. 1994;70(5):463–473.

21. Mates J. A model of synchronization of motor acts to a stimulus sequence - II. Stability analysis, error estimation and simulations. Biological Cybernetics. 1994;70(5):475–484.

22. van der Steen MCM, Keller PE. The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. Frontiers in Human Neuroscience. 2013;7(253):1–15.

23. Large EW, Jones MR. The dynamics of attending: How people track time-varying events. Psychological Review. 1999;106(1):119–159.

24. Large EW, Almonte FV, Velasco MJ. A canonical model for gradient frequency neural networks. Physica D. 2010;239(12):905–911.

25. Temperley D. An evaluation system for metrical models. Computer Music Journal. 2004;28(3):28–44.

26. Temperley D. A unified probabilistic model for polyphonic music analysis. Journal of New Music Research. 2009;38(1):3–18.

27. Vuust P, Witek MAG. Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. Frontiers in Psychology. 2014;5(1111):1–14.

28. Large EW, Fink P, Kelso JAS. Tracking simple and complex sequences. Psychological Research. 2002;66(1):3–17.

29. Repp BH. Sensorimotor synchronisation; a review of the tapping litterature. Psychonomic Bulletin & Review. 2005;12(6):969–992.

30. Fujioka T, Trainor LJ, Large EW, Ross B. Internalized timing of isochronous sounds Is represented in neuromagnetic beta oscillations. Journal of Neuroscience. 2012;32(5):1791–1802.

31. Repp BH, Su YH. Sensorimotor synchronization: A review of recent research (2006-2012). Psychonomic Bulletin and Review. 2013;20(3):403–452.

32. Chauvigné LAS, Gitau KM, Brown S. The neural basis of audiomotor entrainment: an ALE meta-analysis. Frontiers in Human Neuroscience. 2014;8:776.

33. Buzsáki G, Wang XJ. Mechanisms of gamma oscillations. Annual Review of Neuroscience. 2012;35:203–25.

34. Cheyne D, Ferrari P. MEG studies of motor cortex gamma oscillations: evidence for a gamma "fingerprint" in the brain? Frontiers in Human Neuroscience. 2013;7:575.

35. Church RM, Meck WH, Gibbon J. Application of scalar timing theory to individual trials. Journal of Experimental Psychology: Animal Behavior Processes. 1994;20(2):135–155.

36. Naud R, Houtman DB, Rose GJ, Longtin A. Counting on dis-inhibition: a circuit motif for interval counting and selectivity in the anuran auditory system. Journal of Neurophysiology. 2015;114:114: 2804–2815.

37. Rose GJ. The numerical abilities of anurans and their neural correlates: Insights from neuroethological studies of acoustic communication. Philosophical Transactions of the Royal Society B: Biological Sciences. 2017;373(1740).

38. Chamberland S, Timofeeva Y, Evstratova A, Volynski K, Tóth K. Action potential counting at giant mossy fiber terminals gates information transfer in the hippocampus. Proceedings of the National Academy of Sciences of the United States of America. 2018;115(28):7434–7439.

39. Loehr JD, Large EW, Palmer C. Temporal coordination and adaptation to rate change in music performance. Journal of Experimental Psychology: Human Perception and Performance. 2011;37(4):1292–1309.

40. Aschersleben G, Prinz W. Delayed Auditory Feedback in Synchronization. Journal of Motor Behavior. 1997;29(1):35–46.

41. Semjen A, Schulze HH, Vorberg D. Timing precision in continuation and synchronization tapping. Psychological Research Psychologische Forschung. 2000;63(2):137–147.

42. Repp BH. Processes underlying adaptation to tempo changes in sensorimotor synchronization. Human Movement Science. 2001;20(3):277–312.

43. Repp BH. Compensation for subliminal timing perturbations in perceptual-motor synchronization. Psychological Research Psychologische Forschung. 2000;63(2):106–128.

44. Golomb D, Wang XJ, Rinzel J. Propagation of spindle waves in a thalamic slice model. Journal of Neurophysiology. 1996;75(2).

45. Ermentrout B. An adaptive model for synchrony in the firefly Pteroptyx malaccae. Journal of Mathematical Biology. 1991;29:571–585.

46. Dunlap K. Reaction to rhythmic stimuli with attempt to synchronize. Psychological Review. 1910;17(6):399–416.

47. Delignières D, Lemoine L, Torre K. Time intervals production in tapping and oscillatory motion. Human Movement Science. 2004;23(2):87–103.

48. Ding M, Chen Y, Scott Kelso JA. Statistical analysis of timing errors. Brain and Cognition. 2002;48(1):98–106.

49. Miyake I. Researches on rhythmic action. Studies from the Yale Psychological Laboratory. 1902;10:1–48.

50. Woodrow H. The effect of rate of sequence upon the accuracy of synchronization. Journal of Experimental Psychology. 1932;15(4):357–379.

51. Drake C, Botte MC. Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. Perception & Psychophysics. 1993;54(3):277–286.

52. Grahn JA, Rowe JB. Feeling the beat: premotor and striatal interactions in musicians and non-musicians during beat perception. Journal of Neuroscience. 2009;29(23):7540–7548.

53. Nozaradan S, Peretz I, Missal M, Mouraux A. Tagging the neuronal entrainment to beat and meter. The Journal of Neuroscience. 2011;31(28):10234–40.

54. Vorberg D, Schulze HH. Linear phase-correction in synchronization: Predictions, parameter estimation, and simulations. Journal of Mathematical Psychology. 2002;46(1):56–87.

55. Edwards CJ, Alder TB, Rose GJ. Auditory midbrain neurons that count. Nature neuroscience. 2002;5(10):934–6.

56. Teki S, Grube M, Griffiths TD. A unified model of time perception accounts for duration-based and beat-based timing mechanisms. Frontiers in Integrative Neuroscience. 2012;5(90):1–7.

57. Hartcher-O'Brien J, Brighouse C, Levitan CA. A single mechanism account of duration and rate processing via the pacemaker-accumulator and beat frequency models. Current Opinion in Behavioral Sciences. 2016;8:268–275.

58. McAuley JD. Tempo and rhythm. In: Jones MR, Fay RR, Popper AN, editors. Music Perception. Springer; 2010. p. 165–199.

59. Ermentrout B, Terman DH. Mathematical Foundations of Neuroscience. Springer; 2010.

60. Knight BW. Dynamics of encoding in a population of neurons. The Journal of General Physiology. 1972;59(6):734–66.

61. Keener JP, Hoppensteadt FC, Rinzel J. Integrate-and fire models of nerve membrane response to oscillatory input. SIAM Journal on Applied Mathematics. 1981;41(3):503–517.

62. Simen P, Balci F, DeSouza L, Cohen JD, Holmes P. A model of interval timing by neural integration. Journal of Neuroscience. 2011;31(25):9238–9253.

63. Simen P, Vlasov K, Papadakis S. Scale (in)variance in a unified diffusion model of decision making and timing. Psychological Review. 2016;123(2):151–181.

64. Merchant H, Averbeck BB. The computational and neural basis of rhythmic timing in medial premotor cortex. The Journal of Neuroscience. 2017;37(17):4552–4564.

65. Simen P, Rivest F, Ludvig EA, Balci F, Killeen P. Timescale invariance in the pacemaker-accumulator family of timing models. Timing and Time Perception. 2013;1(2):159–188.

66. Grondin S. Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. Attention, Perception & Psychophysics. 2010;72(3):561–582.

67. Michon JA. Studies on subjective duration: I. Differential sensitivity in the perception of repeated temporal intervals. Acta Psychologica. 1964;22:441–450.

68. Collyer CE, Boatright-horowitz SS, Hooper S. A motor timing experiment implemented using a musical instrument digital interface (MIDI) approach. Behavior Research Methods, Instruments, & Computers. 1997;29(3):346–352.

69. Matell MS, Meck WH. Neuropsychological mechanisms of interval timing behavior. BioEssays. 2000;22(1):94–103.

70. Matell MS, Meck WH. Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. Cognitive Brain Research. 2004;21(2):139–170.

71. Coull J, Nobre A. Dissociating explicit timing from temporal expectation with fMRI. Current Opinion in Neurobiology. 2008;18(2):137–144.

72. Paton JJ, Buonomano DV. The neural basis of timing: Distributed mechanisms for diverse functions. Neuron. 2018;98(4):687–705.

73. Snyder JS, Large EW. Gamma-band activity reflects the metric structure of rhythmic tone sequences. Cognitive Brain Research. 2005;24(1):117–126.

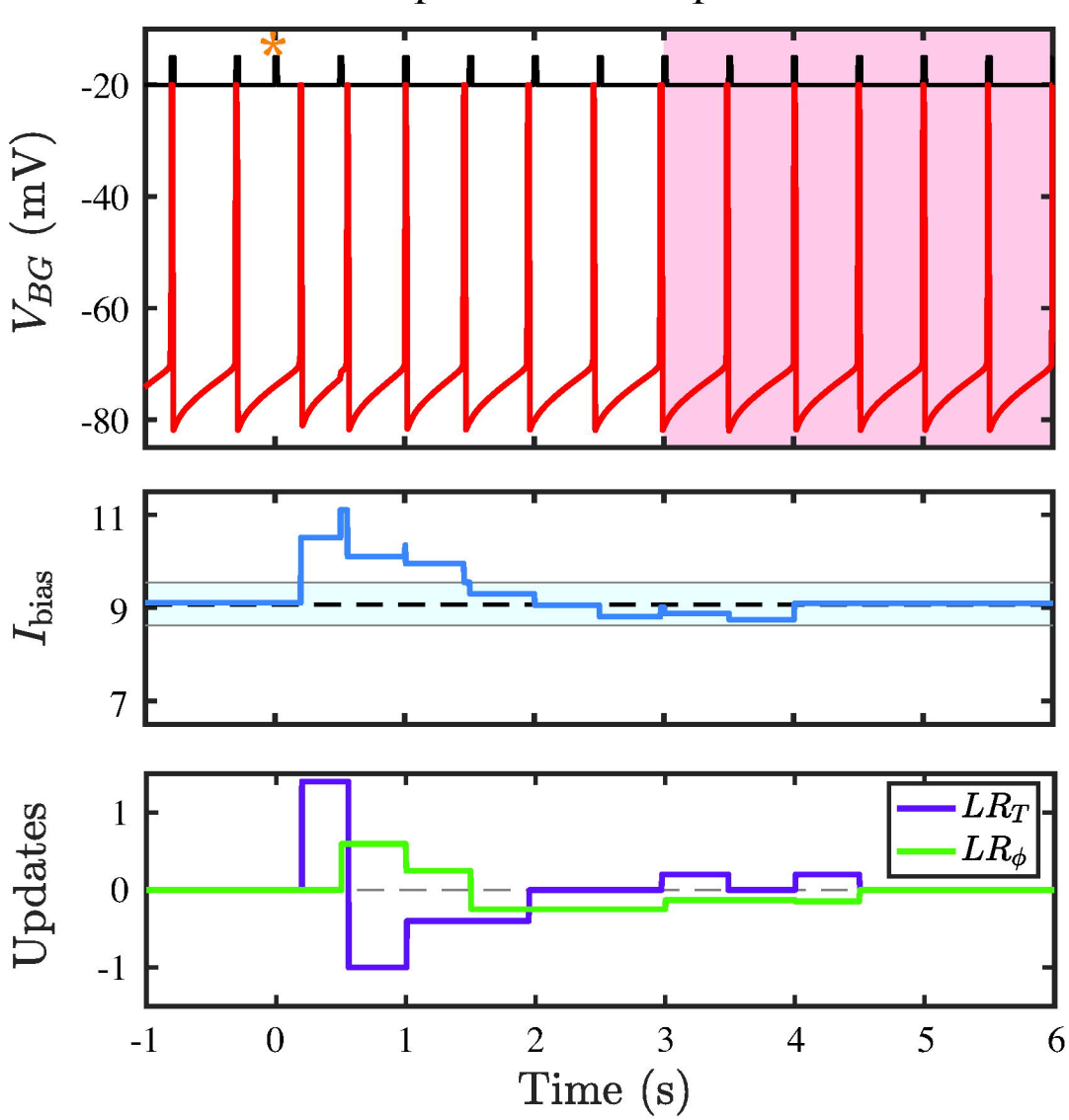
74. Bi GQ, Poo MM. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. The Journal of Neuroscience. 1998;18(24):10464–72.

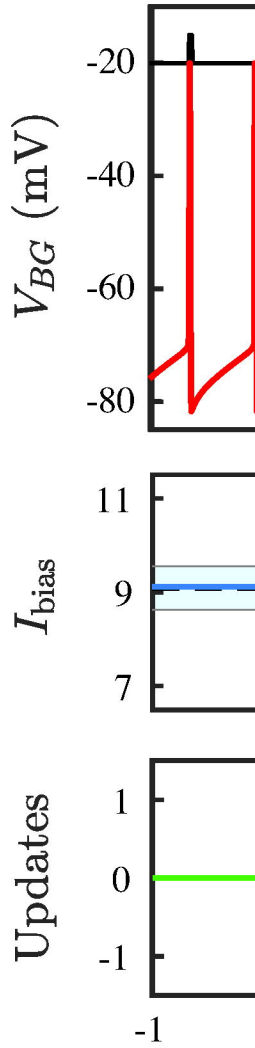
75. Yger P, Stimberg M, Brette R. Fast learning with weak synaptic plasticity. The Journal of Neuroscience. 2015;35(39):13351–62.

76. Vorberg D, Wing A. Modeling variability and dependence in timing. Handbook of Perception and Action. 1996;2:181–262.

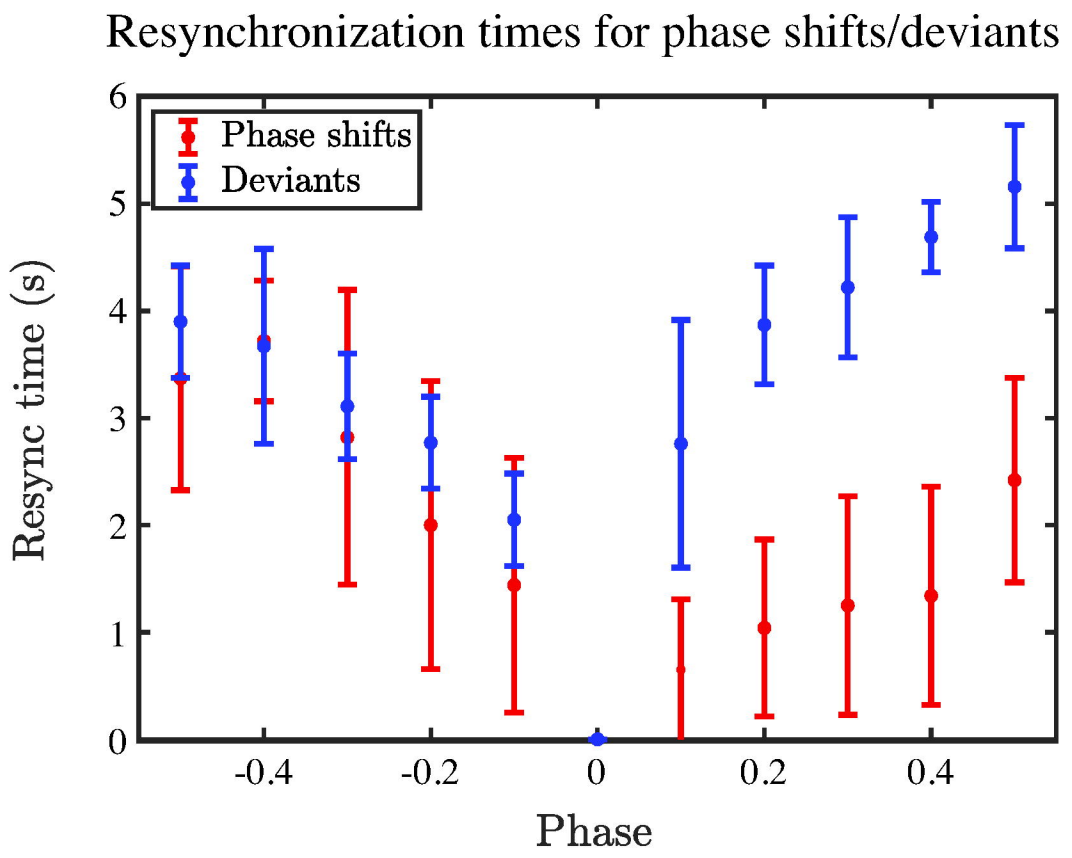
77. Patel AD, Iversen JR. The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. Frontiers in Systems Neuroscience. 2014;8:57.

78. Merchant H, Zarco W, Perez O, Prado L, Bartolo R. Measuring time with different neural chronometers during a synchronization-continuation task. Proceedings of the National Academy of Sciences. 2011;108(49):19784–19789.

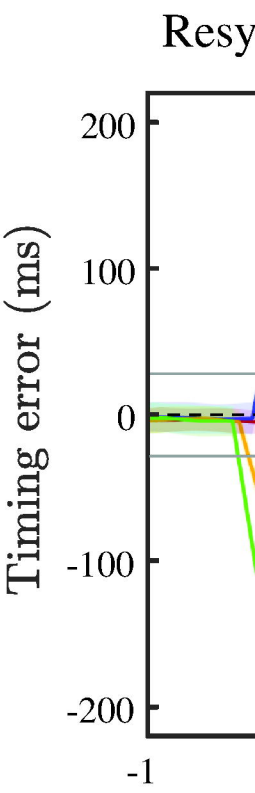

$V_{BG}$ (mV)
-20
-40
-60
-80
$I_{\text{bias}}$
11
9
7
Updates
1
0
-1
$LR_T$
$LR_\phi$
-1
0
1
2
3
4
5
6
Time (s)


$V_{BG}$ (mV)
-20
-40
-60
-80
$I_{\text{bias}}$
11
9
7
Updates
1
0
-1
-1

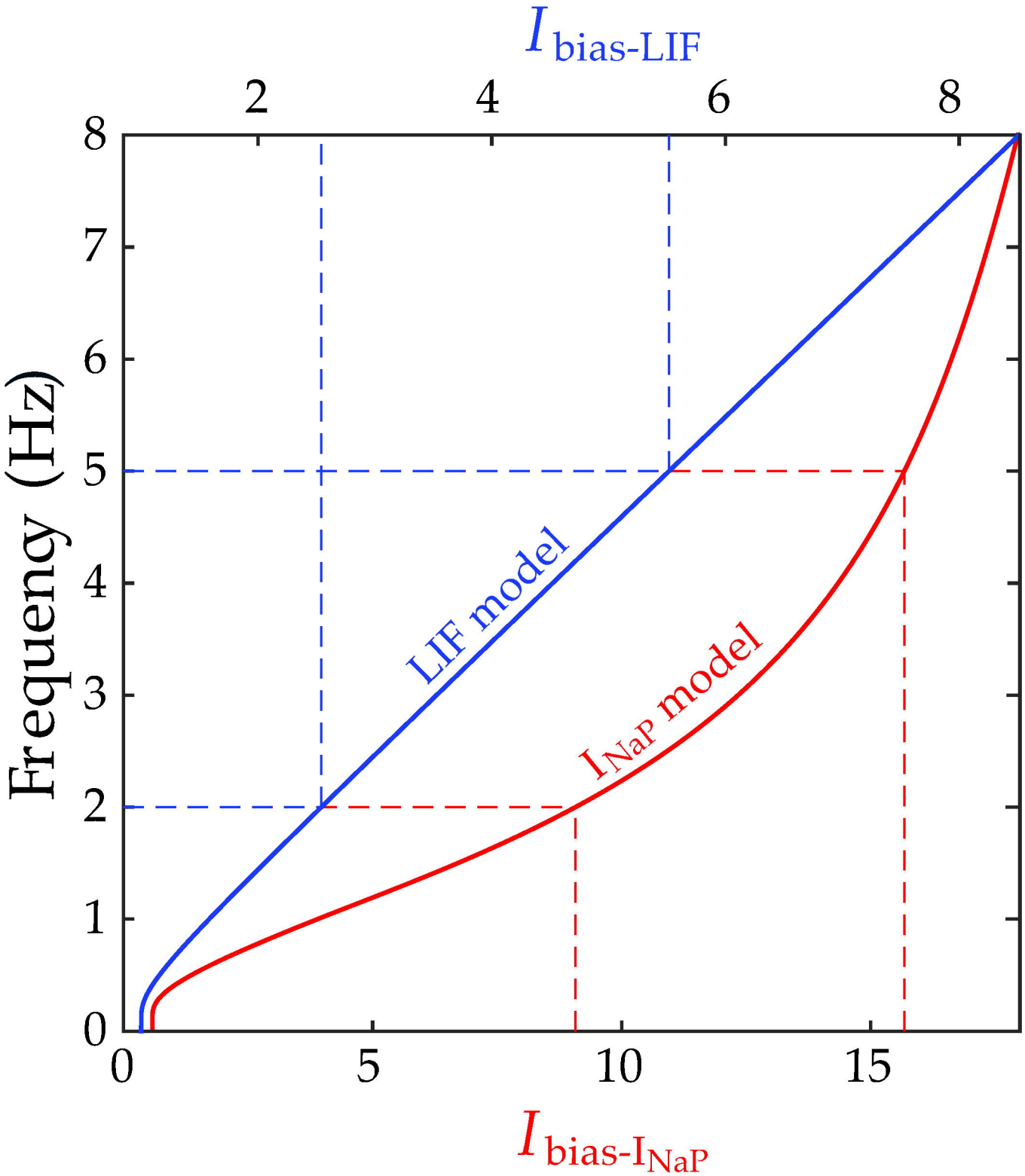
C

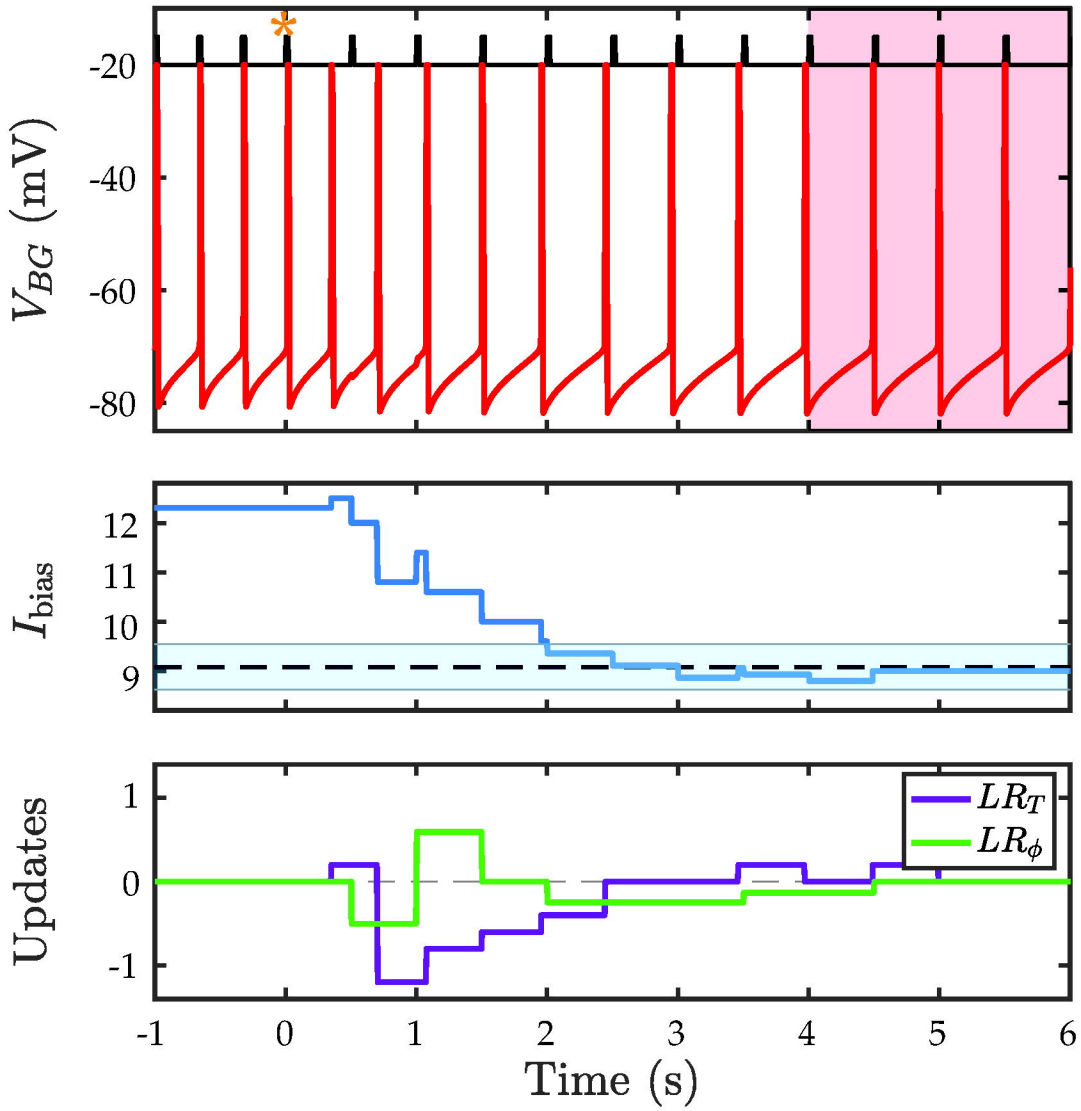

Resynchronization times for phase shifts/deviants
Phase shifts
Deviants
Resync time (s)
0
1
2
3
4
5
6
-0.4
-0.2
0
0.2
0.4
Phase

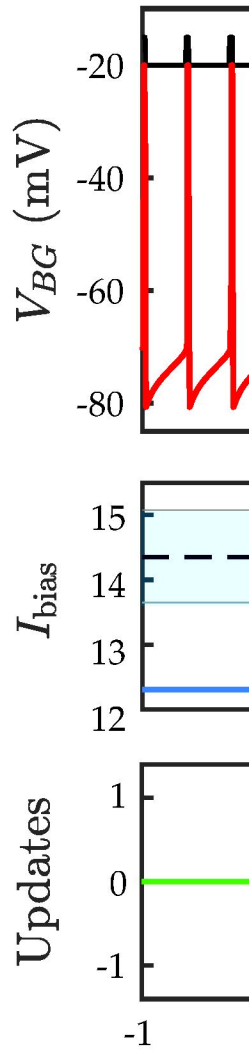
D


Timing error (ms)
200
100
0
-100
-200
-1

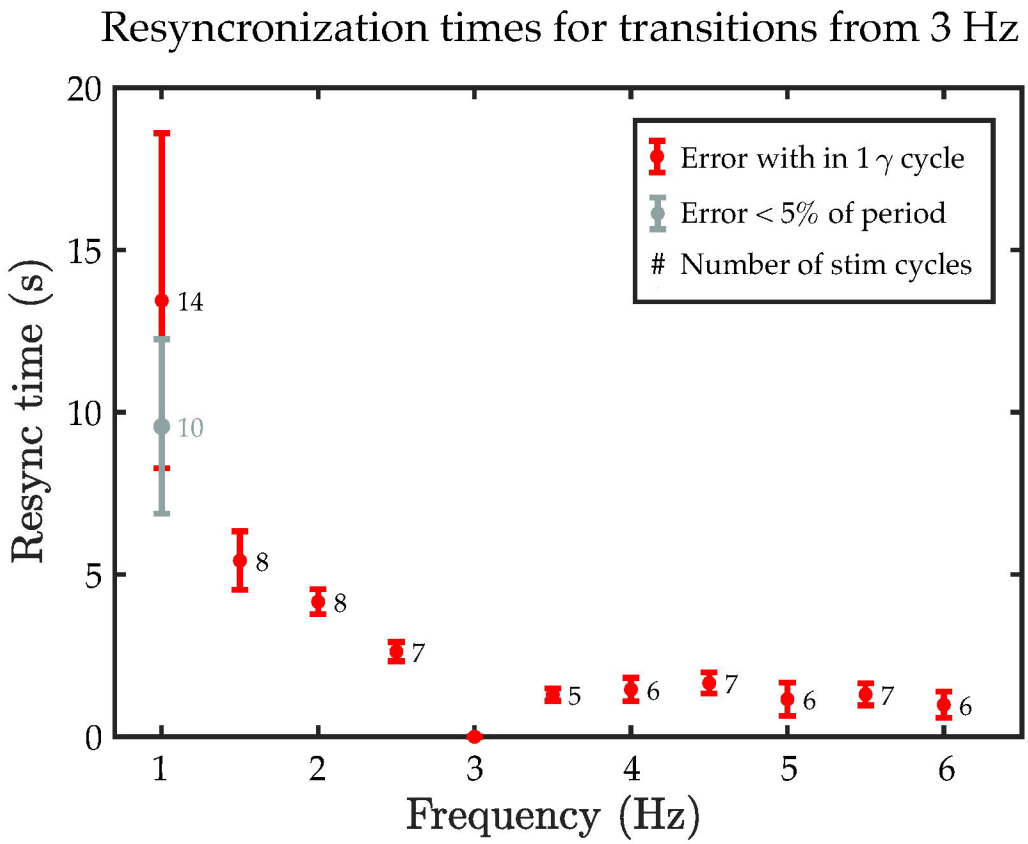
A


Frequency vs $I_{\text{bias}}$ for both models
$I_{\text{bias-LIF}}$
2
4
6
8
Frequency (Hz)
0
1
2
3
4
5
6
7
8
LIF model
$I_{\text{NaP}}$ model
0
5
10
15
$I_{\text{bias-I}_{\text{NaP}}}$

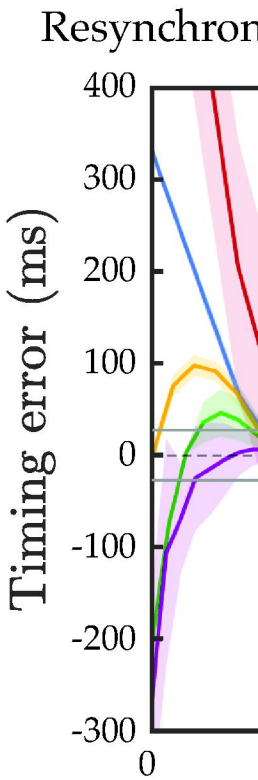

$V_{BG}$ (mV)
-20
-40
-60
-80
$I_{\text{bias}}$
12
11
10
9
Updates
1
0
-1
$LR_T$
$LR_\phi$
-1
0
1
2
3
4
5
6
Time (s)

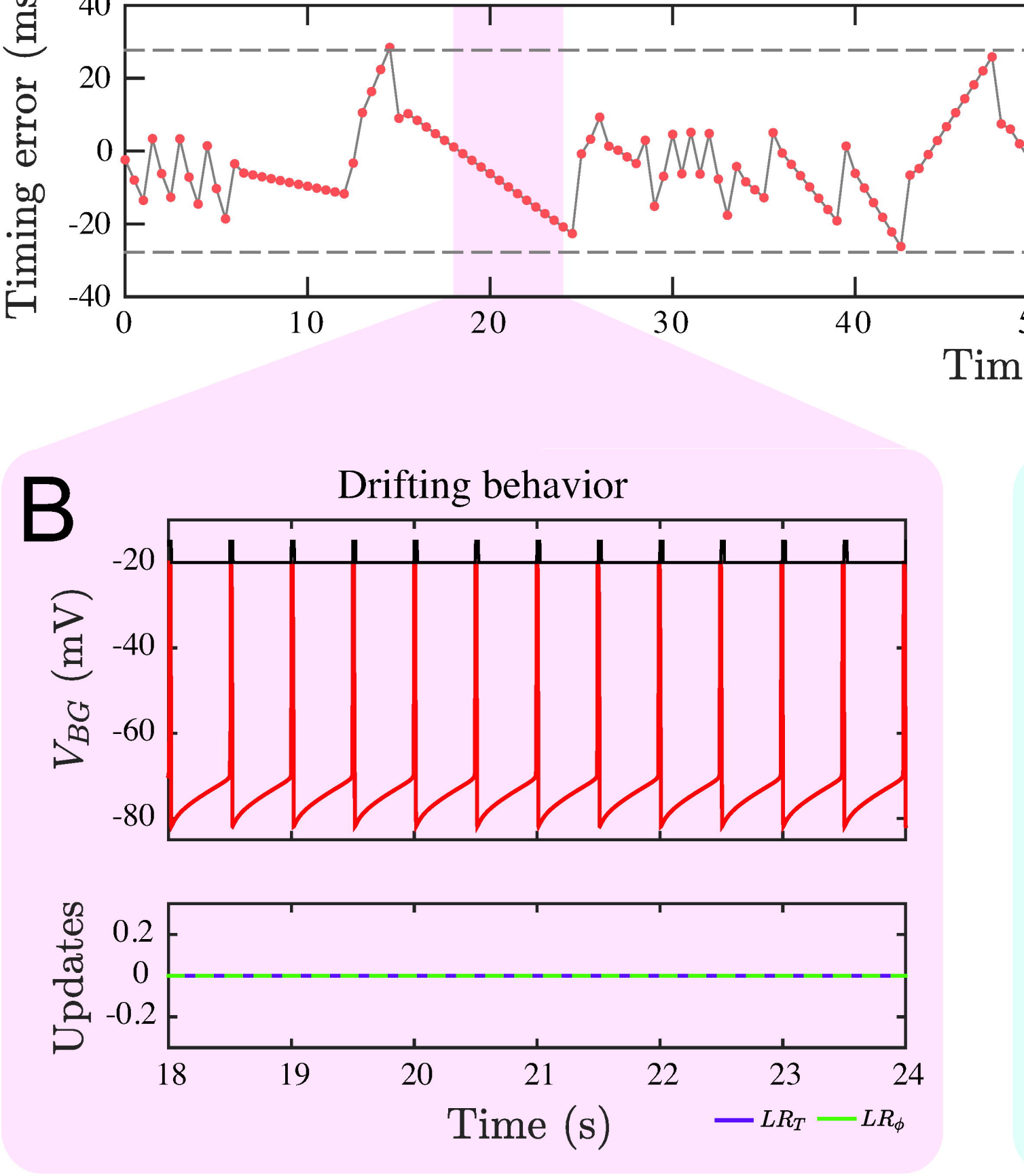

$V_{BG}$ (mV)
-20
-40
-60
-80
$I_{\text{bias}}$
15
14
13
12
Updates
1
0
-1
-1

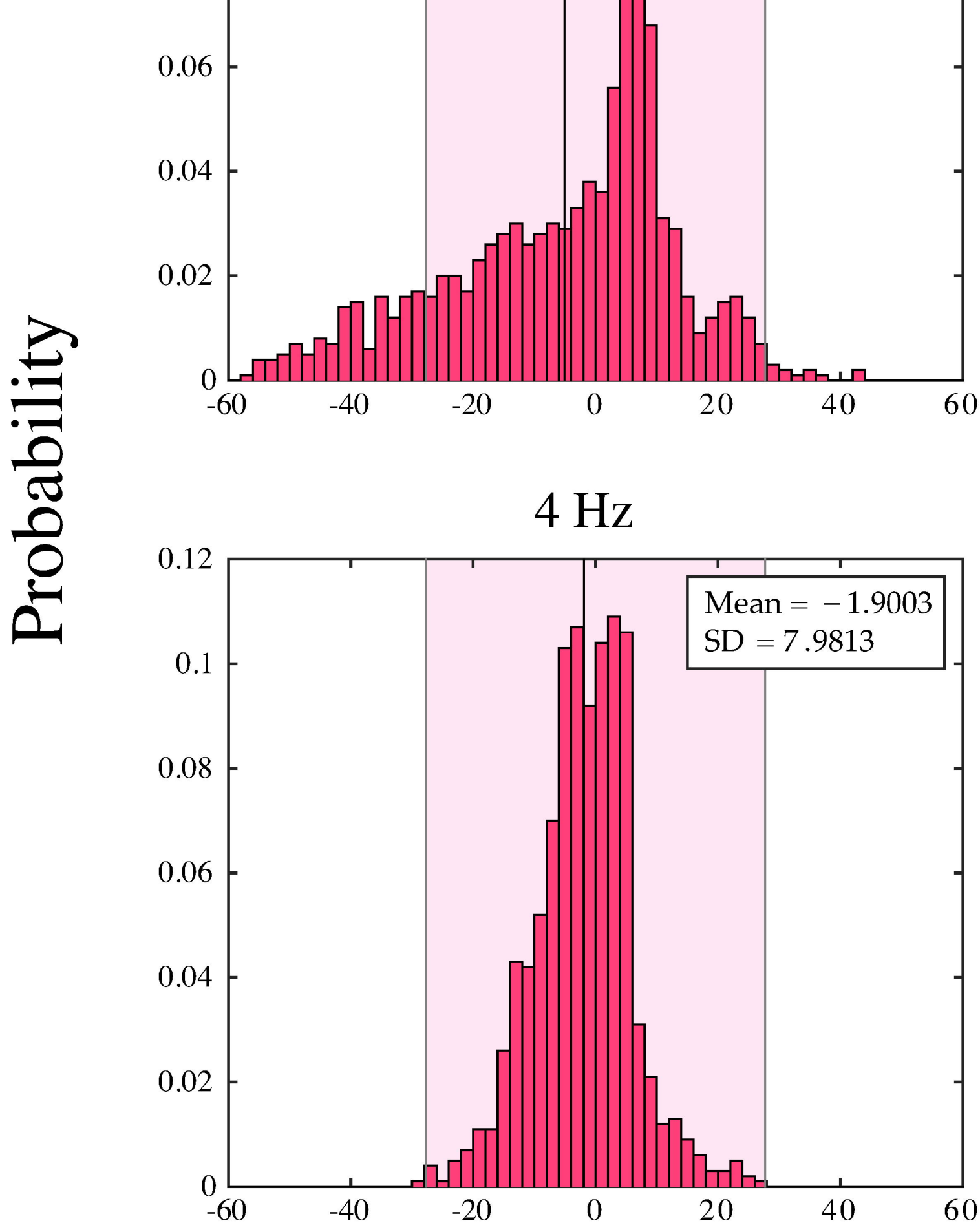
C

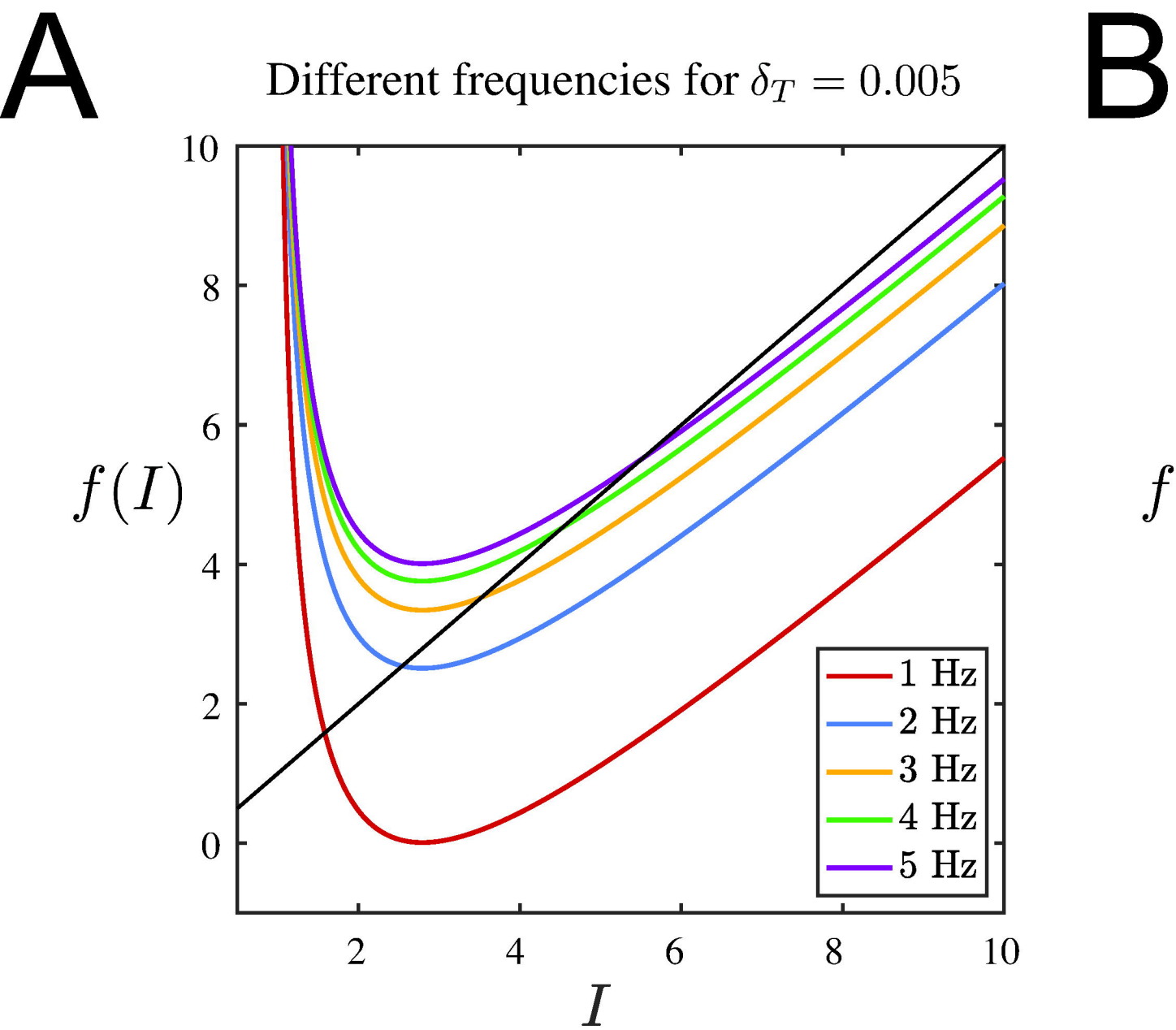

Resyncronization times for transitions from 3 Hz
Error with in 1 $\gamma$ cycle
Error < 5% of period
# Number of stim cycles
Resync time (s)
20
15
10
5
0
14
10
8
8
7
5
6
7
6
7
6
1
2
3
4
5
6
Frequency (Hz)

D


Resynchron
Timing error (ms)
400
300
200
100
0
-100
-200
-300
0

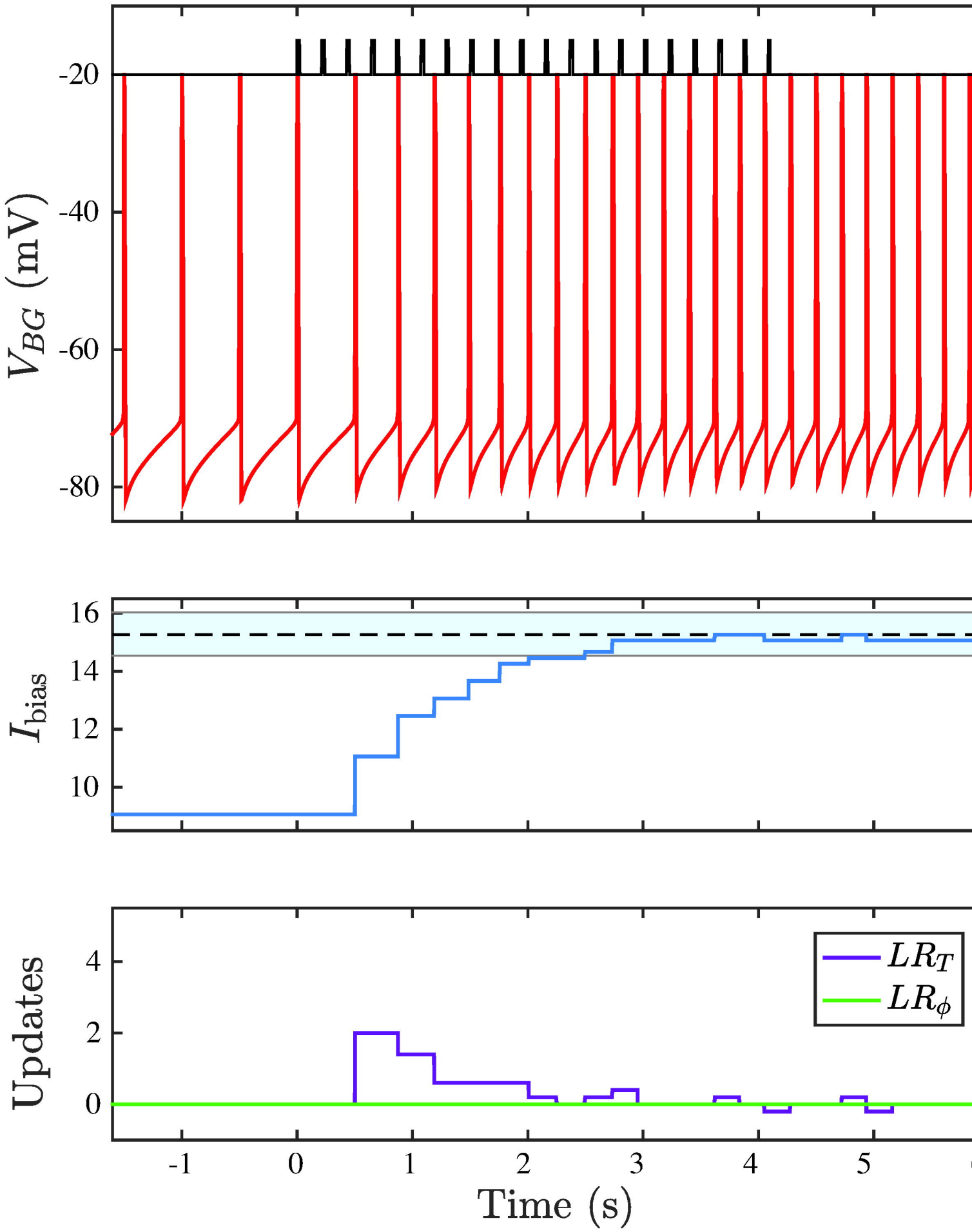

Timing error (ms)
40
20
0
-20
-40
0
10
20
30
40
Tim
B
Drifting behavior
$V_{BG}$ (mV)
-20
-40
-60
-80
Updates
0.2
0
-0.2
18
19
20
21
22
23
24
Time (s)
$LR_T$
$LR_\phi$

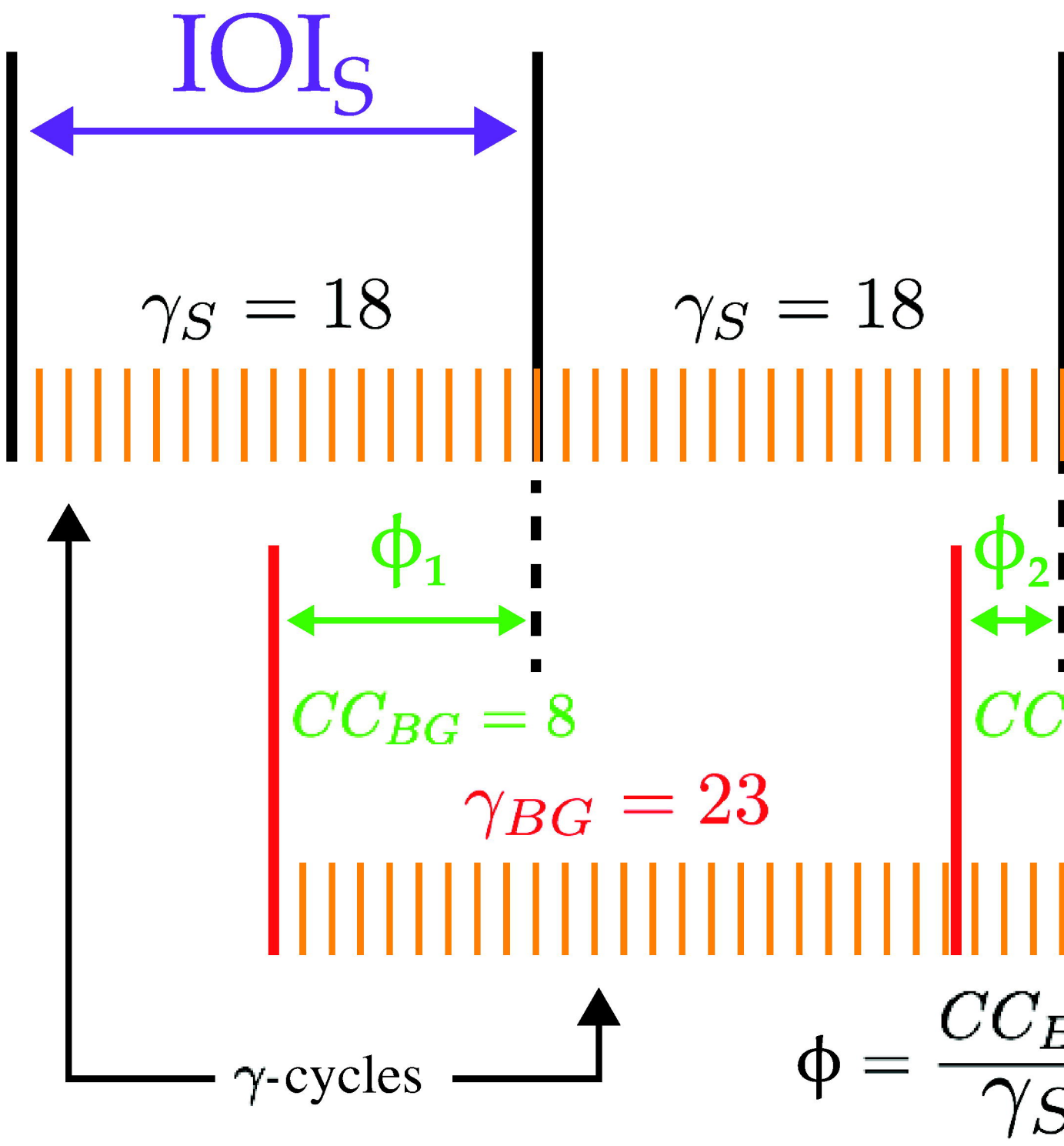

Probability
0.06
0.04
0.02
0
-60
-40
-20
0
20
40
60
4 Hz
0.12
0.1
0.08
0.06
0.04
0.02
0
Mean = −1.9003
SD = 7.9813


A
Different frequencies for $\delta_T = 0.005$
$f(I)$
$I$
1 Hz
2 Hz
3 Hz
4 Hz
5 Hz
B
$f(I)$

A

Learning an isochronous beat, period ($LR_T$) only



B

Explanation of the $\gamma$ counts