

# Linear summation of outputs in a balanced network model of motor cortex

Charles Capaday<sup>1,2\*</sup> and Carl van Vreeswijk<sup>1</sup>

<sup>1</sup> Center for Neurophysics, Physiology and Pathology, Centre National de la Recherche Scientifique UMR 8119, Université Paris-Descartes, Paris, France, <sup>2</sup> Department of Neurorehabilitation Engineering, University Medical Center Göttingen, Georg-August University, Germany

## OPEN ACCESS

### Edited by:

Sandro Romani,  
Howard Hughes Medical Institute  
Janelia Research Campus, USA

### Reviewed by:

Sen Song,  
Tsinghua University, China  
Alberto Bernacchia,  
Jacobs University Bremen, Germany

### \*Correspondence:

Charles Capaday,  
Department of Neurorehabilitation  
Engineering, University Medical Center  
Göttingen, Georg-August University,  
Von-Siebold-Str. 3, D-37075  
Göttingen, Germany  
charles.capaday@ccapcable.com

**Received:** 18 March 2015

**Accepted:** 13 May 2015

**Published:** 05 June 2015

### Citation:

Capaday C and van Vreeswijk C  
(2015) Linear summation of outputs in  
a balanced network model of motor  
cortex.  
*Front. Comput. Neurosci.* 9:63.  
doi: 10.3389/fncom.2015.00063

Given the non-linearities of the neural circuitry's elements, we would expect cortical circuits to respond non-linearly when activated. Surprisingly, when two points in the motor cortex are activated simultaneously, the EMG responses are the linear sum of the responses evoked by each of the points activated separately. Additionally, the corticospinal transfer function is close to linear, implying that the synaptic interactions in motor cortex must be effectively linear. To account for this, here we develop a model of motor cortex composed of multiple interconnected points, each comprised of reciprocally connected excitatory and inhibitory neurons. We show how non-linearities in neuronal transfer functions are eschewed by strong synaptic interactions within each point. Consequently, the simultaneous activation of multiple points results in a linear summation of their respective outputs. We also consider the effects of reduction of inhibition at a cortical point when one or more surrounding points are active. The network response in this condition is linear over an approximately two- to three-fold decrease of inhibitory feedback strength. This result supports the idea that focal disinhibition allows linear coupling of motor cortical points to generate movement related muscle activation patterns; albeit with a limitation on gain control. The model also explains why neural activity does not spread as far out as the axonal connectivity allows, whilst also explaining why distant cortical points can be, nonetheless, functionally coupled by focal disinhibition. Finally, we discuss the advantages that linear interactions at the cortical level afford to motor command synthesis.

**Keywords:** motor cortex, balanced networks, linear summation, neural transfer functions, rate models, spiking models

## Introduction

Microstimulation of a motor cortical point can evoke EMG output in several muscles. This can be quantified as a response vector. When two such points are simultaneously microstimulated, the evoked EMG vector is the linear sum of the EMG vectors evoked by microstimulation of each point on its own (Ethier et al., 2006). The result was unexpected and puzzling, given non-linearities in synaptic transmission and neural transfer functions. Additionally, the corticospinal transfer function—the relation between unit activity in the motor cortex (MCx) and motoneuron activation—is close to linear (Townsend et al., 2006; Capaday et al., 2013). This implies that the neural interactions within the motor MCx may be effectively linear.

However, another possible explanation is that the distance between paired points studied by Ethier et al. (2006) was greater than the distance over which they interact. Several experimental observations make this suggestion unlikely. The distances between pairs of points studied by Ethier et al. (2006) ranged between 0.66 and 5.7 mm (mean = 2.65 mm, SD = 1.52 mm). The axon collaterals of motor cortical neurons extend up to 6–7 mm away from their soma and are studded with synaptic boutons all along their course (Capaday et al., 2009). Spiking at a motor cortical point  $\sim 0.4$  mm in radius induces spiking in a surrounding area of  $\sim 1.5$  mm in radius (Capaday et al., 2011). Thus, two points up to 3 mm apart, or perhaps more, share an overlapping territory. It follows that the distance between cortical points studied by Ethier et al. (2006) was on average within that over which they do interact. Taken together, these considerations suggest the more interesting possibility. The MCx circuitry may be wired to produce linear interactions between cortical points (Capaday et al., 2013).

Balanced neural networks as originally proposed by van Vreeswijk and Sompolinsky (1996, 1998) involve a feedback dependent balance between synaptic excitation and inhibition such that, despite non-linear unit properties, the population output is a linear function of the input (Figure 1A). The population also responds to inputs with a time constant that is much shorter than that of the single units. In a balanced neural network the sum of the excitatory currents from external inputs, as well as from the activity of intrinsic circuit neurons, is balanced nearly exactly by the recurrent inhibitory currents. The basic idea of the balanced neural network is not unlike the principle used in operational amplifiers, where negative feedback of a portion of the output results in a device with linear input–output properties. The balanced network configuration also explains spike time variability. Spiking occurs at times when noise-like fluctuations generated by the dynamics of the circuitry exceed threshold. Spike timing is thus irregular and in fact asynchronous, even if the external input is temporally constant.

Here we present a model of MCx based on interconnected points that operate in a balanced state. The model explains

how simultaneously active cortical points having non-linear unit properties can sum their outputs linearly and we discuss how this may simplify the synthesis of motor commands. We have chosen a level of mathematical description which avoids detailed, uncertain, or unavailable physiological measurements. The emphasis, nonetheless, is on keeping a close correspondence between the Mathematics and the Neurophysiology so as to obtain mechanistic explanations and testable physiological predictions. The model also allows us to understand the consequences of controlling the response gain of a cortical point by disinhibition and the mechanism that limits the spatial spread of activity. We also show that conduction and synaptic time delays do not destabilize the asynchronous state—i.e., time delays do not lead to synchronized activity. This extends the principles of balanced network operation.

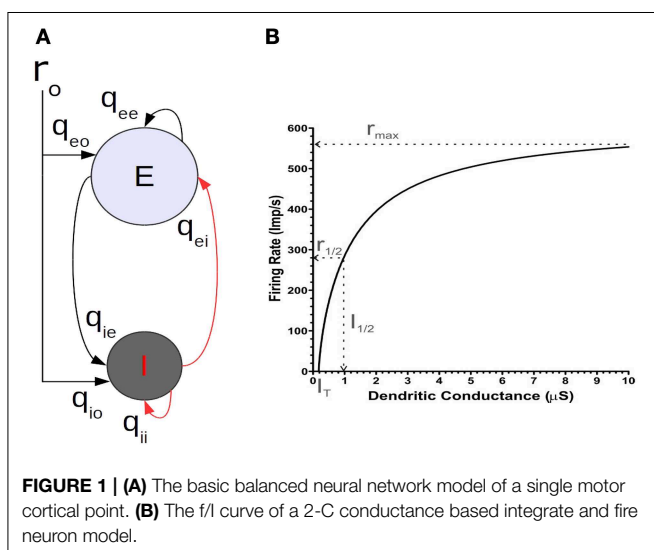
## Materials and Methods

There is considerable knowledge of the intrinsic, cable and repetitive firing properties of layer five neurons in the MCx (e.g., Crill and Schwindt, 1983; Stafstrom et al., 1984a,b). By contrast, little is known about the properties of neurons in other layers and in particular, those of inhibitory neurons. Similarly, the dynamics of synaptic transmission in the MCx *in vivo*, or *in vitro*, have not been investigated in sufficient detail. Here, we assume that synaptic transmission efficacy, that is the balance of synaptic facilitation and depression, is at steady-state over the time interval of movement duration (Abbott and Regehr, 2004). Taken together, the above considerations led us to use rate models as a simple description which allows us to understand the mechanisms of linear summation analytically. Rate models require fewer assumptions and many parameters, such as synaptic strengths, can be derived self-consistently (e.g., see Wilson, 1999). However, rate models cannot capture the fine structure of single unit spike timing and do not account for all types of spike synchronization. Therefore, we also simulated key aspects using networks of spiking model neurons.

In the models we will present, a cortical point consists of a population  $e$  of excitatory neurons and a population  $i$  of inhibitory neurons (Figure 1A). The two populations are reciprocally connected and each receives an external input representing a motor command. Additionally, each population is recurrently connected to itself. That is, the excitatory population is auto excited and the inhibitory one auto inhibited. Importantly, both neuron populations are driven by simultaneous excitatory and inhibitory currents, as has been shown in physiological experiments (e.g., Haider et al., 2006; Okun and Lampl, 2008). We will analyze the interactions of several such points that are fully interconnected by  $e$ -cell collaterals. As a first step, to uncover the basic mechanism by which the balanced network configuration leads to linearization of non-linear unit properties, we first consider in detail the operations of an isolated point.

## Basics of our Rate Model

We explain how temporal changes of firing rate depend on the input current. The firing rate  $r_a$  of a given neuron population



**FIGURE 1 | (A)** The basic balanced neural network model of a single motor cortical point. **(B)** The  $f/I$  curve of a 2-C conductance based integrate and fire neuron model.

$a$  ( $a = e$  for the excitatory population, or  $i$  for the inhibitory population) is given by

$$\tau_a \frac{dr_a}{dt} = -r_a + f_a(I_a), \quad (1)$$

where  $\tau_a$  is the time constant,  $I_a$  is the total input current and  $f_a(I_a)$  represents the frequency/current ( $f/I$ ) relation, or transfer function (e.g., see Shriki et al., 2003; Sompolinsky and White, 2004). We constrained the function  $f_a(I_a)$  of each population to that obtained *in vitro* from neurons of the cat motor cortex (Stafstrom et al., 1984b). This also constrains the range of synaptic connectivity strengths. Thus, the spike rate dynamics of a population of  $n$  recurrently connected neurons is reduced to a single system described by Equation (1) having the same effective synaptic strengths, and whose firing rate represents the population average. In our model, the activity of the excitatory neuron population represents the corticospinal output.

We considered the following two points to obtain the form of the  $f_a(I_a)$  relation. First, that the neurons constituting a given cortical point have an activation threshold and second, a non-linear transfer function. In this report, we use a transfer function motivated by our study of the neural mechanisms that allow modulation of a single neuron's firing rate gain by feedforward inputs (Capaday and van Vreeswijk, 2006). There we showed that, considering a two compartments (2-C) model, the firing rate gain can be modulated by a mixture of excitatory and inhibitory conductances in the dendritic compartment. Of equal importance, our analysis demonstrated that the transfer function of any spatially extended neuron model must be non-linear. To a first order approximation, the equation relating the synaptic input current at the soma  $I_s$  and dendrite  $I_d$  to the firing rate  $r$  of our 2-C model is given by

$$r = \frac{1}{C_m(V_T - V_r)} \left[ I_s + \frac{g_c}{g_c + g_D} I_D \right], \quad (2)$$

where  $C_m$  is the membrane capacitance,  $V_T$  the spiking threshold and  $V_r$  the resting potential. Note how the dendritic synaptic current  $I_D$  is attenuated by a factor  $g_c/(g_c + g_D)$ , where  $g_c$  is the conductance coupling the dendritic compartment to the somatic compartment and  $g_D$  is the dendritic conductance. Additionally, note that the firing rate given by Equation (2) is a decelerating function of the dendritic inputs (**Figure 1B**). We incorporate this effect in the  $f_a(I_a)$  relation of Equation (1) by a saturating transfer function, namely a first order Rushton–Naka function

$$f_a(I_a) = \left| \frac{(I_a - I_{T,a})}{I_{1/2,a} + (I_a - I_{T,a})} r_{max,a} \right|_+, \quad (3)$$

This function relates the firing rate to the input current, where  $I_{1/2,a}$  is the current required to reach the half-saturation rate  $r_{max,a}/2$  and  $I_{T,a}$  is the current threshold. Note the half rectification symbol  $| \cdot |_+$ , making  $f_a(I_a)$  a positive valued function. Following the result of Stafstrom et al. (1984b),  $r_{max,a}$  was set to 250 Spk/s for both populations. From our analysis of the 2-C model,  $I_{1/2,a}$ , the half-saturation current, was

determined to be approximately 25 Spk/s. For the numerical simulations we present here,  $I_{T,a}$  was set to zero because in the study by Ethier et al. (2006) all cortical points were stimulated at supra-threshold intensities. This choice is without loss of generality.

### Model of a Single Cortical Point

Referring to **Figure 1A**, the net current  $I_e$  driving the excitatory neurons is given by

$$I_e = q_{eo}r_o + q_{ee}r_e - q_{ei}r_i. \quad (4)$$

While the net current driving the inhibitory population satisfies

$$I_i = q_{io}r_o + q_{ie}r_e - q_{ii}r_i. \quad (5)$$

Here  $q_{ab}$  is the charge injected in population  $a$  per spike in population  $b$ . The  $q$ -values thus represent the synaptic weights and have units of *charge/spike*. Consequently, the product  $q_{ab}r_a$  is a synaptic current (*charge/s*). Note that these synaptic weights represent the combined effects of all neurons in population  $b$  on a typical neuron in population  $a$ . Now, also take into account that only a few tens of presynaptic spikes can bring the postsynaptic neuron to threshold. Accordingly, because cortical neurons receive inputs from thousands of others, the magnitude of the synaptic currents  $I_a$  ( $a = e$ , or  $i$ ) will be large relative to the half-saturation current  $I_{1/2,a}$ .

### Rate Model of Interacting Cortical Points

To understand quantitatively the interaction between motor cortical points we considered a fully interconnected three point model, with each point described by a balanced network (**Figure 5**). The connections between points are made by axon collaterals of the  $e$ -cell population at each point. The connections between the cortical points of the model correspond to the recurrent long range connections identified in the cat MCx (Capaday et al., 2009). Each point  $k$  receives an external input  $r_{o,k}$  ( $k = 1, 2$ , or  $3$ ). The output of the  $e$ -cell population at each point projects to the  $e$ -cell and  $i$ -cell populations of all other points. The inputs,  $I_{e,k}$  and  $I_{i,k}$ , to the excitatory and inhibitory cells, respectively, of point  $k$  are given by

$$I_{a,k} = q_{ao}r_{o,k} + q_{ae}r_{e,k} - q_{ai}r_{i,k} + \sum_{l \neq k} w_{a,kl}r_{e,l}. \quad (6)$$

where  $r_{o,k}$  is the rate of the external input to point  $k$ ,  $r_{a,k}$  is the rate of population  $a = e$ , or  $i$  at point  $k$ , and  $w_{a,kl}$  is the strength of the intracortical connections from the  $e$ -cells of point  $l$  to the neurons in population  $a$  at point  $k$ . The first three terms on the right hand side of Equation (6) correspond to the within point currents of Equations (4) and (5). The terms under the summation symbol are the inputs from the other points. The synaptic strengths  $w_{a,kl}$  decrease with the separation between points, in keeping with the monotonic decrease of the density of synaptic boutons projecting out of a motor cortical point (Capaday et al., 2009). By contrast, the local (i.e., within point) synaptic connection strengths (e.g.,  $q_{ee}$ ,  $q_{ie}$ , etc.) are identical for all points. Importantly, inhibition remains local, that is the

*i*-cells do not project to other cortical points. This is also in keeping with a generally agreed upon organizational principle of cortical architecture (e.g., Bacci et al., 2005; Markram, 2010; Fino et al., 2013), notwithstanding that some inhibitory interneurons may have longer range projections (e.g., Tomioka et al., 2005). In addition, as previously suggested (Schneider et al., 2002) each point contains a population of what we term master inhibitory neurons (*MI*-neurons) whose function is to inhibit the local *i*-cells that feedback inhibit the local *e*-cells (Figure 5). These *MI*-neurons control, as we will show, the firing rate gain of a cortical point. The presence of specific *MI*-neurons, specialized to innervate other GABAergic interneurons, have been identified in the hippocampus (Freund and Antal, 1988; Acsády et al., 1996). Connections between various classes of inhibitory neurons is also a prominent feature in the neocortex (e.g., Fino et al., 2013). In particular, the recent finding that somatostatin expressing interneurons of the visual cortex do not inhibit each other, yet strongly inhibit all other inhibitory populations (Pfeffer et al., 2013), supports well the idea of *MI*-neurons. The addition of *MI*-neurons modifies the inhibitory feedback current onto the *e*-cell population by scaling the transfer function of the *i*-cell population according to the gain control mechanism described above. Thus, the maximal firing rate of the *i*-cell population  $r_{max,i}$  in Equation (3) was simply scaled by a factor  $G_{i-cell}$ , where  $0 \leq G_{i-cell} \leq 1$ .

### Spiking Model of Interacting Points with Delays

Rate models capture well the dynamics of neuronal populations, as long as the units within a population do not synchronize and the rate changes are not too rapid. For the single cortical point model rapid rate changes are not an issue because we are concerned with steady-state firing rates. However, our rate analysis cannot accurately determine whether the interaction between cortical points induces synchronous activity because of conduction delays between the points. Furthermore, as previously explained, rate models cannot capture the fine structure of single unit spike timing and do not account for all types of spike synchronization (e.g., see Roxin et al., 2005). To investigate these issues we did simulations of a network consisting of multiple points each of which contains populations of excitatory and inhibitory spiking neurons. The standard approach here would be to use networks of single compartment model neurons, either integrate and fire, or conductance based. But the  $f/I$  relationship of such cells is close to linear, particularly at higher rates, as already discussed. Thus, linearization by feedback via recurrent inhibitory connections is not meaningful in such networks. Consequently, we modeled the excitatory neurons as 2-C model neurons, consisting of a passive dendrite coupled to a soma having integrate and fire dynamics following Capaday and van Vreeswijk (2006). As shown there, these simple model neurons have strongly non-linear input-output curves when the input is a synaptic conductance change at the dendritic compartment, or at both the dendritic and somatic compartment. Recall that the firing rate of these 2-C model neurons is approximately given by Equation (2). Inhibitory interneurons tend to be electrotonically more compact. Accordingly we model

the inhibitory cells as single compartment integrate and fire neurons.

Each cortical point consists of  $N_e = 8000$  excitatory and  $N_i = 2000$  inhibitory model neurons. All excitatory and inhibitory neurons in point  $k$  receive an external excitatory input proportional to  $r_{o,k}$ , the firing rate of the external population that projects to the point. The neurons also receive recurrent inputs from, on average, 1000 excitatory and 1000 inhibitory neurons randomly chosen from the same point (van Vreeswijk and Sompolinsky, 1998). Additionally, the neurons receive inputs from excitatory cells in other points. The excitatory cells in point  $k$  receive inputs from, on average,  $K_e(k, l)$  *e*-cells in point  $l$ , while for the inhibitory neurons that number is  $K_i(k, l)$ . This replicates the network connectivity of the rate model. To account for the finite conduction speed, these inputs arrive with a delay  $\Delta_{kl}$ . The distance between adjacent cortical points was 1 mm, consequently the conduction delay between points was set at 10 ms, as experimentally determined (Capaday et al., 2011).

The parameters of the 2-C excitatory neurons were based on electrophysiological measurements of cat layer five pyramidal cells *in vitro* (Stafstrom et al., 1984a,b). The somatic and dendritic membrane conductances were 0.06 and 0.12  $\mu\text{S}$ , respectively. The coupling conductance between the two compartments was 0.06  $\mu\text{S}$ . The somatic and dendritic membrane capacitances were 0.175 and 0.525 nF. The inhibitory neurons were modeled as single compartments having an input conductance of 0.1  $\mu\text{S}$  and a membrane capacitance of 0.7 nF. For both model neurons the threshold was set at  $-60$  mV, while the resting and reset potentials were set at  $-75$  mV.

The synaptic inputs were modeled as instantaneous conductance changes. For the *e*-cells, 80% of inputs of all types were on the dendritic compartment. The reversal potentials of the excitatory and inhibitory synapses were set at 0 and  $-85$  mV, respectively. A spike from the external inputs causes conductance changes of 10.66 and 1.18 nS in the *e*-cells and *i*-cells, respectively. Spikes from the *e*-cells in the network induce conductance changes of 5.33 and 1.18 nS, respectively, in the *e*-cell and *i*-cell population to which they are connected. The *i*-cells produce a conductance change per spike of 56.0 nS in the *e*-cells and 8.30 nS in the *i*-cells.

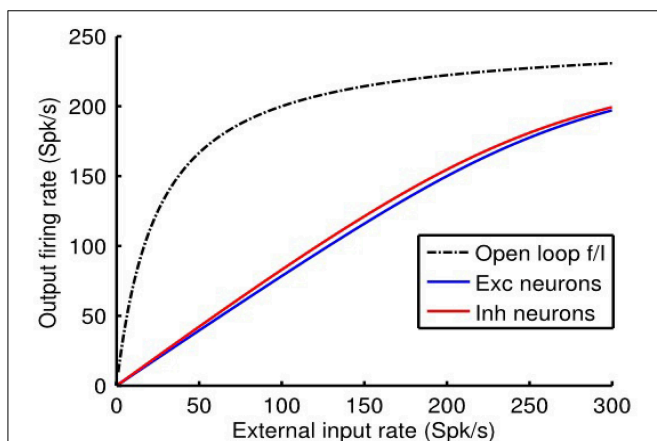
## Results

The results are presented as follows. In the first section we show that despite the non-linear  $f/I$  curves in the dynamic equations describing a single cortical point, the balanced network configuration produces nearly linear input-output curves. The reasons are explained analytically in the subsequent section. In the third section we demonstrate that the interactions between synaptically connected cortical points remain nearly linear for any combination of external or intracortical inputs, even when the strength of the local feedback inhibition is reduced. The response gain of a cortical point is also increased in the latter condition. Additionally, we show that by such an increase of its response gain, a cortical point which receives little or no effective input from a distant point can, nonetheless, be functionally coupled to it. An analytical treatment of the three point model

explaining the mechanism of linear summation and its resilience to reduction of feedback inhibition is presented in the following section. Lastly, we demonstrate with the spiking model that conduction delays between cortical points does not lead to spike synchrony, or oscillations.

### Input–Output Characteristics of the Single Cortical Point Model

We reiterate the fact that when the *e*-cell and *i*-cell populations are uncoupled (i.e.,  $q_{ie} = q_{ei} = 0$ ) and in the absence of any auto-feedback (i.e.,  $q_{ee} = q_{ii} = 0$ ) the open-loop  $f_a(I_a)$  curve of each population is non-linear (Figure 1B), following the Rushton–Naka transfer function Equation (3). Figure 2 shows how the steady-state output of the balanced network described by Equation (6) is related to the external input. The dashed curve is the open-loop  $f_a(I_a)$  curve of the *e*-cell, or *i*-cell, population shown for reference. The salient outcome resulting from the dynamic operation of the balanced network is that the  $f_a(I_a)$  curve of each population becomes markedly linear (Figure 2). The gain is essentially constant over the range of external inputs, but considerably lower at low rates (<50 Spk/s) compared to the open-loop gain. Setting all inhibitory connection strengths to zero (i.e.,  $q_{ei} = q_{ii} = 0$ ) results in a saturated near maximal output, even for a minimal input (Figure 3). For the *e*-cell population this is due to the strong recurrent excitatory feedback  $q_{ee}$  unchecked by inhibition. For the *i*-cell population the effect is due to the resulting strong drive by the *e*-cell population and the absence of inhibitory auto-feedback. The model thus captures the instability characteristic of focal ictal activity produced by application of GABA<sub>A</sub> receptor antagonists (e.g., Capaday et al., 2011). This observation and the fact that focal ictal activity can occur in an isolated cortical slab strongly support the idea that intracortical recurrent excitation must be strong and balanced by commensurate inhibition. The model is thus in the class of inhibition-stabilized networks (Vogels et al., 2005).



**FIGURE 2 | Response of the single cortical point balanced network model to external input.** Note the near linear  $f(I)$  curves of the *e*-cell and *i*-cell populations. By contrast, the dashed black curve depicts the non-linear response of the output *e*-cell population in the open loop mode. Model parameter values were  $q_{ee} = 0.67$ ,  $q_{ei} = -1.7$ ,  $q_{ie} = 1$ , and  $q_{ii} = -2$ .

In summary, the dynamics of the balanced network linearized a non-linear system. The mechanisms underlying this transformation are dealt with in the following section.

### Analysis of the Single Cortical Point Model

Because the transfer functions of the excitatory and inhibitory populations are non-linear we would not, *a priori*, expect them to respond linearly with the external input. However, as we will show, the negative feedback between the populations and, importantly, the auto-inhibition of the inhibitory population produces linearization of the response. If the external population that drives the point fires with a constant rate  $r_o$ , then after a transient, the rates of the excitatory and inhibitory populations reach a constant value. At this fixed point, the rates satisfy the following equations

$$r_e = \left| \frac{I_e - I_{T,e}}{I_{1/2,e} + I_e - I_{T,e}} r_{max,e} \right|_+, \tag{7}$$

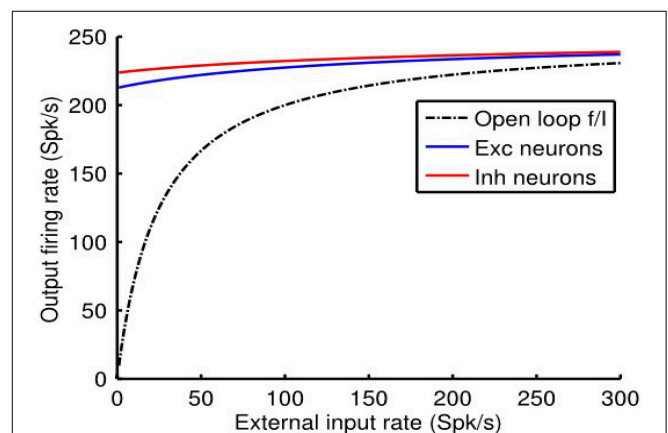
$$r_i = \left| \frac{I_i - I_{T,i}}{I_{1/2,i} + I_i - I_{T,i}} r_{max,i} \right|_+,$$

where the inputs,  $I_e$  and  $I_i$  are given by

$$I_e = q_{eo}r_o + q_{ee}r_e - q_{ei}r_i \text{ and } I_i = q_{io}r_o + q_{ie}r_e - q_{ii}r_i. \tag{8}$$

One can actually solve the set of Equation (7) analytically for  $r_e$  and  $r_i$ , but this involves finding the roots of a quartic equation whose form is not very revealing. Instead we will derive an approximate solution, which much more clearly explains how the linearization comes about. It also informs us on the conditions which the network parameters have to satisfy to get linearization of the response.

Under normal physiological conditions, the rates  $r_e$  and  $r_i$  should not be close to their maximum value for reasonable external input rates  $r_o$ . Looking at Equation (7) we can deduce that this means that  $I_e$  and  $I_i$  should be small compared to  $I_{1/2,e} + I_{T,e}$  and  $I_{1/2,i} + I_{T,i}$ , respectively. However, the synaptic



**FIGURE 3 | Without feedback inhibition the cortical point is unstable, firing at near maximal rate for a minimal input.** Model parameter values as in Figure 2, except that  $q_{ei} = 0$  and  $q_{ii} = 0$ .

strengths  $q_{ab}$  are large. Consequently, the total external input  $q_{eo}r_o$ , the total recurrent excitation  $q_{ee}r_e$  and the total recurrent inhibition  $-q_{ei}r_i$  of the excitatory population will all be much larger in magnitude than  $I_{1/2,e} + I_{T,e}$ . The only way in which the net input  $I_e$  can be sufficiently small is if the recurrent inhibition approximately cancels the feedforward input and recurrent excitation of the excitatory population. Thus, the rates should be such that

$$q_{eo}r_o + q_{ee}r_e \approx q_{ei}r_i \tag{9}$$

Using the same argument for the inputs into the inhibitory population, we obtain a second equation that the rates should satisfy

$$q_{io}r_o + q_{ie}r_e \approx q_{ii}r_i \tag{10}$$

So we have two requirements that have to be approximately satisfied. Combining the preceding two equations gives the approximate solutions for  $r_e$  and  $r_i$ , namely

$$r_e \approx \frac{q_{eo}q_{ii} - q_{io}q_{ei}}{q_{ei}q_{ie} - q_{ee}q_{ii}}r_o = A_e r_o, \quad \text{and} \quad r_i \approx \frac{q_{eo}q_{ie} - q_{io}q_{ee}}{q_{ei}q_{ie} - q_{ee}q_{ii}}r_o = A_i r_o. \tag{11}$$

For a positive external input  $r_o > 0$ ,  $r_e$  and  $r_i$  should be positive. Thus, the weights  $q_{ab}$  should be such that  $A_e$  and  $A_i$  are positive. The derivation makes the point that even though the transfer functions of the two populations are non-linear, the recurrent inputs make the response of the two populations approximately linear with the external input rate. In fact, in this approximation these rates do not depend on the properties of the open-loop transfer function at all, but only on the synaptic weights,  $q_{ab}$ . This is because the large  $q$ -values force the network to an operating point where the total external input and excitatory recurrent feedback are roughly balanced by the recurrent inhibition in both the excitatory and inhibitory populations. In other words, at the operating point  $I_e$  and  $I_i$  will be dynamically zero. Additionally, Equations (11) imply that if we increase all  $q_{ab}$ -values by the same factor (i.e., in proportion), the behavior of the differential Equations (1) does not change, but it can be shown that the linearization is improved.

We can use the fact that at the operating point of Equations (1)  $I_e$  and  $I_i$  are  $\approx 0$  to derive their linear approximation. To do this we calculate the Jacobian matrix  $Df$  of the system of Equations (1) evaluated at  $I_e$  and  $I_i \approx 0$ , the result is

$$Df = \begin{bmatrix} \frac{r_{max,e}q_{ee}}{\tau_e I_{1/2,e}} & \frac{-r_{max,e}q_{ei}}{\tau_e I_{1/2,e}} \\ \frac{r_{max,i}q_{ie}}{\tau_i I_{1/2,i}} & \frac{-r_{max,i}q_{ii}}{\tau_i I_{1/2,i}} \end{bmatrix}. \tag{12}$$

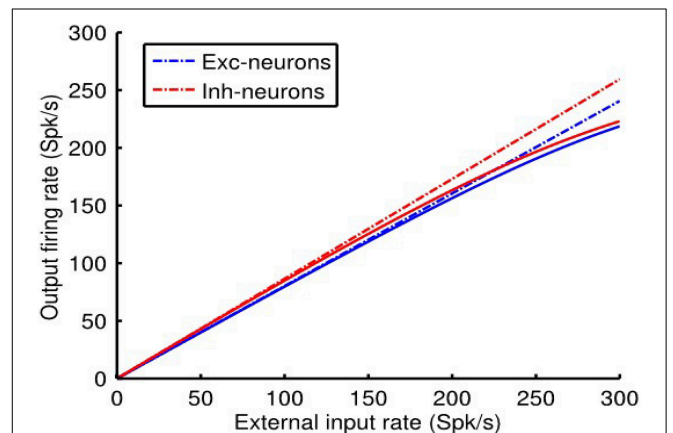
Inspection of the Jacobian shows that the linear approximation has same synaptic weights as the non-linear system, except for a scaling factor  $\gamma_a = r_{max,a}/\tau_a I_{1/2,a}$  ( $a = e, \text{ or } i$ ). We can thus write the linear equivalent system as,

$$\begin{aligned} \tau_e \frac{dr_e}{dt} &= -r_e + \gamma_e f(I_e), \\ \tau_i \frac{dr_i}{dt} &= -r_i + \gamma_i f(I_i), \end{aligned} \tag{13}$$

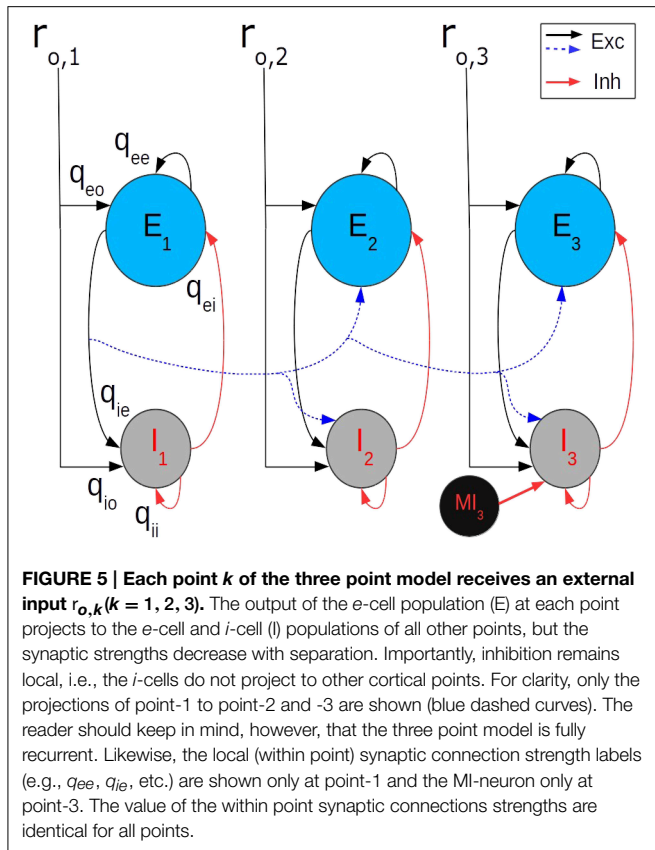
where  $f(I_e)$  and  $f(I_i)$  are the linear functions given by Equations (8).

Referring to **Figure 4**, it can be seen that the linear approximation closely follows the output of the non-linear balanced network model over the entire range of external inputs. As expected, however, there is a deviation at high firing rates because the linear approximation is not constrained by  $r_{max,a}$ , whereas the non-linear system is. That is, in the linear approximation  $r_{max,a}$  directly determines the scale factor  $\gamma_a = r_{max,a}/I_{1/2,a}$ . In the non-linear case, however, it is limiting because of the form of the Rushton–Naka function.

As a last point on the analysis of the single cortical point model we explain the conditions necessary to maintain stability and robustness. For the cortical point model to be stable, the eigenvalues  $\lambda_1$  and  $\lambda_2$  of the Jacobian matrix must be negative. By considering the trace and determinant of the Jacobian matrix, it can be shown that stability requires  $q_{ie}q_{ei} > q_{ee}q_{ii}$  and that,  $\gamma_i q_{ii} > \gamma_e q_{ee}$ . In words, the product of the cross-feedback terms must be greater than the product of the auto-feedback terms and auto-inhibition should be sufficiently large compared to auto-excitation. This can be understood in neurophysiological terms as follows. Consider, for example, that  $q_{ee}$  is very strong, this will drive the  $e$ -cells toward firing rate saturation. If  $q_{ie}$  is weak, the  $i$ -cell population will not increase its firing rate sufficiently to counteract the high activity level of the  $e$ -cells, unless  $q_{ei}$  is very strong. This explains why it is the product  $q_{ie}q_{ei}$  which is important. If instead,  $q_{ii}$  is strong, then again the  $e$ -cell population will not receive sufficient inhibition unless either  $q_{ie}$ , or  $q_{ei}$  are strong. The two terms are once more linked in a multiplicative way. Note that this first requirement insures that the network activity is not oscillatory (i.e., the two populations firing in alternating bursts). The second stability requirement that  $\gamma_i q_{ii} > \gamma_e q_{ee}$  insures that the feedback inhibition of the  $e$ -cell population remains linear over the range of inputs. This prevents the activity level of either population from saturating.



**FIGURE 4 | Comparisons of linearization produced by the dynamics of the non-linear balanced network model vs. that of the linear system approximation.** Further details in the text. Model parameter values were  $q_{ee} = 0.67$ ,  $q_{ei} = -1.7$ ,  $q_{ie} = 1$ , and  $q_{ii} = -2$ .

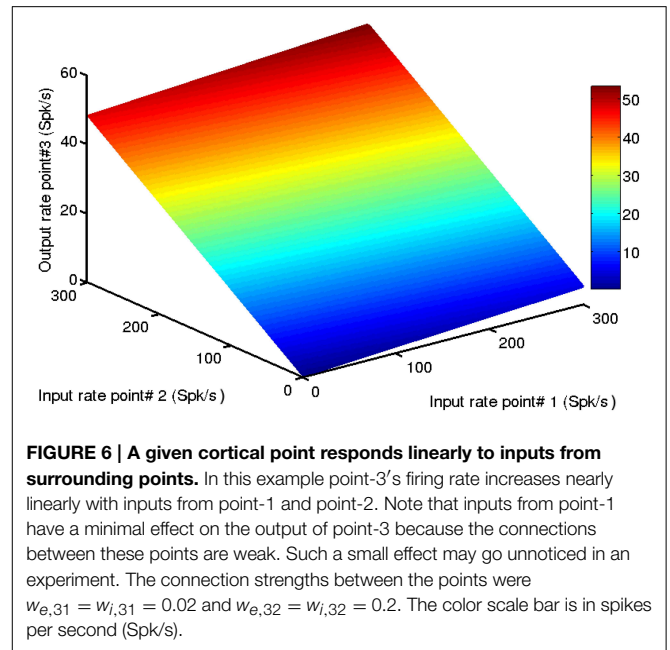


Summarizing, the cortical point model responds nearly linearly to external inputs because the auto-inhibition of the inhibitory population results in a linear transfer function for this population and consequently linear feedback inhibition of the excitatory population.

### Input–Output Characteristics of Coupled Cortical Points

The three point model afforded us the opportunity to determine the details of linear summation of simultaneously active cortical points and how this leads to linear summation of the evoked EMG vectors. It also allowed us to understand the consequences of controlling the response gain of a cortical point by disinhibition and the mechanism that limits spatial spread of activity. We begin with the details of linear summation.

**Figure 6** shows the effects of activity at point-1 and point-2 on the output of point-3. Importantly, in this example point-3 is not driven by a direct external input, only by the activity of points with which it is connected. The firing rate response of point-3, to any combination of inputs from the other two points, lies on a nearly perfect plane. This demonstrates, as should be expected from the results of the previous section, that a given cortical point responds linearly to intracortical inputs (i.e., to activity from cortical points with which it is connected), in addition to external inputs. What may be less evident is that, because the long range connections strengths between points are weaker than the local within-point connections, a given cortical point operates on



the strongly linear portion of its transfer function for intracortical inputs. Note in **Figure 6** how the output of point-3 is limited to about 45 Spk/s for a maximal input from point-2 of 300 Spk/s. It also responds weakly, but also linearly, to inputs from the more distal point-1. Recalling that the points are fully interconnected, the response of point-3 also depends on its feedback onto points 1 and 2. To summarize, a given cortical point's gain for intracortical inputs is smaller than its gain for external inputs. This makes it operate on the strongly linear portion of the transfer function. In general, a linear system is defined as follows. If one obtains a response  $r_1$  to input  $I_1$  and a response  $r_2$  to input  $I_2$ , then a linear system is strictly defined by the following relation

$$aI_1 + bI_2 = ar_1 + br_2$$

for all pairs of inputs  $I_1$  and  $I_2$  and any combinations of the constants  $a$  and  $b$ . From the aforementioned it should be apparent that our three point model conforms to the definition of linearity.

The reader will recall that in both of the above examples the output is taken from a point which in physiological terms lies in the surround fringe of the activated point (i.e., points that are not directly activated). However, and importantly, it is the firing of all cortical points, whether directly activated or part of the indirectly activated fringe, that is transmitted down the corticospinal tract. Now, as stated in the introduction, since the corticospinal transfer function is essentially linear, it follows that the cortical firing rate vector  $\vec{r}$  is related to the muscle activation vector  $\vec{m}$  by the following simple model

$$\vec{m} = W\vec{r} \tag{14}$$

where the matrix  $W$  contains the weighting coefficients. This model states that each cortical point has fixed connection weights

to one or more motoneuron pools and that each motoneuron pool sums the inputs from one or more cortical points. The simple linear summation of muscle responses occurs precisely because the intracortical interactions are linear, this would not occur otherwise. In the following section we present the analysis that explains why the intracortical interactions are linear.

### Analysis of the Coupled Cortical Points Model

In the analysis of the effective linear response of the single point model we saw that, because the coupling strengths  $q_{ab}$  are large, the network evolves to a state in which the recurrent inhibitory input almost perfectly cancels the external input and the recurrent excitation for both the excitatory and inhibitory populations. In the system of interacting cortical points the same principle applies. The system evolves to a state in which, at each cortical point  $k$ , the recurrent inhibition approximately cancels the external input, the recurrent excitation, as well as the inputs from other points. As in the single point model, this is true for the  $e$ -cell and  $i$ -cell populations.

Requiring that the recurrent inhibition balances the various excitatory inputs to the  $e$ -cells at point  $k$  yields, using Equation (6), that the inhibitory current approximately satisfies

$$q_{ei}r_{i,k} = q_{eo}r_{o,k} + q_{ee}r_{e,k} + \sum_{l \neq k} w_{e,kl}r_{e,l}, \quad (15)$$

while, from the balance of excitation and inhibition for the  $i$ -cells in population  $k$  we obtain approximately

$$q_{ii}r_{i,k} = q_{io}r_{o,k} + q_{ie}r_{e,k} + \sum_{l \neq k} w_{i,kl}r_{e,l}. \quad (16)$$

Eliminating  $r_{i,k}$  from these two equations we obtain

$$\hat{q}_{eo}r_{o,k} + \hat{q}_{ee}r_{e,k} + \sum_{l \neq k} \hat{w}_{e,kl}r_{e,l} = 0, \quad (17)$$

with,

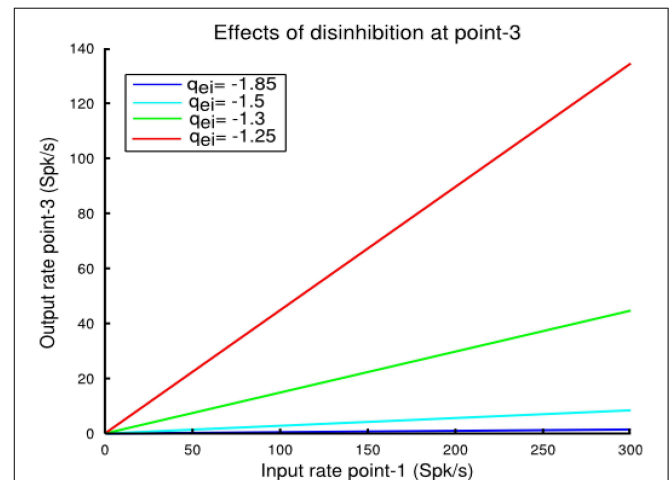
$$\hat{q}_{eo} = q_{eo} - \frac{q_{ei}}{q_{ii}}q_{io}, \quad \hat{q}_{ee} = q_{ee} - \frac{q_{ei}}{q_{ii}}q_{ie}, \quad \text{and} \quad \hat{w}_{e,kl} = w_{e,kl} - \frac{q_{ei}}{q_{ii}}w_{i,kl}.$$

This set of equations can be rewritten in vector notation as  $\hat{q}_{eo} \vec{r}_o + M \vec{r}_e = \vec{0}$ , where  $\vec{r}_o$  and  $\vec{r}_e$  are vectors whose  $k$ -th elements are  $r_{o,k}$  and  $r_{e,k}$ , respectively. The effective connectivity matrix's elements are  $M_{kk} = \hat{q}_{ee}$  and  $M_{kl} = \hat{w}_{e,kl}$ , for  $l \neq k$ . The rates of the excitatory populations are then simply given by  $\vec{r}_e = -\hat{q}_{eo}M^{-1}\vec{r}_o$ . Thus, effectively, the rate of the  $e$ -cell populations depend linearly on the external rates only. As a result, the response of the system to simultaneous activation of two or more points is just the sum of the responses to activation of each point individually.

### Local Control of Synaptic Inhibition and i-Cells

The more important insight gained from the coupled points model pertains to the consequences of local control of inhibitory synaptic strength and inhibitory feedback by the  $MI$ -neurons. When a cortical point is stimulated and a second cortical

point disinhibited by focal iontophoretic release of a  $GABA_A$  antagonist, such as Bicuculline, the response is the same as if both points were stimulated together (Schneider et al., 2002; Ethier et al., 2006). Blocking of  $GABA_A$  receptors is equivalent to reducing  $q_{ei}$  and  $q_{ii}$  in the model. However, we realized that, as the linear approximation Equations (11) and (17) predict, changing the inhibitory synaptic strength parameters in proportion does not change the responsiveness of a cortical point to external or intrinsic inputs. Note that this is only approximately true for the non-linear system as can be appreciated from **Figure 4**. In any case, Equations (11) and (17) predict that, for example, to increase the responsiveness of a cortical point (i.e., its gain),  $q_{ei}$  must be decreased more than  $q_{ii}$ . The consequences of this are shown in **Figure 7**. In that example we demonstrate how local activity at point-1 affects activity at the more distal point-3 as the strength of the local inhibition  $q_{ei}$  at point-3 is reduced. In this condition with the strength of local inhibition at point-3 set to a value of  $q_{ei} = -1.85$ , inputs from point-1 have a minimal effect on the output of point-3. As the strength of  $q_{ei}$  is reduced, the firing rate gain of point-3 is increased. Note that because the inhibitory coefficient  $q_{ii}$  was not changed, the  $q_{ei}/q_{ii}$  ratio decreased. (i.e., differentially controlled). The conclusion is that reduction of local inhibition allows for the functional coupling between points in a graded manner only if the  $q_{ei}/q_{ii}$  ratio is reduced. This has important implications for cortical Physiology in general and, in particular, for the mechanism by which iontophoretic release of  $GABA_A$  receptor antagonists allows cortical points to be coupled. These issues will be dealt with in the discussion.



**FIGURE 7 |** In this example we show how local activity at cortical point-1 affects activity at a the more distal point-3 as inhibition strength at that point is reduced. The coupling strength between the two points was set to  $w_{e,31} = w_{i,31} = 0.02$  and  $w_{e,13} = w_{i,13} = 0.02$ . In this condition, with the strength of the local feedback inhibition set at a value  $q_{ei} = -1.85$ , inputs from point-1 have a minimal effect on the output of point-3 (blue line). As  $q_{ei}$  is reduced at point-3, the firing rate gain of this point is increased. Note, however, that the value of  $q_{ii}$  was not changed. As explained in the text, functional coupling of two points by focal disinhibition only works when  $q_{ei}$  and  $q_{ii}$  are changed differentially, not in proportion.



Reduction of inhibitory synaptic strength as described in the preceding section is non-specific, affecting both auto-inhibition of *i*-cells and feedback inhibition of *e*-cells and thus mimics GABA<sub>A</sub> receptor blockade. To link these experimental observations to physiological function, we have suggested that the feedback inhibition of *e*-cells by *i*-cells may be controlled by *MI*-neurons (Schneider et al., 2002; Capaday, 2004; Capaday et al., 2013). The idea is that these interneurons are specialized at inhibiting the *i*-cell population thereby decreasing their response gain and consequently increasing the gain of the *e*-cells (Figure 5). In our model the *MI*-neuron population controls the firing rate gain  $G_{i-cell}$  of the *i*-cell population to all inputs. This gain control mechanism is based on our 2-C neuron models as explained in the methods. In the example shown in Figure 8 we demonstrate how control of the local *i*-cell population by *MI*-neurons at point-2 modulates its response gain to inputs from point-1. The coupling strength between the two points was set at  $q_{12} = q_{21} = 0.2$  and local inhibitory synaptic strengths,  $q_{ii}$  and  $q_{ei}$ , as in previous examples. Note that when  $G_{i-cell} = 1$  the *MI*-neurons exert no effect on the local *i*-cell population. As  $G_{i-cell}$  is reduced, so too is the effectiveness of the inhibitory feedback on the *e*-cells and consequently the response gain increases. As can be seen in the figure, point-2 responds linearly to inputs from point-1, and with increasing gain as  $G_{i-cell}$  is reduced. At a value nearing  $G_{i-cell} \approx 0.5$ , a deviation from linearity is apparent beyond  $\sim 150$  Spk/s. Thereafter, the response of point-2 follows an accelerating function, as in the example with  $G_{i-cell} = 0.3$ . Thus, gain modulation by the *MI*-neurons maintains response linearity, but only over a limited range of  $0.5 \leq G_{i-cell} \leq 1$ . This effectively means no more than a gain doubling given the parameters of our model. The essential reason for this limitation is the maximum firing rate of the inhibitory population  $r_{max,i}$  limited to 300 Spk/s when  $G_{i-cell} = 1$ .

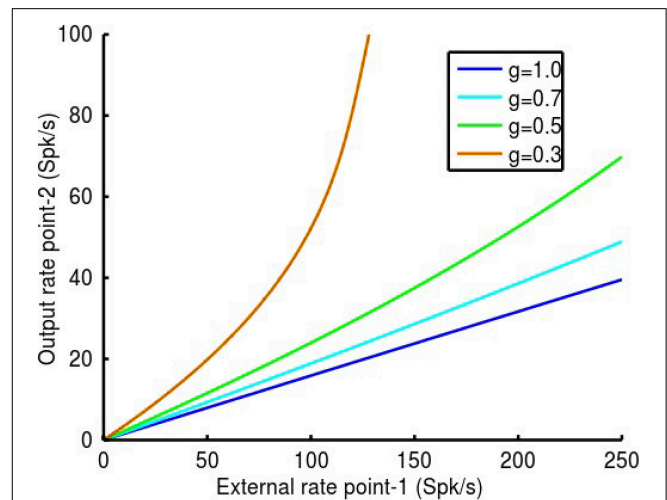
We end this section by presenting the conditions necessary to maintain linearity when *MI*-neurons modulate the response gain of a cortical point. Considering Equation (7) in their linearized version, meaning that  $I_a$  is much less than  $I_{1/2,a}$ , we can write the steady-state firing rate of the *e*-cell population as

$$r_e \approx \frac{q_{eo}(q_{ii} + \alpha_i) - q_{io}q_{ei}}{q_{ei}q_{ie} - (q_{ee} - \alpha_e)(q_{ii} + \alpha_i)} r_o,$$

where  $\alpha_e = I_{1/2,e}/r_{max,e}$  and  $\alpha_i = I_{1/2,i}/G_{i-cell}r_{max,i}$ . One can appreciate that as  $G_{i-cell}$  is reduced,  $\alpha_i$  increases and consequently the gain term in Equation (18) increases. This linear approximation breaks down when the inhibitory rate  $r_{max,i}$  approaches its maximum, as previously explained. The effect of the *MI*-neurons is to decrease the maximum firing rate of the *i*-cell population to  $G_{i-cell}r_{max,i}$ .

### Spiking Model with Conduction Time Delays

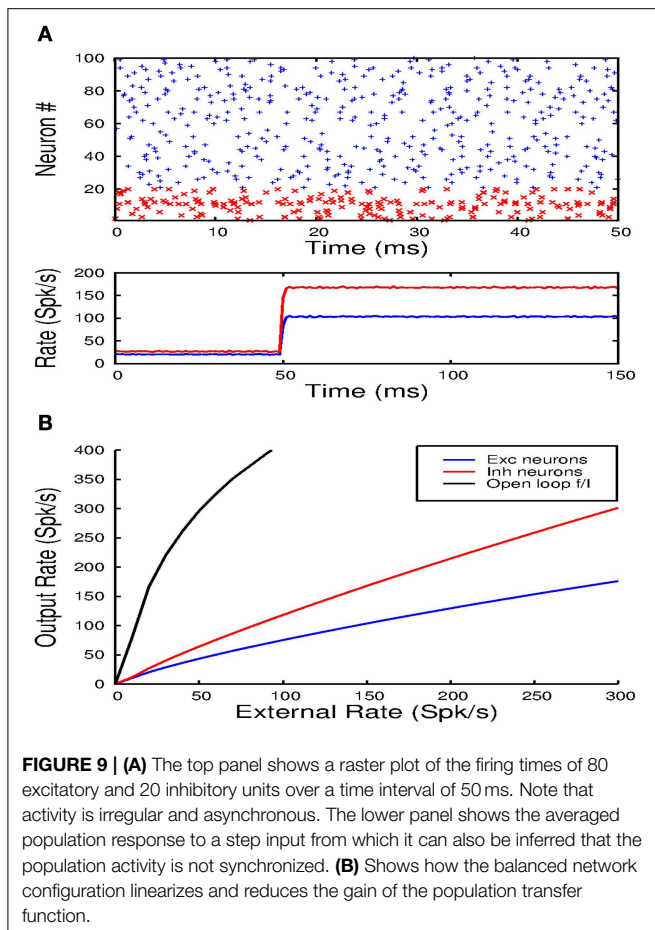
In these simulations of spiking networks, 80% of the synaptic inputs, of any kind, arrive on the dendritic compartment of the *e*-cells and 20% on the soma. Figure 9A shows the response of the network to a step input. The top panel shows a raster plot indicating the spike times of 80 excitatory neurons (blue) and 20 inhibitory cells (red) chosen at random. The lower panel shows



**FIGURE 8 |** In this example we show how *MI*-neurons control of the local *i*-cell population at point-2 modulates its response gain to inputs from point-1. The coupling strength between the two points was set at  $q_{12} = q_{21} = 0.2$  and local inhibitory synaptic strengths,  $q_{ii}$  and  $q_{ei}$ , as in previous examples. Note that when the *MI*-neurons exert no effect on the local *i*-cell s,  $G_{i-cell} = 1$ . As  $G_{i-cell}$  is reduced, so too is the effectiveness of the local inhibitory feedback and consequently the response gain increases. As can be seen in the figure, point-2 responds linearly to inputs from point-1, and with increasing gain as  $G_{i-cell}$  is reduced. At a value nearing  $G_{i-cell} \approx 0.5$ , a deviation from linearity is apparent beyond  $\sim 150$  Spk/s. Thereafter, the response of point-2 becomes an accelerating function, as in the example with  $G_{i-cell} = 0.3$ . Thus, gain modulation by the *MI*-neurons maintains response linearity, but as expected, only over a limited range of  $0.5 \leq G_{i-cell} \leq 1$ . This effectively means not much more than a gain doubling.

the population averaged firing rate determined from the spike counts in 1 ms bins. The raster plot shows that the neurons fire irregularly and asynchronously. This is confirmed by the population averaged rate. After the onset of the step input, the average firing rates of the excitatory and inhibitory cells rapidly increase to a new steady-state value. When this value is reached, the number of spikes in each 1 ms time bin only varies by a small amount, indicating that activity is asynchronous. In Figure 9B the average firing rate in the steady-state is plotted against the strength of the step input for the excitatory (blue) and inhibitory (red) population. The open loop response for the excitatory cells (black) is also shown. Without local inhibitory feedback the excitatory rate first rises rapidly, after which the slope decreases progressively. With local inhibitory feedback the rates of both cell populations vary near linearly with the input rate. Due to the neuronal thresholds, the slope at the initial part of the curves is slightly higher. Nevertheless, as in the rate model, the feedback strongly linearizes the population output.

Next we simulated a network of three interacting cortical points, with a distance of 1 mm and a conduction delay of 10 ms between adjacent points. Figure 10A shows a raster plot of the spiking activity of 80 excitatory and 20 inhibitory neurons in each of the points when point-1 and -3 are activated by step inputs. There is a strong response in points-1 and -3 and a weaker response in point-2. Nevertheless, despite the conduction delays, the activity is clearly asynchronous within and across points.



Additionally, the neurons in the *e*-cell and *i*-cell populations fire irregularly. The three panels in **Figure 10B** show the firing rates of the *e*-cells in the three points as a function of the strength of the input when point-1 is activated (red), when point-3 is activated (green) and when point-1 and -3 are equally activated (blue). In all cases, the response of *e*-cell population at each of the points varies nearly linearly with the input strength. Moreover, at each of the points, the response when both point-1 and -3 are activated is very close to the sum of responses to activation of these points separately.

## Discussion

We have demonstrated a mechanism by which motor cortical points interact and respond to external command inputs linearly, despite non-linear neuron transfer functions. The balanced network configuration puts the motor cortical points in a linear negative feedback regime, as follows. At any given point, auto-inhibition of the *i*-cells effectively linearizes their transfer functions. Consequently, the corresponding *e*-cells receive linear negative feedback, which in turn linearizes their response. The resulting linear intracortical interactions yield linear summation of EMG output vectors, this would not occur otherwise. These conclusions were derived from analysis

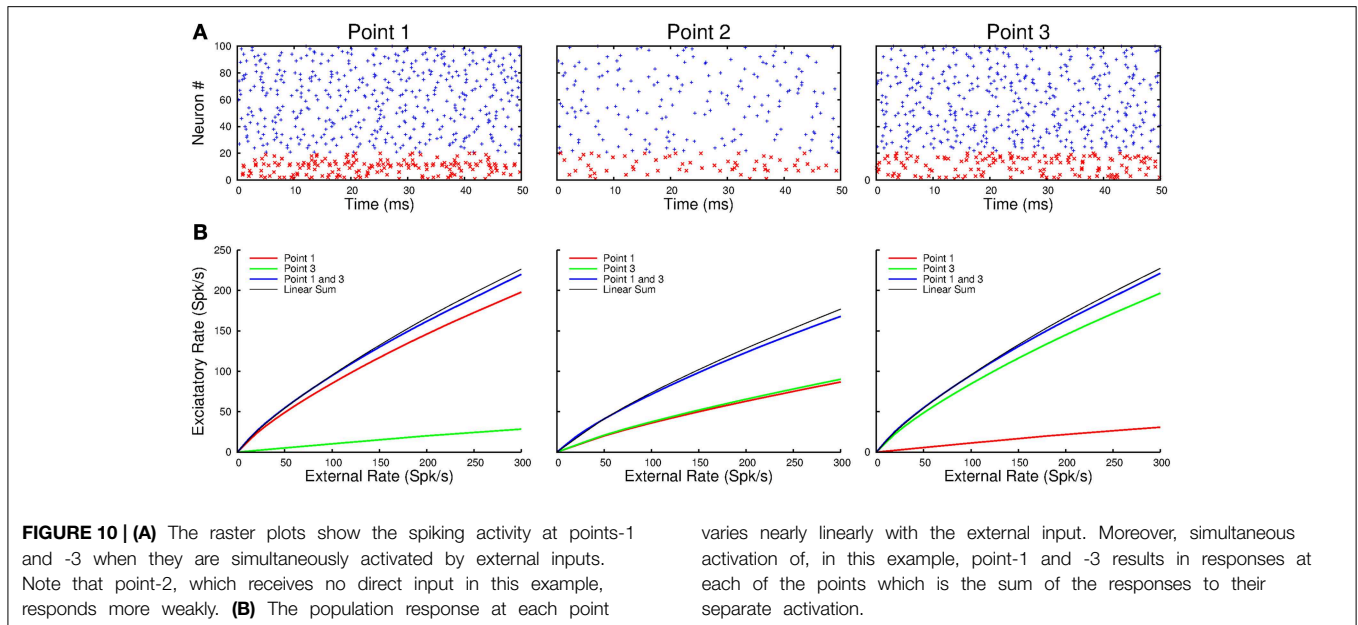
of a rate model. Additionally, numerical simulations of large populations of spiking neurons were done. These were connected according to the same balanced network principles. In this way, we demonstrated that conduction and synaptic time delays between cortical points does not lead to spike synchrony, or oscillations. We have also described how *MI*-neurons can control the response gain of a cortical point and thereby allow spatially separate points to be functionally coupled. However, the effectiveness of this mechanism appears limited to approximately a doubling of the gain, beyond which the response becomes non-linear. In summary, we have shown how the balanced network configuration may contribute to one important operational principle of MCx function, the linear summation of its outputs.

The results of a recent analysis of the behavior of excitatory and inhibitory neurons in various cortical areas, including the MCx, lend further support to our model (Dehghani et al., 2014). The units were separated into putative excitatory or inhibitory neurons based on their extracellular spike shape. The temporal activity of the two populations was shown to be strongly correlated in the way predicted by the balanced network configuration and this over multiple time scales ranging from a few milliseconds to tens of seconds.

## Novel Features of the Model its Limitations and Predictions

We introduced a network consisting of 2-C model neurons to capture an important physiological non-linearity, the relation between synaptic conductance and firing rate (Capaday and van Vreeswijk, 2006). A simple first order Rushton–Naka function describes well this relation derived from these 2-C models. The original formulation of the balanced network configuration by van Vreeswijk and Sompolinsky (1996) was based on binary units, contained no propagation or synaptic delays and the network is effectively a single cortical point. Despite the simplicity of binary units, the mathematical analysis of their interaction turned out to be involved (van Vreeswijk and Sompolinsky, 1998). Here we used a rate model which makes the mechanisms underlying linearization more apparent and the mathematics more approachable. For example, the conditions for stability of the network are more more easily derived and conform to intuition. These stability requirements make well-defined and testable experimental predictions that should further our knowledge of cortical circuitry. What's more, we have extended the balanced network principle to a network of interconnected cortical points whose interaction can be controlled by master inhibitory neurons. Lastly, the original balanced network formulation did not consider the effects of conduction and synaptic delays on network stability. Here we have shown that time delays do not destabilize the asynchronous state—i.e., time delays do not lead to synchronized activity. The reasons are that on the one hand, the delays to the *e*-cell and *i*-cell populations are equal and that on the other hand, the recurrent interactions within each point mitigate spike synchronization. This extends the principles of balanced network operation.

We had previously suggested that the muscle representations in different cortical points could be coupled together by



**FIGURE 10 | (A)** The raster plots show the spiking activity at points-1 and -3 when they are simultaneously activated by external inputs. Note that point-2, which receives no direct input in this example, responds more weakly. **(B)** The population response at each point

varies nearly linearly with the external input. Moreover, simultaneous activation of, in this example, point-1 and -3 results in responses at each of the points which is the sum of the responses to their separate activation.

selective excitation and release from inhibition, so as to produce movement related muscle activation patterns (Schneider et al., 2002). The results presented here show that the mechanism is plausible. Inhibition of the *i*-cells by the *MI*-neurons does increase the response of the *e*-cells allowing the disinhibited cortical point to be functionally coupled linearly with others. However, our analysis reveals that for the interaction to remain linear the gain increase must be modest, two to threefold in the present model. Whether in fact this mechanism is part of the process of motor command generation remains to be determined. A more refined model, that takes into account differential connectivity of *i*-cell population types, may accommodate a greater range of gain control while maintaining linearity. In any case, the range over which gain changes are possible need to be determined experimentally.

Our last point in this section is to explicitly state our model's most important prediction. The population response of any given motor cortical point will be the linear sum of its responses to separate inputs. For example, if a cortical point receives input from two different points, then the output of that point will be the linear sum of its response to each input on its own.

### Explanation of Related Experimental Results

Focal iontophoresis of GABA<sub>A</sub> receptor antagonists should reduce the auto inhibition of *i*-cells to the same extent as the feedback inhibition of the *e*-cells. Our model predicts that this should only have a minimal effect on the firing rate of the *e*-cell population. Consequently, coupling of cortical points by disinhibition should not occur. Yet the experimental evidence is clear, focal disinhibition of a cortical point allows it to be functionally coupled to the activity of a distant cortical point (Schneider et al., 2002). One possibility is that there is indeed a differential effect of GABA<sub>A</sub> receptor antagonists on  $q_{ei}$  and  $q_{ii}$ , as we have modeled it here. Speculating, one may

suggest that the binding affinities of receptor subtypes in the two populations may be different, or that the antagonists may have differential access to the receptors of each population. A second possibility arises from the stability requirement that  $\gamma_i q_{ii} > \gamma_e q_{ee}$ . At some concentration of GABA<sub>A</sub> receptor antagonists,  $q_{ii}$  will become smaller than  $q_{ee}$ . The cortical point will then become unstable and discharge in synchronized periodic bursts. The consequences are different in each case. In the first case, the response at the disinhibited point will scale with the input as we have shown here. In the second case we have a triggered effect, the input triggers the instability and the point responds with a burst whose size is largely independent of the input strength. Only further experiments can unravel which of these mechanisms are involved. Nonetheless, the model predicts that during the phase of increased excitability when the neural activity is still asynchronous the disinhibited point will respond proportionately to inputs. By contrast, when the point is spontaneously bursting, the inputs will only reset the burst phase with relatively little effect on burst amplitude. Both possibilities can explain why the outputs of a microstimulated point and a second point made spontaneously active by focal disinhibition sum linearly (Ethier et al., 2006).

A third experimental observation explained by our model is why spike activity at one point does not spread over the full length of the extant connections (Capaday et al., 2011). Yet, upon disinhibition a more distal point can be recruited (Schneider et al., 2002). A given point has a small gain to distal inputs, consequently its response may be below threshold. When this point is disinhibited, its response gain increases sufficiently to respond to distal inputs with substantial spike activity leading to a motor output. It should be noted that this explanation is inconsistent with the long standing concept of surround inhibition, which would lead to sublinear summation.

## Functional Advantage of Linearity for Motor Command Generation

What functional advantage does linear summation at the cortical level afford to the synthesis of motor commands? We propose that linear summation of the muscle synergies represented in different cortical points provides a simple prediction of the consequences of their simultaneous activation. This may simplify the generation of motor commands. Furthermore, assuming that the movement modules represented in the MCx are finite in number, interpolation is necessary to have a continuous space of possible movements. If the interactions within the MCx were non-linear, the interpolation process would be more challenging. We make no claim that the musculo-skeletal system is linear or controlled linearly, only that the motor cortical commands themselves are derived by linear combination. Results from a wide variety of approaches and experimental preparations show that a given class of movements, including locomotor movements and postural adjustments, are controlled by a small set of muscle synergies (reviewed by Bizzi and Cheung, 2013). Common to all the various theoretical formulations (e.g., D'Avella et al., 2006) is the linear combination of basic output modules. Neurophysiological evidence for the existence of neural circuits underlying these modules have been reported for pre-motoneuronal spinal networks and the

MCx (Bizzi et al., 1991; Giszter et al., 1993; Schneider et al., 2001; Ethier et al., 2006; Graziano, 2006). What is more, the spinal modules and cortical points combine their respective outputs linearly (Mussa-Ivaldi et al., 1994; Ethier et al., 2006). More recently, it was suggested that the MCx selects and combines spinal modules temporally to produce the muscle activation patterns of reaching movements (Overduin et al., 2012).

## Epilogue

It may be suggested that linearization can be effected by two opposing non-linearities, one intracortical and one intraspinal. By considering established physiological and anatomical principles, we have provided evidence for an alternative explanation. Two consecutive linear stages account for the phenomenon.

## Acknowledgments

CC was a “Professeur invité” of the Université Paris-Descartes. We thank Dr. Claude Meunier and Prof. W.G. Darling for their comments and suggestions on a draft of the manuscript. This work was partly funded by the Agence Nationale de la Recherche, ANR 09-SYS-C002.

## References

- Abbott, L. F., and Regehr, W. G. (2004). Synaptic computation. *Nature* 431, 796–803. doi: 10.1038/nature03010
- Acsády, L., Görcs, T. J., and Freund, T. F. (1996). Different populations of vasoactive intestinal polypeptide-immunoreactive interneurons are specialized to control pyramidal cells or interneurons in the hippocampus. *Neuroscience* 73, 317–334. doi: 10.1016/0306-4522(95)00609-5
- Bacci, A., Huguenard, J. R., and Prince, D. A. (2005). Modulation of neocortical interneurons: extrinsic influences and exercises in self-control. *Trends Neurosci.* 28, 602–610. doi: 10.1016/j.tins.2005.08.007
- Bizzi, E., and Cheung, V. C. K. (2013). The neural origin of muscle synergies. *Front. Comput. Neurosci.* 7:51. doi: 10.3389/fncom.2013.00051
- Bizzi, E., Mussa-Ivaldi, F. A., and Giszter, S. (1991). Computations underlying the execution of movement: a biological perspective. *Science* 253, 287–291. doi: 10.1126/science.1857964
- Capaday, C. (2004). The integrated nature of motor cortical function. *Neuroscientist* 10, 207–220. doi: 10.1177/107385403262109
- Capaday, C., Ethier, C., Brizzi, L., Sik, A., van Vreeswijk, C., and Gingras, D. (2009). On the nature of the intrinsic connectivity of the cat motor cortex: evidence for a recurrent neural network topology. *J. Neurophysiol.* 102, 2131–2141. doi: 10.1152/jn.91319.2008
- Capaday, C., Ethier, C., van Vreeswijk, C., and Darling, W. G. (2013). On the functional organization and operational principles of the motor cortex. *Front. Neural Circuits* 7:66. doi: 10.3389/fncir.2013.00066
- Capaday, C., and van Vreeswijk, C. (2006). Direct control of firing rate gain by dendritic shunting inhibition. *J. Integr. Neurosci.* 5, 199–222. doi: 10.1142/S021963520600115X
- Capaday, C., van Vreeswijk, C., Ethier, C., Ferkinghoff-Borg, J., and Weber, D. (2011). Neural mechanism of activity spread in the cat motor cortex and its relation to the intrinsic connectivity. *J. Physiol.* 589, 2515–2528. doi: 10.1113/jphysiol.2011.206938
- Crill, W. E., and Schwandt, P. C. (1983). Active currents in mammalian central neurons. *Trends Neurosci.* 6, 236–240. doi: 10.1016/0166-2236(83)90101-7
- D'Avella, A., Portone, A., Fernandez, L., and Lacquaniti, F. (2006). Control of fast-reaching movements by muscle synergy combinations. *J. Neurosci.* 26, 7791–7810. doi: 10.1523/JNEUROSCI.0830-06.2006
- Dehghani, N., Peyrache, A., Telenczuk, B., Le Van Quyen, M., Halgren, E., Cash, S. S., et al. (2014). Multiscale Balance of Excitation and Inhibition in Single-Unit ensemble Recordings in Human and Monkey Neocortex. arXiv:1410.2610.
- Ethier, C., Brizzi, L., Darling, W. G., and Capaday, C. (2006). Linear summation of cat motor cortex outputs. *J. Neurosci.* 26, 5574–5581. doi: 10.1523/JNEUROSCI.5332-05.2006
- Fino, E., Packer, A. M., and Yuste, R. (2013). The logic of inhibitory connectivity in the neocortex. *Neuroscientist* 19, 228–237. doi: 10.1177/107385412456743
- Freund, T. F., and Antal, M. (1988). GABA-containing neurons in the septum control inhibitory interneurons in the hippocampus. *Nature* 336, 170–173. doi: 10.1038/336170a0
- Giszter, S. F., Mussa-Ivaldi, F. A., and Bizzi, E. (1993). Convergent force fields organized in the frog's spinal cord. *J. Neurosci.* 13, 467–491.
- Graziano, M. (2006). The organization of behavioral repertoire in motor cortex. *Annu. Rev. Neurosci.* 29, 105–134. doi: 10.1146/annurev.neuro.29.051605.112924
- Haider, B., Duque, A., Hasenstaub, A. R., and McCormick, D. A. (2006). Neocortical network activity *in vivo* is generated through a dynamic balance of excitation and inhibition. *J. Neurosci. Off. J. Soc. Neurosci.* 26, 4535–4545. doi: 10.1523/JNEUROSCI.5297-05.2006
- Markram, H. (2010). “Microcircuitry of the neocortex,” In *Handbook of Brain Microcircuits*, eds G. Shepherd and S. Grillner (New York, NY: Oxford University Press), 22–30. doi: 10.1093/med/9780195389883.003.0003
- Mussa-Ivaldi, F. A., Giszter, S. F., and Bizzi, E. (1994). Linear combinations of primitives in vertebrate motor control. *Proc. Natl. Acad. Sci. U.S.A.* 91, 7534–7538. doi: 10.1073/pnas.91.16.7534
- Okun, M., and Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nat. Neurosci.* 11, 535–537. doi: 10.1038/nn.2105
- Overduin, S. A., D'Avella, A., Carmena, J. M., and Bizzi, E. (2012). Microstimulation activates a handful of muscle synergies. *Neuron* 76, 1071–1077. doi: 10.1016/j.neuron.2012.10.018

- Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., and Scanzani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat. Neurosci.* 16, 1068–1076. doi: 10.1038/nn.3446
- Roxin, A., Brunel, N., and Hansel, D. (2005). Role of delays in shaping spatiotemporal dynamics of neuronal activity in large networks. *Phys. Rev. Lett.* 94:238103. doi: 10.1103/PhysRevLett.94.238103
- Schneider, C., Devanne, H., Lavoie, B. A., and Capaday, C. (2002). Neural mechanisms involved in the functional linking of motor cortical points. *Exp. Brain Res.* 146, 86–94. doi: 10.1007/s00221-002-1137-2
- Schneider, C., Zytnicki, D., and Capaday, C. (2001). Quantitative evidence for multiple widespread representations of individual muscles in the cat motor cortex. *Neurosci. Lett.* 310, 183–187. doi: 10.1016/S0304-3940(01)02105-X
- Shriki, O., Hansel, D., and Sompolinsky, H. (2003). Rate models for conductance-based cortical neuronal networks. *Neural Comput.* 15, 1809–1841. doi: 10.1162/08997660360675053
- Sompolinsky, S. H., and White, W. O. (2004). “Theory of large recurrent networks: from spikes to behavior,” in *Methods and Models in Neurophysics: Lecture Notes of the Les Houches Summer School 2003*, eds C. C. C. Chow, B. G. B. Gutkin, D. H. D. Hansel, C. Meunier and J. Dalibard (Elsevier), 267–340.
- Stafstrom, C. E., Schwindt, P. C., and Crill, W. E. (1984a). Cable properties of layer V neurons from cat sensorimotor cortex *in vitro*. *J. Neurophysiol.* 52, 278–289.
- Stafstrom, C. E., Schwindt, P. C., and Crill, W. E. (1984b). Repetitive firing in layer V neurons from cat neocortex *in vitro*. *J. Neurophysiol.* 52, 264–277.
- Tomioka, R., Okamoto, K., Furuta, T., Fujiyama, F., Iwasato, T., Yanagawa, Y., et al. (2005). Demonstration of long-range GABAergic connections distributed throughout the mouse neocortex. *Eur. J. Neurosci.* 21, 1587–1600. doi: 10.1111/j.1460-9568.2005.03989.x
- Townsend, B. R., Paninski, L., and Lemon, R. N. (2006). Linear encoding of muscle activity in primary motor cortex and cerebellum. *J. Neurophysiol.* 96, 578–2592. doi: 10.1152/jn.01086.2005
- van Vreeswijk, C., and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274, 1724–1726.
- van Vreeswijk, C., and Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Comput.* 10, 1321–1371. doi: 10.1162/089976698300017214
- Vogels, T. P., Rajan, K., and Abbott, L. F. (2005). Neural network dynamics. *Annu. Rev. Neurosci.* 28, 357–376. doi: 10.1146/annurev.neuro.28.061604.135637
- Wilson, H. R. (1999). *Spikes, Decisions, and Actions*. New York, NY: Oxford University Press.

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

Copyright © 2015 Capaday and van Vreeswijk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.