Increased spinal excitation causes decreased locomotor complexity

Myriam Lauren de Graaf[1,2,3*], Luis Mochizuki[4], Heiko Wagner[1,2], and Charlotte Le Mouel[1,5]

[1] *Dept. of Movement Science, University of Münster, Horstmarer Landweg 62b, 48149 Münster, Germany*
[2] *Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Münster, Germany*
[3] *Center for Nonlinear Science (CeNoS), University of Münster, Corrensstraße 2, 48149, Münster, Germany*
[4] *School of Arts, Sciences and Humanities, University of São Paulo, Brazil*
[5] *Institut des Systèmes Intelligents et de Robotique (ISIR), Sorbonne Université, CNRS, Paris, F-75005, France*
[*] *Corresponding author: mdegraaf@uni-muenster.de*

Abstract

Animals display rich and coordinated motor patterns during walking and running, that are generated and controlled within the central nervous system. Previous computational and experimental results suggest that the balance between excitation and inhibition in neural circuits may be critical for generating such structured motor patterns. In this paper, we explore the influence of this balance on the ability of a reservoir computing artificial neural network to learn human locomotor patterns, using mean-field theory analysis and simulations. We varied the numbers of neurons, connection percentages and connection strengths of excitatory and inhibitory neuron populations, and introduced the anatomical *imbalance* that quantifies their combined overall effect. Our results indicate that network dynamics and performance depend critically on the anatomical imbalance in the network. Inhibition-dominated networks work well, displaying balanced input to the neurons and good firing rate variability. Excitation-dominated networks, however, lead to saturated firing rates, thereby reducing the effective dimensionality of the networks and leading to simplified motor output patterns. This suggests that motor pattern generation may be robust to increased inhibition but not increased excitation in neural networks.

***Keywords*—** EI balance, artificial neural networks, excitation, inhibition, locomotion, motor control

1 Introduction

During walking and running, humans and other animals create rich and coordinated motor patterns. With increasing locomotor speed, the stepping frequency increases and the pattern of muscle contraction changes, in terms of both the amplitude of muscle contractions and their timing within the step cycle (Ivanenko et al., 2006; Ivanenko et al., 2004; Lacquaniti et al., 2012). The muscle activation patterns are generated within the spinal cord by locomotor central pattern generators (CPGs) (Delcomyn, 1980; Grillner & Wallén, 1985; Grillner et al., 1998; Ijspeert, 2008). The spinal cord is capable of not only generating the rhythm of locomotion, but also transforming this simple rhythm into the complex muscle contraction patterns observed during locomotion.

Separate spinal circuits have been identified that deal with these processes of rhythm generation on the one hand and pattern formation on the other. Rhythm-generating neurons are located more medially and project locally, whereas pattern-forming neurons are located more laterally and project to the lateral edge of the spinal cord, where the motoneuron pools are situated (Griener et al., 2013). Assuming separate rhythm generation and pattern formation circuits in modeling approaches can successfully replicate experimental phenomena such as non-resetting deletions (Lafreniere-Roula & McCrea, 2005), i.e. the finding that rhythm can be maintained independent of the pattern formation (McCrea & Rybak, 2007, 2008; Rybak et al., 2006), as well as the effects of sensory stimulation on locomotor patterns (McCrea & Rybak, 2007; Rybak et al., 2006).

Rhythm generation by CPGs is most often conceptually modelled using the half-center model (explained in more detail in Latash, 2012). This model assumes reciprocal inhibition between two neural pools (the half-centers), one of which excites extensor motoneurons and the other flexor motoneurons (Jankowska et al., 1967). When the extensor pool is active, it inhibits the flexor pool. After a while, activity in the extensor pool decreases by some self-inhibitory process, such as fatigue (Brown, 1911; Latash, 2012; McCrea & Rybak, 2008). This releases the flexor pool from inhibition, allowing it to activate and concurrently inhibit the extensor pool until it, in turn, fatigues. The result is a rhythm of alternating activity between the flexor and extensor pools.

Pattern formation has received comparatively less attention than rhythm generation. Some models propose coupled oscillators at each joint (unit burst generators, Grillner, 1981), or two half-centers acting at each joint (Li et al., 2017), which allows hip, knee and ankle extensors to vary slightly out of phase. However, these models still assume a strict alternation between flexors and extensors, thus failing to capture co-activation and the more complex activations of bi-articular muscles (Markin et al., 2012). Later models therefore, extended the pattern formation networks to also account for bi-articular muscles (Shevtsova et al., 2016).

None of these conceptual models, however, are able to explain the concurrent inhibitory and excitatory neural activity observed in nature. According to the half-center models, the input to a neuron in a given pool should be excitatory when that neuron's pool is active, and inhibitory when the opposite pool is active. However, intracellular recordings of spinal neurons show that motorneurons receive simultaneous excitatory and inhibitory inputs when their pool is active, rather than alternating excitation and inhibition (Berg et al., 2007; Petersen et al., 2014). A hallmark of this balance between excitatory and inhibitory inputs (EI balance) is a firing rate distribution across the neural population which is positively skewed (Petersen & Berg, 2016). These skewed firing rates cannot be recreated by reciprocal inhibition, as reciprocal inhibition leads to disinhibition of the active half-center and thus to run-away activity (Berg et al., 2019). Recurrent inhibition must thus be present in spinal neural circuits. Since the half-center models do not incorporate recurrent inhibition, they are unable to investigate its effects.

The role of inhibition in CPGs has been studied experimentally by blocking inhibition in animal models. When inhibition is blocked, the generated rhythm is undisturbed, but the flexors and extensors are active simultaneously and vary in phase (A. H. Cohen & Harris-Warrick, 1984; Cowley & Schmidt, 1995; Lanuza et al., 2004; Talpalar et al., 2013). This indicates that inhibition is important for the alternating pattern between flexion and extension (Britz et al., 2015; J. Zhang et al., 2014). Similarly, inhibitory neurons have been shown to play an important role in alternation between the left and right sides of the body (Jankowska, 2008; Kiehn, 2006; Lanuza et al., 2004; Quinlan & Kiehn, 2007; Talpalar et al., 2013). Inhibition thus plays a critical role in pattern formation of CPGs.

To further investigate the role of recurrent inhibition in pattern formation, we propose to use recurrent neural networks. Simulation studies have shown that such networks are capable of producing complex, high-dimensional signals, even in the absence of input, or with a very simple input signal (Funahashi & Nakamura, 1993). In reservoir computing (Verstraeten et al., 2007), a subtype of recurrent neural networks, linear readouts of the network's activity are trained to reproduce a variety of target signals, while the internal connection weights remain static. Reservoir computing has been used to model how biological neural networks achieve a variety of tasks (Boström et al., 2013; Hinaut & Dominey, 2013; Wyffels & Schrauwen, 2009). Moreover, it has proven highly effective in reproducing features of human locomotion such as the trajectories of markers placed on the human body (Sussillo & Abbott, 2009) and joint angles (Hoellinger et al., 2013; Wyffels & Schrauwen, 2009).

An additional advantage of using artificial neural network models is that their behaviour can be analyzed via mean-field theory (MFT), which describes the behaviour of large, complex systems in terms of their average effects, rather than each of their individual components. MFT has been applied to investigate neural network dynamics, e.g., around the transition to chaos (Kadmon & Sompolinsky, 2015), or in networks with balanced excitation and inhibition (Harish & Hansel, 2015). Recurrent neural networks thus provide a versatile and insightful framework for studying neural pattern formation (Hoellinger et al., 2013; Wyffels & Schrauwen, 2009).

By modelling a network with recurrently connected neurons that mimic the excitatory and inhibitory interactions observed in spinal networks, we can investigate the importance of recurrent inhibition in the generation and modulation of motor patterns. While the functional EI balance in both biological and artificial neural networks has been extensively studied (reviewed in, e.g., Hennequin et al., 2017; Herstel & Wierenga, 2021; Isaacson & Scanziani, 2011; Liang et al., 2024; Sadeh & Clopath, 2021), to our knowledge,

none of these studies systematically explored the underlying anatomical ratios between the excitatory and inhibitory populations in recurrent neural networks.

When individual neurons are modelled as rate units, they typically do not obey Dale's Law (Eccles, 1976): their outgoing connection weights are most often drawn from a random distribution centred around zero (Boström et al., 2013; Jaeger, 2001; Sussillo & Abbott, 2009), causing them to exert both excitatory and inhibitory influences. In spiking neural networks, the neurons do obey Dale's law, but they are often incorporated in a 4:1 ratio, reflecting the respective population sizes found in the cortex (Marín, 2012; Marom & Shahaf, 2002; Meinecke & Peters, 1987; Sahara et al., 2012; Wonders & Anderson, 2006). However, none of these studies have focused on motor pattern generation, where spinal cord-specific EI ratios might be more relevant. Findings in rats indicate that the anatomical EI ratio might be smaller in the spinal cord (Todd & Sullivan, 1990). This indicates that the 4:1 ratio is not omnipresent, and that inhibition might play a larger role in spinal motor pattern generation.

The goal of our paper is to study the role of recurrent inhibition and anatomical EI balance in motor pattern formation. We adapt an existing reservoir computing model (Sussillo & Abbott, 2009) to have distinct excitatory and inhibitory neural populations, and introduce a new composite measure (*imbalance*) that quantifies their relative influence on the network dynamics in terms of their size and connectivity. We systematically alter the imbalance to determine its influence on the network dynamics and muscle activation patterns generated by the network.

2 Methods

We modelled pattern formation in the spinal cord by training simulated neural networks to produce locomotor muscle activations obtained from human experiments. We investigated the influence of the ratio of excitation to inhibition in these networks, by varying the numbers of neurons (N), connection percentages (p), and connection strengths (g) for the two populations. We relate the network performance to the intrinsic network dynamics, obtained through Mean Field Theory (Mastrogiuseppe & Ostojic, 2017, 2018; Rajan et al., 2010; Sompolinsky et al., 1988).

Simulations and data analysis were performed in Matlab (Version R2023a, The MathWorks, Inc., Natick, Massachusetts, United States).

2.1 Network

The investigated networks consisted of an excitatory population E with neuron number N_E (red triangles in the reservoir, Figure 1, lower panel) and an inhibitory population I with neuron number N_I (blue circles in the reservoir, Figure 1, lower panel). Any given neuron i in either population receives an external input, output feedback and recurrent input. The external input equals $J^{In}U(t)$, where $U(t) = 1 + \sin(\omega t)$ with ω equal to the target stride frequency (Figure 1, bottom left), and J^{In} are the input weights drawn from a Gaussian distribution with mean 0 and variance $v_{in} = \frac{1}{3}$. The output feedback equals $J_i^{Fb}Y(t)$, where $Y(t)$ is the 17-dimensional output of the network (Figure 1, bottom right) and the feedback weights J^{Fb} are drawn from a uniform distribution between -1 and 1. The connection weights J_{ij}^E from any excitatory neuron j to any neuron i have a probability p_E of being non-zero. For non-zero weights, the weight equals the absolute value of a variable drawn from a Gaussian distribution with mean zero and variance V_E^0. The variance V_E^0 is chosen such that the variance v_E of the excitatory weights is independent of p_E and scales with N_E according to $v_E = g_E^2/N_E$ (see Section A.5), where g_E scales the excitatory connection strength (Sompolinsky et al., 1988; Sussillo & Abbott, 2009). Likewise, inhibitory weights J_{ij}^I have a probability p_I of being non-zero. If non-zero, their value is the negative of the absolute value of a Gaussian random variable with mean zero and variance V_I^0, chosen such that the variance of the inhibitory weights is $v_I = g_I^2/N_I$.

The internal state x_i of each neuron i is initialized randomly, drawn from the standard normal distribution, and follows the dynamics:

$$\tau \frac{d}{dt} x_i(t) = -x_i(t) + \sum_{j=1}^{N_E} J_{ij}^E \phi(x_j^E(t)) + \sum_{j=1}^{N_I} J_{ij}^I \phi(x_j^I(t)) + J_i^{Fb} Y(t) + J_i^{In} U(t) \tag{1}$$

with time constant $\tau = 0.01$. The neurons' firing rate equals $\phi(x_i)$, where ϕ is $\tanh(\cdot)$ if x_i is positive, and

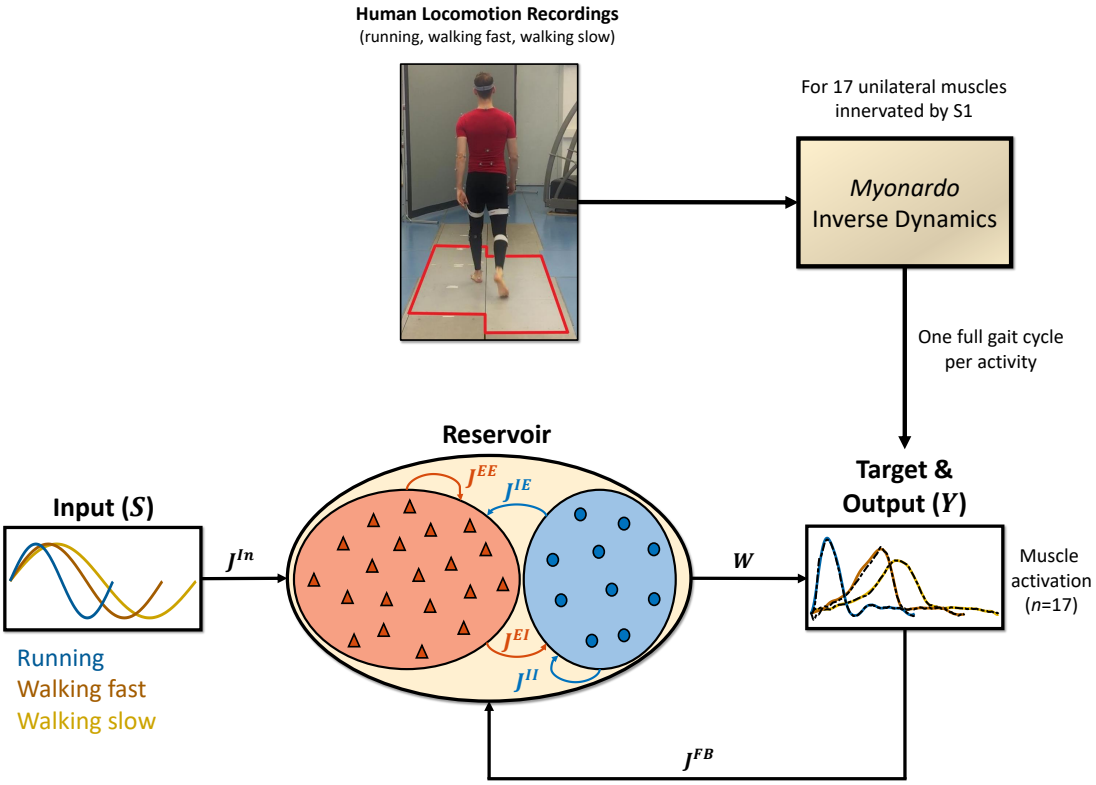

Figure 1 – Schematic overview of methods. *Human locomotor signals were recorded from one adult during slow walking, faster walking, and running (top middle). Ground reaction forces were recorded using six Kistler Force Platforms (outlined in red) and kinematics were recorded using a Qualisys measuring system. The recorded kinetics and kinematics were the input to the inverse dynamics calculations (top right), which provided the required muscle activation for 17 unilateral muscles. These muscle activations were used as the (17-dimensional) target output of the neural network (bottom right). The neural network (bottom middle) consists of two connected excitatory (red, triangles) and inhibitory (blue, circles) populations. In the network, blue lines denote inhibitory connections, while red lines denote excitatory connections. The input to the network (bottom left) was a sinusoid with a frequency corresponding to the gait frequency for each activity: running (blue), fast walking (brown), and slow walking (yellow).*

zero otherwise. The output of the network is calculated from the neuron firing rates as follows:

$$Y_i(t) = \sum_{j=1}^{k} W_{ij} \cdot \phi\left(x_j(t)\right) \tag{2}$$

where W_{ij} is the output weights matrix, that is initialized as zeroes and evolves throughout training (see Section 2.3.3 and Appendix C).

Our network is an adaptation of an echo state neural network (Sussillo & Abbott, 2009) with two major modifications. First, synaptic weights from any given neuron are either all positive or all negative, whereas, in the original network, all recurrent weights are drawn from a Gaussian distribution of mean zero and variance $\frac{g^2}{pN}$, therefore a given neuron can have both positive and negative output weights (Sussillo & Abbott, 2009). Second, the firing rates of all neurons were constrained to be positive, whereas in the original network, the firing rate is $\tanh(x_i)$ which can be either positive or negative depending on the sign of x_i. As a result, individual neurons in our network can be considered either excitatory (red triangles in Figure 1, lower panel) or inhibitory (blue circles in Figure 1, lower panel).

2.2 Mean Field Theory analysis

We performed a theoretical analysis of the networks using Mean Field Theory (MFT); a technique where a description of the network dynamics can be derived self-consistently by averaging over the random parts of the network connectivity (Mastrogiuseppe & Ostojic, 2017, 2018; Rajan et al., 2010; Sompolinsky et al.,

1988). For simplicity, we omitted the output feedback term in this study, but see Mastrogiuseppe and Ostojic (2018) for a full treatment of the output feedback. A detailed description of our MFT approach can be found in Appendix A, but we will present a short overview here.

Briefly, in the MFT approach, the input to each neuron is modelled as a Gaussian random process. By averaging across different realizations of the random connectivity (denoted by [.]), we can express the mean μ and the variance Δ of the input as a function of the mean firing rate $[\phi]$ and mean squared firing rate $[\phi^2]$ (see Appendix Section A.2) as follows:

$$\mu = (m_E N_E + m_I N_I)\left[\phi\right] \tag{3}$$

$$\Delta = (v_E N_E + v_I N_I)[\phi^2] + v_{in} U(t)^2 \tag{4}$$

where m_E and m_I are the mean values of the excitatory and inhibitory weights; and v_E and v_I are their variances. The population activity thus depends critically on two parameters: the mean $m_E N_E + m_I N_I$ and the variance $v_E N_E + v_I N_I$ of the sum of the recurrent weights to each neuron. We define the ratio between the mean recurrent input μ and the average firing rate as the anatomical imbalance A:

$$A = m_E N_E + m_I N_I \tag{5}$$

Next, we assume that the network is in a quasi-stationary state (i.e. the input signal $U(t) = 1 + \sin(\omega t)$ varies slowly compared to the time constant τ). In this state, the activation x_i of each neuron has the same mean and variance as its input. We explicitly model this activation as a Gaussian variable of mean μ and variance Δ. In this way, we can express $[\phi]$ and $[\phi^2]$ as a function of the input mean μ and variance Δ (see Appendix Section A.3):

$$\left[\phi\right] = \int Dz \cdot \phi\left(\mu + \sqrt{\Delta} z\right) \tag{6}$$

$$\left[\phi^2\right] = \int Dz \cdot \phi^2\left(\mu + \sqrt{\Delta} z\right) \tag{7}$$

where $\int Dz = \int_{-\infty}^{+\infty} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz$ is derived from the probability distribution of the Gaussian process z. We solve this set of four equations with four unknowns to determine self-consistently the stationary solution given by $\mu_0, \Delta_0, [\phi]_0, [\phi^2]_0$.

Finally, we perform a stability analysis and show that the stationary solution loses stability for large values of $v_E N_E + v_I N_I$ (see Appendix Section A.6). We determine the maximal value of $v_E N_E + v_I N_I$ for which networks are stable for all values of imbalance A and restrict ourselves to this stable range (see Figure S.4A). We find that, with this restriction, network performance does not depend on $v_E N_E + v_I N_I$ (see Figure S.7A). Note that $v_E N_E + v_I N_I$ is varied via the overall strength $g_{tot} = \sqrt{v_E N_E + v_I N_I}$, as explained in Section 2.3).

2.3 Simulations

The differential equation governing the change in neural activation (Equation 1) was solved numerically using Euler-forward integration to get the new neuron activations for each time step ($\Delta t = 0.005s$).

2.3.1 Network parameters

For the specific weight distributions in our networks (see Section A.5), the anatomical imbalance can be expressed as a function of the neuron number (N), the connection percentage (p), and the connection strengths (g) of the two populations according to:

$$A = m_E N_E + m_I N_I = \sqrt{\frac{2}{\pi}}\left(g_E \sqrt{\frac{p_E}{N_E\left(1 - \frac{2}{\pi} p_E\right)}} - g_I \sqrt{\frac{p_I}{N_I\left(1 - \frac{2}{\pi} p_I\right)}}\right) \tag{8}$$

We investigated networks with imbalances ranging from −15 to +15 in steps of 1, by modifying either N, p, or g for both populations. This analysis was carried out for various levels of the total network size $N_{tot} = N_E + N_I$, the mean connection percentage $p_{tot} = \frac{N_E p_E + N_I p_I}{N_E + N_I}$, or the total connection strength $g_{tot} = \sqrt{v_E N_E + v_I N_I} = \sqrt{g_E^2 + g_I^2}$.

First, the imbalance was modified via N, with total network sizes of N_{tot} = 300, 400, 500, 750, 1000, 1500 and 2000. We chose a minimum network size of 300 as networks with fewer neurons were found to have poor performance (see Figure S.7). The connection percentage was set constant at $p_E = p_I = 0.1$, corresponding to connection ratios found in *ex vivo* neural networks (Marom & Shahaf, 2002), and the connection strength at $g_E = g_I = 1.5$, as this has been shown to produce good network performance (Sussillo & Abbott, 2009).

Second, the imbalance was modified via p, with the mean connection percentage p_{tot} ranging from 0.05 to 0.5 in steps of 0.05. The total neuron number was set constant at $N_{tot} = 750$, with neuron ratios varying from 4:1 to 1:4, namely: N_E:N_I = 600:150, 500:250, 375:375, 250:500 and 150:600. The connection strength was set constant at $g_E = g_I = 1.5$.

Finally, imbalance was modified via g, with the total connection strength g_{tot} ranging from 0.5 to 2.25 in steps of 0.25. Additionally, $g_{tot} \approx 2.12$ was included, corresponding to $g_E = g_I = 1.5$, i.e. the connection strengths used when varying N and p. Networks with $g_{tot} \geq 2.5$ were excluded from further analysis as these were predicted to be unstable by the mean-field theory analysis (see Figure S.4). The total network size was set constant at $N_{tot} = 750$, with neuron ratios varying from 4:1 to 1:4, as above. The connection percentage was set constant at $p_E = p_I = 0.1$.

For each combination of parameter settings, twenty networks were trained and tested. The results were averaged over these twenty networks to decrease the influence of random initialization.

2.3.2 Locomotor muscle activations

The neural networks presented in the previous sections were trained to produce locomotor muscle activations obtained from human experiments.

Kinematics and ground reaction forces were recorded for three locomotor activities: walking at two different (self-chosen) speeds and running. Kinetics were recorded using six Kistler force platforms (Type 9287CA, 90×60 cm, Kistler Instrumente AG, Winterthur, Switzerland - outlined in red in the photo in Figure 1). Whole-body kinematics were recorded at 200 Hz (Qualisys, Göteborg, Sweden) using a modified Plug-In Gait marker set (Vicon Motion Systems Ltd, Oxford, UK). A single participant (male, 26 years old, 1.96m tall, and weighing 86kg) walked and ran across the length of the force plates. For each of the three activities, a single stride from a right heel strike to the following right heel strike was selected. The heel strikes were detected by finding the lowest point of the right heel marker, as recorded by the Qualisys system. This data was obtained as part of another experiment that has been approved by the local ethics committee of the Faculty of Sports Science and Psychology of the University of Münster (#2019-10-RD). The participant signed the informed consent form before the start of the measurements.

Muscle activations, ranging from 0 to a theoretical maximum of 1, were calculated from the recorded kinetics and kinematics using the inverse dynamics function of the musculoskeletal model *Myonardo*® (Predimo GmbH. Myonardo [Software]. Münster, Germany). For more information on the musculoskeletal model and the performed calculations, see Appendix B and (Wagner et al., 2022). From the model output, we selected the 17 muscles on the right side of the body that are innervated by the S1 spinal segment (Sharrard, 1964): m. gluteus maximus, m. gluteus medius, m. gluteus minimus, m. biceps femoris (caput longum and breve), m. semitendinosus, m. semimembranosus, m. tensor fasciae latae, m. piriformis, m. gastrocenemius, m. soleus, m. peroneus longus and brevis, m. flexor hallucis longus, m. flexor digitorum longus, m. extensor hallucis longus, and m. extensor digitorum longus. This resulted in a 17-dimensional target signal for each of the 3 locomotor activities.

These single-stride target signals were then used as building blocks to assemble the full target signals used to train and test the networks. The training target signal consisted of 15 strides, with each of the three activities repeated for five consecutive strides. The test target signal consisted of 20 strides, with the three activities randomly interleaved. A different test signal was created for each of the twenty network instances. The same set of twenty signals was used across all parameter settings. The transitions between two consecutive strides were smoothed by leaving a one-sample gap between the two strides, which was then filled through linear interpolation. The full target signals were then filtered using a 20 Hz low-pass bi-directional Butterworth filter (2nd order). Finally, the first and last 50 samples were cut from the full signals to remove any unwanted filter artifacts. Only complete cycles were used to evaluate network performance, thereby excluding the transient phase of the network from the evaluation.

2.3.3 Training and testing the networks

The output weights of the networks (W_{ij}) were trained using the recursive least squares algorithm (Haykin, 2014, see also Appendix C). Unlike the internal connection weights, the output weights of the network were not restricted to be positive during training and can be either negative or positive. All networks were trained over 5 consecutive training iterations, with the network activations and their derivative reinitialized at the start of each iteration.

2.3.4 Outcome parameters

Performance quantification The performance of each network was quantified as the percentage of successfully reproduced strides. First, the root-mean-square error (RMSE) between the actual (dashed black lines, Figure 2) and target (coloured lines, Figure 2) output was calculated for each muscle and each of the 20 strides in the test signal. The pattern produced by each muscle during each stride was then classified as either a success if RMSE ≤ 0.05 (shaded in green in Figure 2) or a failure if RMSE > 0.05 (shaded in red in Figure 2).

Network dimensionality The richness of the internal network dynamics was quantified as the number of network principal components (NPCs) necessary to account for 99% of the variance in the neuron firing rates. The number of principal components in the output and target output were calculated in the same way.

Muscle parameters We calculated the co-activation index (*CAI*) between the muscle activation output patterns. For every muscle pair, the co-activation can be calculated as (Falconer, 1985; Souissi et al., 2017):

$$CAI(t) = \frac{2 \cdot abs(min(Y_{norm,1}(t), Y_{norm,2}(t)))}{abs\left(Y_{norm,1(t)}\right) + abs(Y_{norm,2(t)})} \tag{9}$$

Here, $Y_{norm,1}$ and $Y_{norm,2}$ are the activations of both muscles, normalized to their own maximum over the evaluated time period. The co-activation index is reported as an average over the entire test signal.

Output Variance We additionally calculated the variance between the different locomotor activities, to assess the network's ability to generate distinct signals for each task. After time-normalizing all strides from 0 to 100%, we computed the mean and variance across cycles for each of the three tasks individually. We could then determine the between-task variance by calculating the variance over the three resulting averaged cycles.

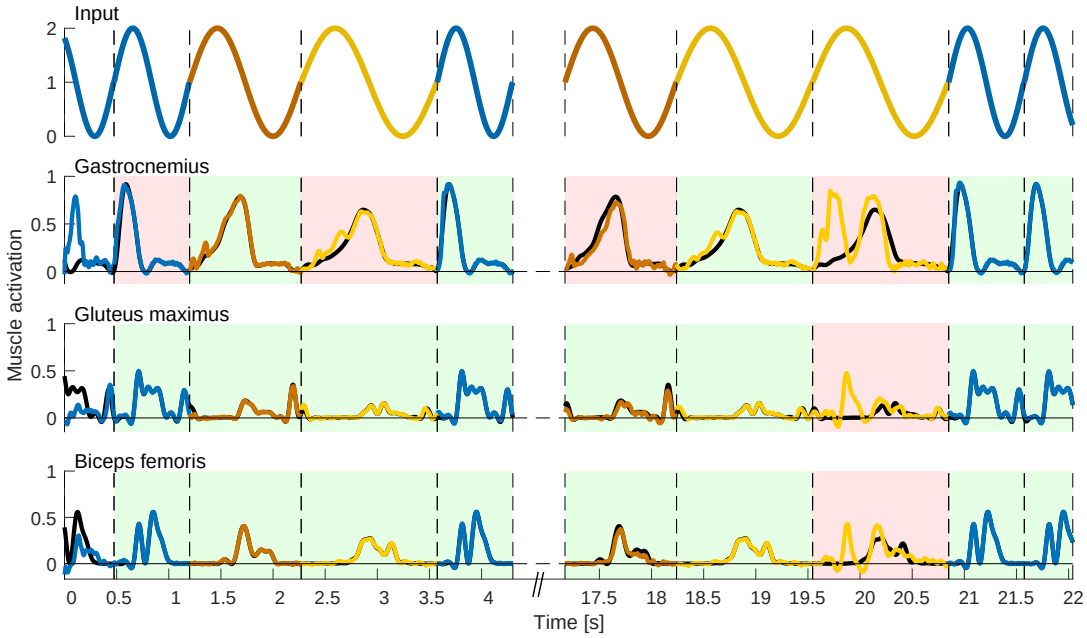

Figure 2 – Performance calculation in an example output. *The network is tested on its ability to reproduce the motor patterns of randomly interleaved gait cycles. The input to the network (top row) and network outputs of three example muscles (bottom rows) are shown for slow walking (yellow), fast walking (orange) and running (blue). The target output is indicated by the solid black line (bottom rows). The transitions between cycles are indicated as dashed vertical black lines. For each cycle and muscle, if the root mean square difference between the target and the actual output is less than 0.05, the cycle is classified as a success (green background). Otherwise, it is classified as a failure (red background). The performance of the network is the percentage of successful cycles over the full set of 17 muscles and 20 cycles. Note that the first cycle is not taken into account for the performance evaluation (white background) as it covers the initialization period of the network. The performance of the depicted network ($N_E = N_I = 375$, $p_E = p_I = 0.1$ and $g_E = g_I = 1.5$) is 71.7% (77.8% if only taking into account the visible cycles).*

3 Results

We present our findings on the impact of anatomical excitatory to inhibitory imbalance in recurrent neural networks trained to replicate locomotor signals. Our results include a theoretical assessment of network dynamics using mean-field theory, complemented by simulation analyses that explore the relationship between the imbalance, effective dimensionality, and motor outputs.

Mean field theory

A theoretical analysis using Mean Field Theory showed that the average input to a neuron ($\mu_i = [I_i]$) depends on the average firing rate φ and on the excitatory and inhibitory neuron numbers (N_E and N_I), connection percentages (p_E and p_I) and connection strengths (g_E and g_I, see Equation 3). We defined the anatomical imbalance (A) between excitation and inhibition in the network as the ratio between the average recurrent input and the average firing rate (Equation 8).

This analysis allowed us to predict the mean and variance of both the recurrent input and the firing rate, which showed a good match to simulations for all studied parameters (see Figures 3 and S.8). The MFT analysis showed that, for the analysed parameter ranges, the internal network dynamics do not depend on the total number of neurons or the mean connection percentage, but only on the imbalance and the total connection strength (Appendix A, Equations 38 and 39). While the total connection strengths have quantitative differences, they do exhibit qualitatively similar behaviour (Figure S.8).

Network Dynamics

As predicted by our mean-field theory analysis (Figure 3, black), we find that the network dynamics depend critically on the excitatory to inhibitory imbalance, regardless of whether it was varied in terms of neuron number (Figure 3, blue), connection percentage (Figure 3, green) or connection strength (Figure 3, yellow). The effect of the imbalance is asymmetrical: increasing the inhibition has a markedly different effect than increasing excitation.

When increasing the excitation in the network, we see a rapid, almost linear increase of the recurrent input with Imbalance A (Figure 3A). Here, the incoming excitatory current (dashed lines in Figure 3B) overpowers the incoming inhibitory current (dotted lines in Figure 3B). This large excitatory current causes the mean firing rate to saturate (Figure 3C), which not only keeps the neuronal input high, but also results in a drastic drop in the firing rate variance (Figure 3D). This is also evidenced by a strong negative skewness in the firing rate distributions (see Figure 3E, and example histograms in (Figure S.9).

In contrast, for inhibitory networks ($A < 0$), the mean input received by the neurons is negative but small (Figure 3A), with the recurrent inputs from both the excitatory and inhibitory populations relatively balanced and close to zero (Figure 3B). As a consequence, small variations in the inputs to individual neurons can drive variations in the neurons' firing rate, thereby preventing the firing rates from saturating toward their minimum value of zero (Figure 3C). The firing rate variance therefore also does not drop to zero (Figure 3D) and the firing rate distribution has a moderately positive skew (Figure 3E).

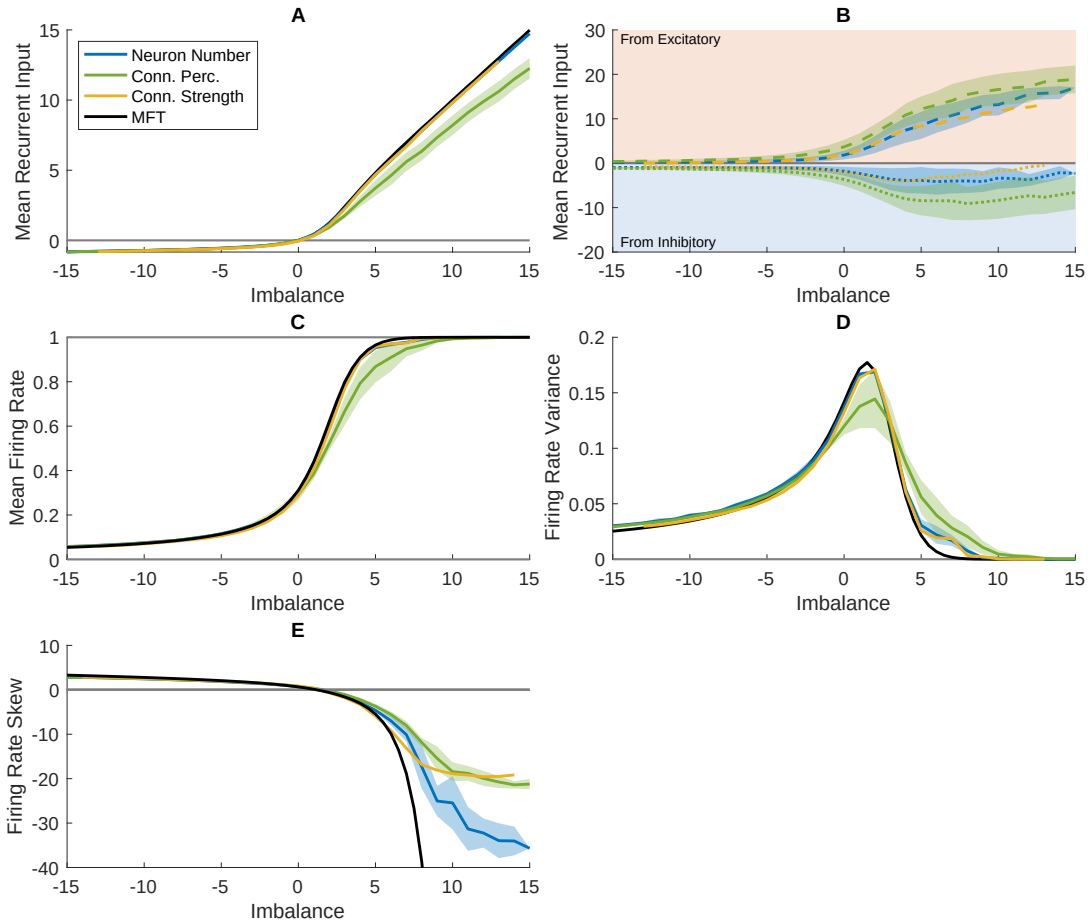

Figure 3 – Network Dynamics as a function of Imbalance **A**, *with (A) showing the mean input to individual neurons, (B) the mean input from the excitatory (positive; in the red-shaded area) and inhibitory (negative; in the blue-shaded area) populations, (C) the mean firing rate, (D) the firing rate variance, and (E) the skew of the firing rate distribution. The average overall levels of N_{tot} (blue) and p_{tot} (green) are depicted, with the shaded area showing the standard deviation. Only one level of g_{tot} is shown ($\sqrt{2 \cdot 1.5^2}$, matching the total connection strength used when varying the neuron number and connection percentage), as the total level influences the relationship between Imbalance and the firing rate (see Figure S.8). The black lines show the mean-field theory (MFT) predictions for each imbalance with $g_{tot} = \sqrt{2 \cdot 1.5^2}$. All simulation results have been averaged over 20 network instances.*

Dimensionality and Performance

Varying the imbalance led to a large variety in the effective dimensionality of the networks, with the number of NPCs ranging from 1 to 171 in individual networks. Due to the low variability in excitation-dominated networks, the dimensionality of the network drops drastically when excitation increases (see Figure 4A). This is accompanied by a concurrent drop in network performance (Figure 4B). While the

dimensionality in inhibitory networks is also lower than in balanced networks (Figure 4A), the number of principal components in the network stays well above the number of principal components in the target data (i.e. 9 PCs explaining > 99% of the output variance, as indicated by the pink line in Figure 4A,C). This means that the dimensionality of the inhibition-dominated networks is theoretically large enough to reproduce all dimensions present in the target output, as reflected by their relatively high performance (Figure 4B).

Indeed, we see that network performance depends critically on the effective dimensionality of the networks (see Figure 4C). For our selection of stable networks, we see an approximately linear increase in performance as the number of NPCs grows before plateauing and slowly dropping (Figure 4C). There thus seems to be an optimal window for the number of NPCs around approximately 40.This corresponds to 2.8-6.4 times the number of output principal components. While the influence of the imbalance on the effective dimensionality does depend on g_{tot}, N_{tot}, and p_{tot} (see Figure S.10), the influence of imbalance on performance is largely independent of these factors (Figure S.7, solid lines). This difference can be explained by the relatively wide plateau and slow drop-off in Figure 4C), indicating that a change in the number of effective dimensions doesn't necessarily lead to a large change in the performance.

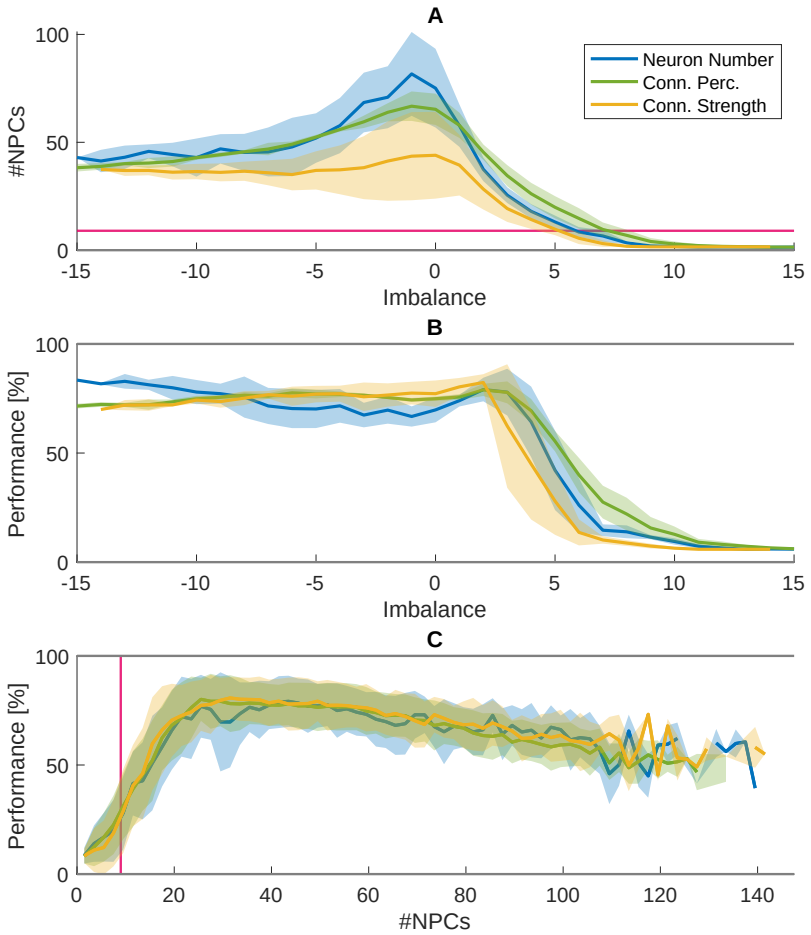

Figure 4 – Effective dimensionality and performance of the networks. *The imbalance A is plotted against (A) the number of network principal components (NPCs) and (B) the performance of the network. In (C), the network performance is plotted as a function of the number of NPCs. The average over 20 network instances and all included levels of N_{tot}, p_{tot} and g_{tot} are depicted, with the shaded area showing the standard deviation over these levels. In (A) and (C), the pink line indicates the number of principal components in the target output (i.e. 9).*

Motor patterns

Increasing the imbalance toward inhibition also has a markedly different effect on the motor outputs as compared to changing it towards excitation. Figure 5 shows the target as well as the actual network motor outputs for four example muscles for slow walking (red), fast walking (yellow) and running (blue). Balanced networks and networks with negative imbalance (first and second column in Figure 5) can produce distinct motor outputs for the three tasks. In contrast, in excitation-dominated networks (third, fourth and fifth column in Figure 5), the motor outputs for the three tasks become increasingly similar, which is also evidenced by the decrease in the variance between the locomotor tasks (see Figure 6B). In addition, the outputs of the different muscles become more similar. This is evidenced by an increase in the co-activation index *CAI* (see Figure 6A). The reduced variability and dimensionality in the excitatory networks are thus accompanied by a simplification of the networks' outputs.

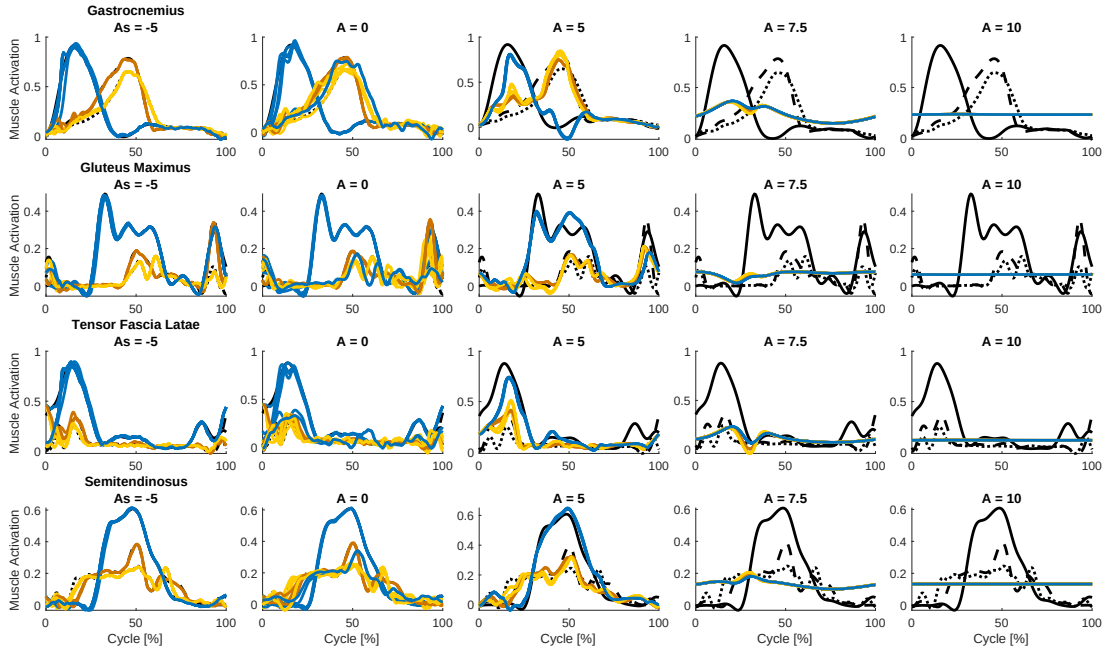

Figure 5 – Example output signals plotted as a function of the cycle percentage*, with 0% and 100% corresponding to right heel strikes. Each column shows signals belonging to a random example network with an imbalance of $A = -5, 0, +5$ and $+10$ from left to right. Each row shows the output corresponding to a different muscle. From top to bottom, these are the m. gastrocnemius, m. gluteus maximus, m. tensor fascia latae and m. semintendinosus. Black lines indicate the target output for each of the three activities, with solid lines representing running, dashed lines fast walking and dotted lines slow walking. Coloured lines indicate the actual network output, where blue depicts running, yellow fast walking and red slow walking. The example networks had a success percentage of 85%, 71%, 50%, 8% and 6% respectively.*

4 Discussion

This study aimed to explore how the ratio of recurrent inhibition to excitation affects the dynamics, performance and motor outputs of RNN-based CPG models, with a focus on pattern formation. Reservoir computing models were trained to recreate muscle activation patterns derived from human locomotion, using various ratios of recurrent excitation to inhibition. We performed a mean-field analysis and derived a composite measure called imbalance that quantifies the overall ratio of inhibition to excitation of the recurrent connections in the network. The mean-field analysis predicted that a high imbalance (i.e. where excitation $\gg$ inhibition) would lead to saturated firing rates and reduce the variability in the network. Conversely, inhibition-dominated networks were expected to maintain higher variability. Our simulations confirmed these predictions and additionally showed that inhibition is necessary both for good network performance and to recreate experimentally observed firing rate distributions. In contrast, networks with increased excitation tended to produce simplistic and repetitive motor patterns.

Many studies have investigated the balance between excitation and inhibition, both in biological and

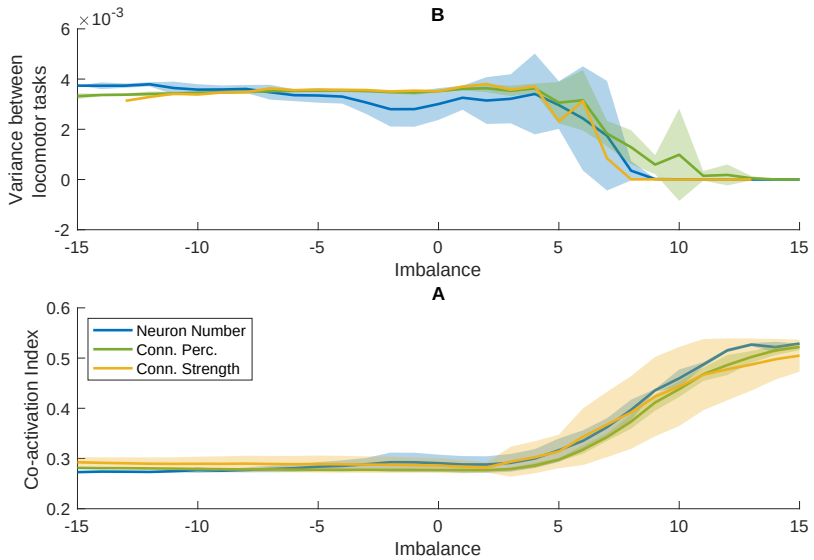

Figure 6 – Increased excitation leads to a simplification of the motor output, both between muscles as well as between locomotor tasks. *(A) average co-activation index between each pair of muscles and (B) variance between the outputs for the different locomotor tasks (i.e. slow walking, fast walking, and running), as a function of the imbalance (A), when changing the number of neurons (blue), connection percentage (green) and connection strength (yellow). Both panels represent averages over 20 network instances and all levels of N_{tot}, p_{tot} and g_{tot}, with the shaded area depicting the standard deviation over these levels. For panel (A), only $g_{tot} = \sqrt{2 \cdot 1.5^2}$ is shown, as the overall connection strength influenced the relationship between the imbalance and the variance (see Figure S.8C.*

artificial neural networks (reviewed in, e.g., Hennequin et al., 2017; Herstel & Wierenga, 2021; Isaacson & Scanziani, 2011; Liang et al., 2024; Sadeh & Clopath, 2021), demonstrating that a functional EI balance is crucial for proper neural functioning. This functional balance, however, is not necessarily the result of a strict underlying anatomical balance or symmetry. In biological networks, the anatomical characteristics of excitatory and inhibitory neurons do not follow a 1-to-1 ratio. For instance, there are typically more excitatory than inhibitory neurons: inhibitory neurons make up only 30 to 45% of the neurons in laminae I to III of the rat spinal cord (Todd & Sullivan, 1990), and only around 20% of the neurons in the mammalian cerebral cortex (Marín, 2012; Marom & Shahaf, 2002; Meinecke & Peters, 1987; Sahara et al., 2012; Wonders & Anderson, 2006). The connection probabilities are also not equal, as, in the cortex, only one in fifteen inputs to a given neuron is inhibitory (Megías et al., 2001; Peters, 2002). Conversely, inhibitory neurons have been found to synapse closer to the axon somata (Beaulieu et al., 1992; Peters, 2002), thereby increasing their effectiveness relative to the excitatory neurons (Chen et al., 2012; Markram et al., 2004). It is unclear, however, whether this compensates for the smaller number of connections, or whether there still is a resultant anatomical imbalance between excitation and inhibition. In this study, we examined the influence of the anatomical ratio between excitatory and inhibitory recurrent connections and showed that, as a natural consequence of the recurrent input, the net current to the neurons can still be balanced, even if the individual anatomical parameters are not symmetrical or imbalanced.

To investigate the effects of the anatomical EI balance, we introduced the anatomical *imbalance*, which quantifies the influence of the excitatory and inhibitory populations in the networks in one single variable. The imbalance has the same impact on the network dynamics, regardless of which of the underlying population characteristics (neuron number, connection percentage, and connection strength) are adjusted. The described relationship between the imbalance, network dynamics and performance only breaks down when other factors disrupt network functioning, e.g. when the connection strength is increased beyond the stable region (see Figure S.7A, $g_{tot} \geq 2.5$), or when the networks become too small (see Figure S.7). The imbalance and strength equations are first expressed in terms of only the mean, variance, and size of the two neuron pools. Therefore, the equations can readily be adapted to networks with different distributions and scaling factors, and we expect our results to generalize to these networks.

Our findings indicate that recurrent inhibition is not only important to maintain network performance, but also vital in replicating features observed in experimental recordings. First, recurrent inhibition

is necessary to recreate the positively skewed firing rate distributions observed in experimental studies. Specifically, networks must exhibit a slight anatomical imbalance favouring inhibition to recreate experimentally observed firing rate skewness. Excitation-dominated networks have negatively skewed distributions, with a prominent peak at the maximum firing rate. Increasing the inhibition broadens the distributions and causes the peak to shift leftward, thereby matching the lognormal firing rate distributions that have been experimentally recorded in the spinal cord (Petersen & Berg, 2016) and various cortical areas (Hromádka et al., 2008; O'Connor et al., 2010; Shafi et al., 2007) across different species. These findings are consistent with previous works, both computational and experimental, that showed that balanced networks have firing rate distributions that are relatively wide and skewed at low average firing rates (Hennequin et al., 2017; Lindén & Berg, 2021; Roxin et al., 2011; Van Vreeswijk & Sompolinsky, 1996; Vogels et al., 2005), whereas excitatory networks exhibit a distribution with a single peak at the maximum firing rate (Lindén & Berg, 2021). Similarly, Roxin et al. (2011) demonstrated *in silico* that increasing the firing rate of inhibitory neurons via the external input transforms the firing rate distribution from negatively skewed to more lognormal distributions, thereby matching experimental results recorded *in vivo*. This shift means that balanced networks can respond to external inputs in a graded manner, rather than the all-or-none firing rate seen in excitatory networks (Lindén & Berg, 2021). Our study adds that the same is true for inhibition-dominated networks.

Second, inhibition is important to prevent a loss of the effective dimensionality of the network. The dimensionality of the network is crucial for network performance (Carroll, 2020a; Carroll & Pecora, 2019; Susman et al., 2021). In our study, networks with too few principal components were unable to replicate the target output signals. This matches with prior fin findings, where increasing the rank of the reservoir matrix (which places an upper bound on the number of meaningful principal components that can be extracted) is associated with a decrease in the test error (Carroll & Pecora, 2019). In excitatory networks, the effective dimensionality of the reservoir is reduced below the dimensionality of the output, and thus insufficient to reproduce the required output patterns. This is caused by run-away activity in the network leading to saturated firing rates, which can be stabilized by sufficient inhibition (Herstel & Wierenga, 2021; Sadeh & Clopath, 2021). These results highlight the importance of recurrent inhibition in maintaining appropriate network dimensionality.

Third, our results indicate that inhibition plays a crucial role in generating diverse and variable motor patterns. High excitation networks were unable to produce different outputs for different muscles, as evidenced by the high co-activation and low number of output principal components. Nor were they able to differentiate between the different locomotor tasks, as evidenced by the drop in the between-task variability. These findings align with *in vivo* studies, where blocking inhibition led to synchronous activation of flexor and extensor motoneurons (Beato & Nistri, 1999; Cazalets et al., 1998; Cowley & Schmidt, 1995; Kiehn, 2006).

The importance of inhibition in the generation of diverse movements is also supported by the reduced complexity of muscle activation patterns that have been observed in various motor disorders. For instance, in cerebral palsy, there is evidence of a decreased inhibition (reviewed by Fogarty, 2023) as well as a reduction in the number of motor modules during walking (Steele et al., 2015; Tang et al., 2015), indicating a reduced locomotor complexity. Condliffe et al. (2016) showed that poor motor ability in cerebral palsy is associated with reduced inhibitory control of motoneurons. Similarly, stroke patients also show a loss of motor modules (Cheung et al., 2012; Clark et al., 2010), which is associated with reduced walking speed (Routson et al., 2013) and a decreased ability to adapt speed, timing, and step characteristics in mobility tasks (Routson et al., 2014). In stroke, however, the role of inhibition is less clear, with inhibition seemingly having a protective impact post-stroke by counteracting excitotoxicity and the associated increase in EI balance (Sydserff et al., 1995), but hampering post-stroke recovery (Grigoras & Stagg, 2021). The early neuroprotective effect of inhibition has been shown to be associated with improved functional ability (Marshall et al., 1999), which matches our results.

Based on these findings, we suggest that dystonia may be the result of reduced inhibition in spinal circuits causing excessive co-contraction. Dystonias are movement disorders characterized by spontaneous muscle activity at rest, large co-contraction of flexors and extensors during voluntary movements such as walking (L. G. Cohen & Hallett, 1988; Rothwell et al., 1983), and reduced reciprocal inhibition (Nakashima et al., 1989; Panizza et al., 1990). Dystonias are thought to result from basal ganglia dysfunction, but neurophysiological evidence for this has remained elusive. The most prevalent genetic form of dystonia is early-onset generalized torsion dystonia, commonly caused by a mutation of the TOR1A gene (Ozelius et al., 1999). When this gene is deleted in spinal circuits in mice, the mice develop dystonia

with symptoms corresponding to those of humans (Pocratsky et al., 2023). These recent experimental results thus suggest that dystonia may be a disorder of spinal circuits, rather than basal ganglia circuits as commonly assumed. We suggest that this disorder may be due to reduced inhibition in spinal circuits causing excessive co-contraction. Indeed, the most effective treatment for dystonia is long-term deep brain stimulation, which has been shown to improve reciprocal inhibition and reduce agonist-antagonist co-contraction (Tisch et al., 2006).

Such findings are not limited to motor disorders alone. In higher neural centres, an experimental disruption of the excitatory to inhibitory (EI) balance towards excitation can result in a variety of pathologies, such as epilepsy (Dichter & Ayala, 1987; Dudek & Sutula, 2007), autism (Casanova et al., 2003; Markicevic et al., 2020; J. L. R. Rubenstein & Merzenich, 2003; J. L. Rubenstein, 2010) and schizophrenia (Murray et al., 2014; Yizhar et al., 2011). Similarly, disinhibition in the spinal cord has been shown to lead to allodynia (Lee et al., 2019; Yaksh, 1989). While these disorders are not classified as motor disorders, they are often accompanied by changes in motor control. In autism spectrum disorder, for example, individuals often display atypical motor behaviours, including atypical posture and gait, hypotonia, and impaired bilateral coordination, balance and fine motor skills (De Jong et al., 2011; Jansiewicz et al., 2006; Zampella et al., 2021) and in schizophrenia, motor abnormalities such as catatonia have been linked to EI imbalance (Walther & Strik, 2012).

While this study provides valuable insights into the importance of recurrent inhibition, several limitations must be acknowledged. First, this study was limited to three activities, measured in just a single participant. Due to this limited activity range, the networks were only trained on three output frequencies. As the frequency for which the network performs optimally depends on factors such as the connection strength or the target amplitude (Susman et al., 2021), our results might not fully generalize to untested frequencies. Second, our results are based on firing rate, rather than spiking, neural networks. It is unclear whether spiking networks exhibit the same robustness to inhibition as our firing rate networks, or if excessive inhibition in spiking networks would result in quiescence. Third, the fixed connectivity within the neural reservoir limits the biological realism of our model. Although computationally efficient, the static reservoir connections preclude our network from synaptic plasticity, which plays an important role in maintaining functional EI balance (Field et al., 2020; Froemke, 2015; Vogels et al., 2011). This thus limits the model's ability to simulate the dynamic adjustments observed in biological neural networks. However, for our study, this fixed structure was necessary to study the effects of the excitation-to-inhibition ratio, as neural plasticity would change the networks' imbalance and thus skew the results. Future work could focus on the interplay between initial imbalance and synaptic plasticity.

In conclusion, our study highlights the role of recurrent inhibition and anatomical EI balance in generating complex and coordinated motor patterns. Networks that were dominated by inhibition displayed robust and complex firing patterns, and were thus able to learn and reproduce human locomotor patterns effectively. In contrast, networks that were dominated by excitation exhibited a reduced effective dimensionality and simplified motor outputs. These findings demonstrate the potential for neural networks to tolerate excessive inhibition, but not increased excitation. The inclusion of sufficient recurrent inhibition was not only essential in generating the locomotor patterns; it was also necessary to reproduce experimental observations of skewed firing rates. Future work should explore a wider range of motor tasks to test and expand our findings, and investigate if they generalize to spiking neural networks.

References

Aljadeff, J., Stern, M., & Sharpee, T. (2015). Transition to Chaos in Random Networks with Cell-Type-Specific Connectivity. *Physical Review Letters*, *114*(8), 088101. https://doi.org/10.1103/PhysRevLett.114.088101

Beato, M., & Nistri, A. (1999). Interaction Between Disinhibited Bursting and Fictive Locomotor Patterns in the Rat Isolated Spinal Cord. *Journal of Neurophysiology*, *82*(5), 2029–2038. https://doi.org/10.1152/jn.1999.82.5.2029

Beaulieu, C., Kisvarday, Z., Somogyi, P., Cynader, M., & Cowey, A. (1992). Quantitative Distribution of GABA-immunopositive and-immunonegative Neurons and Synapses in the Monkey Striate Cortex (Area 17). *Cerebral Cortex*, *2*(4), 295–309. https://doi.org/10.1093/cercor/2.4.295

Berg, R. W., Alaburda, A., & Hounsgaard, J. (2007). Balanced Inhibition and Excitation Drive Spike Activity in Spinal Half-Centers. *Science*, *315*(5810), 390–393. https://doi.org/10.1126/science.1134960

Berg, R. W., Willumsen, A., & Lindén, H. (2019). When networks walk a fine line: Balance of excitation and inhibition in spinal motor circuits. *Current Opinion in Physiology*, *8*, 76–83. https://doi.org/10.1016/j.cophys.2019.01.006

Boström, K. J., Wagner, H., Prieske, M., & De Lussanet, M. (2013). Model for a flexible motor memory based on a self-active recurrent neural network. *Human Movement Science*, *32*(5), 880–898. https://doi.org/10.1016/j.humov.2013.07.003

Britz, O., Zhang, J., Grossmann, K. S., Dyck, J., Kim, J. C., Dymecki, S., Gosgnach, S., & Goulding, M. (2015). A genetically defined asymmetry underlies the inhibitory control of flexor-extensor locomotor movements. *eLife*, *4*, e04718. https://doi.org/10.7554/eLife.04718

Brown, T. G. (1911). The intrinsic factors in the act of progression in the mammal. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, *84*(572), 308–319. https://doi.org/10.1098/rspb.1911.0077

Carroll, T. L. (2020a). Dimension of reservoir computers. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *30*(1), 013102. https://doi.org/10.1063/1.5128898

Carroll, T. L., & Pecora, L. M. (2019). Network structure effects in reservoir computers. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *29*(8), 083130. https://doi.org/10.1063/1.5097686

Casanova, M. F., Buxhoeveden, D., & Gomez, J. (2003). Disruption in the Inhibitory Architecture of the Cell Minicolumn: Implications for Autisim. *The Neuroscientist*, *9*(6), 496–507. https://doi.org/10.1177/1073858403253552

Cazalets, J.-R., Bertrand, S., Sqalli-Houssaini, Y., & Clarac, F. (1998). Gabaergic Control of Spinal Locomotor Networks in the Neonatal Rat. *Annals of the New York Academy of Sciences*, *860*(1), 168–180. https://doi.org/10.1111/j.1749-6632.1998.tb09047.x

Chen, J. L., Villa, K. L., Cha, J. W., So, P. T., Kubota, Y., & Nedivi, E. (2012). Clustered Dynamics of Inhibitory Synapses and Dendritic Spines in the Adult Neocortex. *Neuron*, *74*(2), 361–373. https://doi.org/10.1016/j.neuron.2012.02.030

Cheung, V. C. K., Turolla, A., Agostini, M., Silvoni, S., Bennis, C., Kasi, P., Paganoni, S., Bonato, P., & Bizzi, E. (2012). Muscle synergy patterns as physiological markers of motor cortical damage. *Proceedings of the National Academy of Sciences*, *109*(36), 14652–14656. https://doi.org/10.1073/pnas.1212056109

Clark, D. J., Ting, L. H., Zajac, F. E., Neptune, R. R., & Kautz, S. A. (2010). Merging of Healthy Motor Modules Predicts Reduced Locomotor Performance and Muscle Coordination Complexity Post-Stroke. *Journal of Neurophysiology*, *103*(2), 844–857. https://doi.org/10.1152/jn.00825.2009

Cohen, A. H., & Harris-Warrick, R. M. (1984). Strychnine eliminates alternating motor output during fictive locomotion in the lamprey. *Brain Research*, *293*(1), 164–167. https://doi.org/10.1016/0006-8993(84)91464-1

Cohen, L. G., & Hallett, M. (1988). Hand cramps: Clinical features and electromyographic patterns in a focal dystonia. *Neurology*, *38*(7), 1005–1012. https://doi.org/10.1212/wnl.38.7.1005

Condliffe, E. G., Jeffery, D. T., Emery, D. J., & Gorassini, M. A. (2016). Spinal inhibition and motor function in adults with spastic cerebral palsy. *The Journal of Physiology*, *594*(10), 2691–2705. https://doi.org/10.1113/JP271886

Cowley, K. C., & Schmidt, B. J. (1995). Effects of inhibitory amino acid antagonists on reciprocal inhibitory interactions during rhythmic motor activity in the in vitro neonatal rat spinal cord. *Journal of Neurophysiology*, *74*(3), 1109–1117. https://doi.org/10.1152/jn.1995.74.3.1109

Delcomyn, F. (1980). Neural Basis of Rhythmic Behavior in Animals. *Science*, *210*(4469), 492–498. https://doi.org/10.1126/science.7423199

Dichter, M. A., & Ayala, G. F. (1987). Cellular Mechanisms of Epilepsy: A Status Report. *Science*, *237*(4811), 157–164. https://doi.org/10.1126/science.3037700

Dudek, F. E., & Sutula, T. P. (2007). Epileptogenesis in the dentate gyrus: A critical perspective. In *Progress in Brain Research* (pp. 755–773, Vol. 163). Elsevier. https://doi.org/10.1016/S0079-6123(07)63041-6

Eccles, J. (1976). From electrical to chemical transmission in the central nervous system. *Notes and Records of the Royal Society of London*, *30*(2), 219–230. https://doi.org/10.1098/rsnr.1976.0015

Falconer, K. (1985). Quantitative assessment of co-contraction at the ankle joint in walking. *Electromyography and Clinical Neurophysiology*, *25*(2-3), 135–149.

Field, R. E., D'amour, J. A., Tremblay, R., Miehl, C., Rudy, B., Gjorgjieva, J., & Froemke, R. C. (2020). Heterosynaptic Plasticity Determines the Set Point for Cortical Excitatory-Inhibitory Balance. *Neuron*, *106*(5), 842–854.e4. https://doi.org/10.1016/j.neuron.2020.03.002

Fogarty, M. J. (2023). Inhibitory Synaptic Influences on Developmental Motor Disorders. *International Journal of Molecular Sciences*, *24*(8), 6962. https://doi.org/10.3390/ijms24086962

Froemke, R. C. (2015). Plasticity of Cortical Excitatory-Inhibitory Balance. *Annual Review of Neuroscience*, *38*(Volume 38, 2015), 195–219. https://doi.org/10.1146/annurev-neuro-071714-034002

Funahashi, K.-i., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, *6*(6), 801–806. https://doi.org/10.1016/S0893-6080(05)80125-X

Griener, A., Dyck, J., & Gosgnach, S. (2013). Regional distribution of putative rhythm-generating and pattern-forming components of the mammalian locomotor CPG. *Neuroscience*, *250*, 644–650. https://doi.org/10.1016/j.neuroscience.2013.07.070

Grigoras, I.-F., & Stagg, C. J. (2021). Recent advances in the role of excitation–inhibition balance in motor recovery post-stroke. *Faculty Reviews*, *10*, 58. https://doi.org/10.12703/r/10-58

Grillner, S. (1981, December). Control of Locomotion in Bipeds, Tetrapods, and Fish. In R. Terjung (Ed.), *Comprehensive Physiology* (1st ed., pp. 1179–1236). Wiley. https://doi.org/10.1002/cphy.cp010226

Grillner, S., Ekeberg, Ö., El Manira, A., Lansner, A., Parker, D., Tegnér, J., & Wallén, P. (1998). Intrinsic function of a neuronal network — a vertebrate central pattern generator. *Brain Research Reviews*, *26*(2-3), 184–197. https://doi.org/10.1016/S0165-0173(98)00002-2

Grillner, S., & Wallén, P. (1985). Central Pattern Generators for Locomotion, with Special Reference to Vertebrates. *Annual Review of Neuroscience*, *8*(1), 233–261. https://doi.org/10.1146/annurev.ne.08.030185.001313

Harish, O., & Hansel, D. (2015). Asynchronous Rate Chaos in Spiking Neuronal Circuits. *PLOS Computational Biology*, *11*(7), e1004266. https://doi.org/10.1371/journal.pcbi.1004266

Hatze, H. (1980). A mathematical model for the computational determination of parameter values of anthropomorphic segments. *Journal of Biomechanics*, *13*(10), 833–843. https://doi.org/10.1016/0021-9290(80)90171-2

Haykin, S. S. (2014). *Adaptive filter theory* (Fifth, international edition). Pearson.

Hennequin, G., Agnes, E. J., & Vogels, T. P. (2017). Inhibitory Plasticity: Balance, Control, and Codependence. *Annual Review of Neuroscience*, *40*(1), 557–579. https://doi.org/10.1146/annurev-neuro-072116-031005

Herstel, L. J., & Wierenga, C. J. (2021). Network control through coordinated inhibition. *Current Opinion in Neurobiology*, *67*, 34–41. https://doi.org/10.1016/j.conb.2020.08.001

Herzog, W. (1999). Muscle. In B. M. Nigg & W. Herzog (Eds.), *Biomechanics of the Musculo-skeletal system*. Wiley, Chichester.

Hinaut, X., & Dominey, P. F. (2013). Real-Time Parallel Processing of Grammatical Structure in the Fronto-Striatal System: A Recurrent Network Simulation Study Using Reservoir Computing (T. Boraud, Ed.). *PLoS ONE*, *8*(2), e52946. https://doi.org/10.1371/journal.pone.0052946

Hoellinger, T., Petieau, M., Duvinage, M., Castermans, T., Seetharaman, K., Cebolla, A.-M., Bengoetxea, A., Ivanenko, Y., Dan, B., & Cheron, G. (2013). Biological oscillations for learning walking coordination: Dynamic recurrent neural network functionally models physiological central pattern generator. *Frontiers in Computational Neuroscience*, *7*. https://doi.org/10.3389/fncom.2013.00070

Hromádka, T., DeWeese, M. R., & Zador, A. M. (2008). Sparse Representation of Sounds in the Unanesthetized Auditory Cortex. *PLOS Biology*, *6*(1), e16. https://doi.org/10.1371/journal.pbio.0060016

Ijspeert, A. J. (2008). Central pattern generators for locomotion control in animals and robots: A review. *Neural Networks*, *21*(4), 642–653. https://doi.org/10.1016/j.neunet.2008.03.014

Isaacson, J. S., & Scanziani, M. (2011). How Inhibition Shapes Cortical Activity. *Neuron*, *72*(2), 231–243. https://doi.org/10.1016/j.neuron.2011.09.027

Ivanenko, Y. P., Poppele, R. E., & Lacquaniti, F. (2006). Spinal Cord Maps of Spatiotemporal Alpha-Motoneuron Activation in Humans Walking at Different Speeds. *Journal of Neurophysiology*, *95*(2), 602–618. https://doi.org/10.1152/jn.00767.2005

Ivanenko, Y. P., Poppele, R. E., & Lacquaniti, F. (2004). Five basic muscle activation patterns account for muscle activity during human locomotion. *The Journal of Physiology*, *556*(1), 267–282. https://doi.org/10.1113/jphysiol.2003.057174

Jaeger, H. (2001). *The "echo state" approach to analysing and training recurrent neural networks – with an Erratum note* (GMD Report No. 148). GMD - German National Research Institute for Computer Science. http://www.faculty.jacobs-university.de/hjaeger/pubs/EchoStatesTechRep.pdf

Jankowska, E., Jukes, M. G. M., Lund, S., & Lundberg, A. (1967). The Effect of DOPA on the Spinal Cord 5. Reciprocal organization of pathways transmitting excitatory action to alpha motoneurones of flexors and extensors. *Acta Physiologica Scandinavica*, *70*(3-4), 369–388. https://doi.org/10.1111/j.1748-1716.1967.tb03636.x

Jankowska, E. (2008). Spinal interneuronal networks in the cat: Elementary components. *Brain Research Reviews*, *57*(1), 46–55. https://doi.org/10.1016/j.brainresrev.2007.06.022

Jansiewicz, E. M., Goldberg, M. C., Newschaffer, C. J., Denckla, M. B., Landa, R., & Mostofsky, S. H. (2006). Motor Signs Distinguish Children with High Functioning Autism and Asperger's Syndrome from Controls. *Journal of Autism and Developmental Disorders*, *36*(5), 613–621. https://doi.org/10.1007/s10803-006-0109-y

Jong, M. de, Punt, M., Groot, E. de, Minderaa, R. B., & Hadders-Algra, M. (2011). Minor neurological dysfunction in children with autism spectrum disorder. *Developmental Medicine & Child Neurology*, *53*(7), 641–646. https://doi.org/10.1111/j.1469-8749.2011.03971.x

Kadmon, J., & Sompolinsky, H. (2015). Transition to Chaos in Random Neuronal Networks. *Physical Review X*, *5*(4), 041030. https://doi.org/10.1103/PhysRevX.5.041030

Kiehn, O. (2006). Locomotor Circuits in the Mammalian Spinal Cord. *Annual Review of Neuroscience*, *29*(1), 279–306. https://doi.org/10.1146/annurev.neuro.29.051605.112910

Lacquaniti, F., Ivanenko, Y. P., & Zago, M. (2012). Patterned control of human locomotion. *The Journal of physiology*, *590*(Pt 10), 2189–2199. https://doi.org/10.1113/jphysiol.2011.215137

Lafreniere-Roula, M., & McCrea, D. A. (2005). Deletions of Rhythmic Motoneuron Activity During Fictive Locomotion and Scratch Provide Clues to the Organization of the Mammalian Central Pattern Generator. *Journal of Neurophysiology*, *94*(2), 1120–1132. https://doi.org/10.1152/jn.00216.2005

Lanuza, G. M., Gosgnach, S., Pierani, A., Jessell, T. M., & Goulding, M. (2004). Genetic Identification of Spinal Interneurons that Coordinate Left-Right Locomotor Activity Necessary for Walking Movements. *Neuron*, *42*(3), 375–386. https://doi.org/10.1016/S0896-6273(04)00249-1

Latash, M. L. (2012). Neurophysiological structures. In *Fundamentals of Motor Control* (pp. 171–210). Elsevier. https://doi.org/10.1016/B978-0-12-415956-3.00010-5

Lee, K. Y., Ratté, S., & Prescott, S. A. (2019). Excitatory neurons are more disinhibited than inhibitory neurons by chloride dysregulation in the spinal dorsal horn. *eLife*, *8*, e49753. https://doi.org/10.7554/eLife.49753

Li, W., Szczecinski, N. S., & Quinn, R. D. (2017). A neural network with central pattern generators entrained by sensory feedback controls walking of a bipedal model. *Bioinspiration & Biomimetics*, *12*(6), 065002. https://doi.org/10.1088/1748-3190/aa8290

Liang, J., Yang, Z., & Zhou, C. (2024). Excitation–Inhibition Balance, Neural Criticality, and Activities in Neuronal Circuits. *The Neuroscientist*, 10738584231221766. https://doi.org/10.1177/10738584231221766

Lindén, H., & Berg, R. W. (2021). Why Firing Rate Distributions Are Important for Understanding Spinal Central Pattern Generators. *Frontiers in Human Neuroscience*, *15*, 719388. https://doi.org/10.3389/fnhum.2021.719388

Marín, O. (2012). Interneuron dysfunction in psychiatric disorders. *Nature Reviews Neuroscience*, *13*(2), 107–120. https://doi.org/10.1038/nrn3155

Markicevic, M., Fulcher, B. D., Lewis, C., Helmchen, F., Rudin, M., Zerbi, V., & Wenderoth, N. (2020). Cortical Excitation:Inhibition Imbalance Causes Abnormal Brain Network Dynamics as Observed in Neurodevelopmental Disorders. *Cerebral Cortex*, *30*(9), 4922–4937. https://doi.org/10.1093/cercor/bhaa084

Markin, S. N., Lemay, M. A., Prilutsky, B. I., & Rybak, I. A. (2012). Motoneuronal and muscle synergies involved in cat hindlimb control during fictive and real locomotion: A comparison study. *Journal of Neurophysiology*, *107*(8), 2057–2071. https://doi.org/10.1152/jn.00865.2011

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., & Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nature Reviews Neuroscience*, *5*(10), 793–807. https://doi.org/10.1038/nrn1519

Marom, S., & Shahaf, G. (2002). Development, learning and memory in large random networks of cortical neurons: Lessons beyond anatomy. *Quarterly Reviews of Biophysics*, *35*(1), 63–87. https://doi.org/10.1017/S0033583501003742

Marshall, J. W., Cross, A. J., & Ridley, R. M. (1999). Functional benefit from clomethiazole treatment after focal cerebral ischemia in a nonhuman primate species. *Experimental Neurology*, *156*(1), 121–129. https://doi.org/10.1006/exnr.1998.6994

Mastrogiuseppe, F., & Ostojic, S. (2017). Intrinsically-generated fluctuating activity in excitatory-inhibitory networks (P. E. Latham, Ed.). *PLOS Computational Biology*, *13*(4), e1005498. https://doi.org/10.1371/journal.pcbi.1005498

Mastrogiuseppe, F., & Ostojic, S. (2018). Linking Connectivity, Dynamics, and Computations in Low-Rank Recurrent Neural Networks. *Neuron*, *99*(3), 609–623.e29. https://doi.org/10.1016/j.neuron.2018.07.003

McCrea, D. A., & Rybak, I. A. (2007, January). Modeling the mammalian locomotor CPG: Insights from mistakes and perturbations. In P. Cisek, T. Drew & J. F. Kalaska (Eds.), *Progress in Brain Research* (pp. 235–253, Vol. 165). Elsevier. https://doi.org/10.1016/S0079-6123(06)65015-2

McCrea, D. A., & Rybak, I. A. (2008). Organization of mammalian locomotor rhythm and pattern generation. *Brain Research Reviews*, *57*(1), 134–146. https://doi.org/10.1016/j.brainresrev.2007.08.006

Megías, M., Emri, Z., Freund, T., & Gulyás, A. (2001). Total number and distribution of inhibitory and excitatory synapses on hippocampal CA1 pyramidal cells. *Neuroscience*, *102*(3), 527–540. https://doi.org/10.1016/S0306-4522(00)00496-6

Meinecke, D. L., & Peters, A. (1987). GABA immunoreactive neurons in rat visual cortex. *The Journal of Comparative Neurology*, *261*(3), 388–404. https://doi.org/10.1002/cne.902610305

Murray, J. D., Anticevic, A., Gancsos, M., Ichinose, M., Corlett, P. R., Krystal, J. H., & Wang, X.-J. (2014). Linking Microcircuit Dysfunction to Cognitive Impairment: Effects of Disinhibition Associated with Schizophrenia in a Cortical Working Memory Model. *Cerebral Cortex*, *24*(4), 859–872. https://doi.org/10.1093/cercor/bhs370

Nakashima, K., Rothwell, J. C., Day, B. L., Thompson, P. D., Shannon, K., & Marsden, C. D. (1989). Reciprocal inhibition between forearm muscles in patients with writer's cramp and other occupational cramps, symptomatic hemidystonia and hemiparesis due to stroke. *Brain: A Journal of Neurology*, *112 (Pt 3)*, 681–697. https://doi.org/10.1093/brain/112.3.681

O'Connor, D. H., Peron, S. P., Huber, D., & Svoboda, K. (2010). Neural Activity in Barrel Cortex Underlying Vibrissa-Based Object Localization in Mice. *Neuron*, *67*(6), 1048–1061. https://doi.org/10.1016/j.neuron.2010.08.026

Ozelius, L. J., Page, C. E., Klein, C., Hewett, J. W., Mineta, M., Leung, J., Shalish, C., Bressman, S. B., Leon, D. de, Brin, M. F., Fahn, S., Corey, D. P., & Breakefield, X. O. (1999). The TOR1A (DYT1) gene family and its role in early onset torsion dystonia. *Genomics*, *62*(3), 377–384. https://doi.org/10.1006/geno.1999.6039

Panizza, M., Lelli, S., Nilsson, J., & Hallett, M. (1990). H-reflex recovery curve and reciprocal inhibition of H-reflex in different kinds of dystonia. *Neurology*, *40*(5), 824–828. https://doi.org/10.1212/wnl.40.5.824

Peters, A. (2002). Examining neocortical circuits: Some background and facts. *Journal of Neurocytology*, *31*(3-5 SPEC. ISS.), 183–193. https://doi.org/10.1023/A:1024157522651

Petersen, P. C., & Berg, R. W. (2016). Lognormal firing rate distribution reveals prominent fluctuation–driven regime in spinal motor networks. *eLife*, *5*, e18805. https://doi.org/10.7554/eLife.18805

Petersen, P. C., Vestergaard, M., Jensen, K. H. R., & Berg, R. W. (2014). Premotor Spinal Network with Balanced Excitation and Inhibition during Motor Patterns Has High Resilience to Structural Division. *The Journal of Neuroscience*, *34*(8), 2774–2784. https://doi.org/10.1523/JNEUROSCI.3349-13.2014

Pocratsky, A. M., Nascimento, F., Özyurt, M. G., White, I. J., Sullivan, R., O'Callaghan, B. J., Smith, C. C., Surana, S., Beato, M., & Brownstone, R. M. (2023). Pathophysiology of Dyt1-Tor1a dystonia in mice is mediated by spinal neural circuit dysfunction. *Science Translational Medicine*, *15*(694), eadg3904. https://doi.org/10.1126/scitranslmed.adg3904

Quinlan, K. A., & Kiehn, O. (2007). Segmental, Synaptic Actions of Commissural Interneurons in the Mouse Spinal Cord. *Journal of Neuroscience*, *27*(24), 6521–6530. https://doi.org/10.1523/JNEUROSCI.1618-07.2007

Rajagopal, A., Dembia, C. L., DeMers, M. S., Delp, D. D., Hicks, J. L., & Delp, S. L. (2016). Full-body musculoskeletal model for muscle-driven simulation of human gait. *IEEE Transactions on Biomedical Engineering*, *63*(10), 2068–2079. https://doi.org/10.1109/tbme.2016.2586891

Rajan, K., & Abbott, L. F. (2006). Eigenvalue Spectra of Random Matrices for Neural Networks. *Physical Review Letters*, *97*(18), 188104. https://doi.org/10.1103/PhysRevLett.97.188104

Rajan, K., Abbott, L. F., & Sompolinsky, H. (2010). Stimulus-dependent suppression of chaos in recurrent neural networks. *Physical Review E*, *82*(1), 011903. https://doi.org/10.1103/PhysRevE.82.011903

Rothwell, J. C., Obeso, J. A., Day, B. L., & Marsden, C. D. (1983). Pathophysiology of dystonias. *Advances in Neurology*, *39*, 851–863.

Routson, R. L., Clark, D. J., Bowden, M. G., Kautz, S. A., & Neptune, R. R. (2013). The influence of locomotor rehabilitation on module quality and post-stroke hemiparetic walking performance. *Gait & Posture*, *38*(3), 511–517. https://doi.org/10.1016/j.gaitpost.2013.01.020

Routson, R. L., Kautz, S. A., & Neptune, R. R. (2014). Modular organization across changing task demands in healthy and poststroke gait. *Physiological Reports*, *2*(6), e12055. https://doi.org/10.14814/phy2.12055

Roxin, A., Brunel, N., Hansel, D., Mongillo, G., & Vreeswijk, C. van. (2011). On the Distribution of Firing Rates in Networks of Cortical Neurons. *The Journal of Neuroscience*, *31*(45), 16217–16226. https://doi.org/10.1523/JNEUROSCI.1677-11.2011

Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: Increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, *2*(5), 255–267. https://doi.org/10.1034/j.1601-183X.2003.00037.x

Rubenstein, J. L. (2010). Three hypotheses for developmental defects that may underlie some forms of autism spectrum disorder. *Current Opinion in Neurology*, *23*(2), 118–123. https://doi.org/10.1097/WCO.0b013e328336eb13

Rybak, I. A., Shevtsova, N. A., Lafreniere-Roula, M., & McCrea, D. A. (2006). Modelling spinal circuitry involved in locomotor pattern generation: Insights from deletions during fictive locomotion. *The Journal of Physiology*, *577*(Pt 2), 617–639. https://doi.org/10.1113/jphysiol.2006.118703

Sadeh, S., & Clopath, C. (2021). Inhibitory stabilization and cortical computation. *Nature Reviews Neuroscience*, *22*(1), 21–37. https://doi.org/10.1038/s41583-020-00390-z

Sahara, S., Yanagawa, Y., O'Leary, D. D. M., & Stevens, C. F. (2012). The fraction of cortical GABAergic neurons is constant from near the start of cortical neurogenesis to adulthood. *Journal of Neuroscience*, *32*(14), 4755–4761. https://doi.org/10.1523/JNEUROSCI.6412-11.2012

Shafi, M., Zhou, Y., Quintana, J., Chow, C., Fuster, J., & Bodner, M. (2007). Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience*, *146*(3), 1082–1108. https://doi.org/10.1016/j.neuroscience.2006.12.072

Sharrard, W. J. W. (1964). The Segmental Innervation of the Lower Limb Muscles in Man. *Annals of The Royal College of Surgeons of England*, *35*(2), 106–122. Retrieved February 22, 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2311748/

Shevtsova, N. A., Hamade, K., Chakrabarty, S., Markin, S. N., Prilutsky, B. I., & Rybak, I. A. (2016). Modeling the Organization of Spinal Cord Neural Circuits Controlling Two-Joint Muscles. In B. I. Prilutsky & D. H. Edwards (Eds.), *Neuromechanical Modeling of Posture and Locomotion* (pp. 121–162). Springer New York. https://doi.org/10.1007/978-1-4939-3267-2_5

Shippen, J., & May, B. (2012). A kinematic approach to calculating ground reaction forces in dance. *Journal of Dance Medicine & Science*, *16*(1), 39–43.

Sompolinsky, H., Crisanti, A., & Sommers, H. J. (1988). Chaos in Random Neural Networks. *Physical Review Letters*, *61*(3), 259–262. https://doi.org/10.1103/PhysRevLett.61.259

Souissi, H., Zory, R., Bredin, J., & Gerus, P. (2017). Comparison of methodologies to assess muscle co-contraction during gait. *Journal of Biomechanics*, *57*, 141–145. https://doi.org/10.1016/j.jbiomech.2017.03.029

Steele, K. M., Rozumalski, A., & Schwartz, M. H. (2015). Muscle synergies and complexity of neuromuscular control during gait in cerebral palsy. *Developmental Medicine & Child Neurology*, *57*(12), 1176–1182. https://doi.org/10.1111/dmcn.12826

Susman, L., Mastrogiuseppe, F., Brenner, N., & Barak, O. (2021). Quality of internal representation shapes learning performance in feedback neural networks. *Physical Review Research*, *3*(1), 013176. https://doi.org/10.1103/PhysRevResearch.3.013176

Sussillo, D., & Abbott, L. (2009). Generating Coherent Patterns of Activity from Chaotic Neural Networks. *Neuron*, *63*(4), 544–557. https://doi.org/10.1016/j.neuron.2009.07.018

Sydserff, S. G., Cross, A. J., & Green, A. R. (1995). The neuroprotective effect of chlormethiazole on ischaemic neuronal damage following permanent middle cerebral artery ischaemia in the rat. *Neurodegeneration*, *4*(3), 323–328. https://doi.org/10.1016/1055-8330(95)90022-5

Talpalar, A. E., Bouvier, J., Borgius, L., Fortin, G., Pierani, A., & Kiehn, O. (2013). Dual-mode operation of neuronal networks involved in left–right alternation. *Nature*, *500*(7460), 85–88. https://doi.org/10.1038/nature12286

Tang, L., Li, F., Cao, S., Zhang, X., Wu, D., & Chen, X. (2015). Muscle synergy analysis in children with cerebral palsy. *Journal of Neural Engineering*, *12*(4), 046017. https://doi.org/10.1088/1741-2560/12/4/046017

Thaller, S., & Wagner, H. (2004). The relation between Hill's equation and individual muscle properties. *Journal of Theoretical Biology*, *231*(3). https://doi.org/10.1016/j.jtbi.2004.06.027

Tisch, S., Limousin, P., Rothwell, J. C., Asselman, P., Zrinzo, L., Jahanshahi, M., Bhatia, K. P., & Hariz, M. I. (2006). Changes in forearm reciprocal inhibition following pallidal stimulation for dystonia. *Neurology*, *66*(7), 1091–1093. https://doi.org/10.1212/01.wnl.0000204649.36458.8f

Todd, A. J., & Sullivan, A. C. (1990). Light microscope study of the coexistence of GABA-like and glycine-like immunoreactivities in the spinal cord of the rat. *The Journal of Comparative Neurology*, *296*(3), 496–505. https://doi.org/10.1002/cne.902960312

Verstraeten, D., Schrauwen, B., D'Haene, M., & Stroobandt, D. (2007). An experimental unification of reservoir computing methods. *Neural Networks*, *20*(3), 391–403. https://doi.org/10.1016/j.neunet.2007.04.003

Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C., & Gerstner, W. (2011). Inhibitory Plasticity Balances Excitation and Inhibition in Sensory Pathways and Memory Networks. *Science*, *334*(6062), 1569–1573. https://doi.org/10.1126/science.1211095

Vogels, T. P., Rajan, K., & Abbott, L. (2005). Neural Network Dynamics. *Annual Review of Neuroscience*, *28*(1), 357–376. https://doi.org/10.1146/annurev.neuro.28.061604.135637

Vreeswijk, C. van, & Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science (New York, N.Y.)*, *274*(5293), 1724–1726. https://doi.org/10.1126/science.274.5293.1724

Wagner, H., Boström, K. J., Lussanet, M. H. de, Graaf, M. L. de, Puta, C., & Mochizuki, L. (2022). Optimization Reduces Knee-Joint Forces During Walking and Squatting: Validating the Inverse Dynamics Approach for Full Body Movements on Instrumented Knee Prostheses. *Motor Control*, 1–18. https://doi.org/10.1123/mc.2021-0110

Walther, S., & Strik, W. (2012). Motor Symptoms and Schizophrenia. *Neuropsychobiology*, *66*(2), 77–92. https://doi.org/10.1159/000339456

Winter, D. A. (2009). *Biomechanics and Motor Control of Human Movement*. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470549148

Wonders, C. P., & Anderson, S. A. (2006). The origin and specification of cortical interneurons. *Nature Reviews Neuroscience*, *7*(9), 687–696. https://doi.org/10.1038/nrn1954

Wyffels, F., & Schrauwen, B. (2009). Design of a Central Pattern Generator Using Reservoir Computing for Learning Human Motion. *2009 Advanced Technologies for Enhanced Quality of Life*, 118–122. https://doi.org/10.1109/AT-EQUAL.2009.32

Yaksh, T. L. (1989). Behavioral and autonomic correlates of the tactile evoked allodynia produced by spinal glycine inhibition: Effects of modulatory receptor systems and excitatory amino acid antagonists. *Pain*, *37*(1), 111–123. https://doi.org/10.1016/0304-3959(89)90160-7

Yizhar, O., Fenno, L. E., Prigge, M., Schneider, F., Davidson, T. J., O'Shea, D. J., Sohal, V. S., Goshen, I., Finkelstein, J., Paz, J. T., Stehfest, K., Fudim, R., Ramakrishnan, C., Huguenard, J. R., Hegemann, P., & Deisseroth, K. (2011). Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature*, *477*(7363), 171–178. https://doi.org/10.1038/nature10360

Zampella, C. J., Wang, L. A. L., Haley, M., Hutchinson, A. G., & Marchena, A. de. (2021). Motor Skill Differences in Autism Spectrum Disorder: A Clinically Focused Review. *Current Psychiatry Reports*, *23*(10), 64. https://doi.org/10.1007/s11920-021-01280-6

Zhang, J., Lanuza, G. M., Britz, O., Wang, Z., Siembab, V. C., Zhang, Y., Velasquez, T., Alvarez, F. J., Frank, E., & Goulding, M. (2014). V1 and v2b interneurons secure the alternating flexor-extensor motor

activity mice require for limbed locomotion. *Neuron*, *82*(1), 138–150. https://doi.org/10.1016/j.neuron.2014.02.013

Appendix A Mean Field Theory

We studied the networks analytically through the application of Mean-Field Theory (MFT). With this technique, a description of the network dynamics can be derived self-consistently by averaging over the random parts of the network connectivity (Mastrogiuseppe & Ostojic, 2017, 2018; Rajan et al., 2010; Sompolinsky et al., 1988). Following this approach, we decompose the input received by each neuron into a static component and a component that depends on the random instantiation of the network connectivity. We then average across different realisations of the network connectivity to determine the mean and variance of this random component. This allows us to self-consistently determine the mean and variance of the neuron firing rates.

A.1 Network Dynamics

We consider a recurrent network of N_E excitatory and N_I inhibitory neurons. Each neuron i is characterized by its activation $x_i(t)$ and its activity $\phi(x_i(t))$, corresponding to its firing rate. The dynamics follow:

$$\tau \frac{d}{dt} x_i = -x_i + I_i(t) \tag{10}$$

where $I_i(t)$ is the input received by any neuron i. This input is the sum of three terms: the external input $I_i^{In}(t)$, the excitatory recurrent input $I_i^E(t)$, and the inhibitory recurrent input $I_i^I(t)$. Note that here, we consider the network dynamics in the absence of output feedback. A derivation of the Mean Field Theory in the presence of output feedback can be found in Mastrogiuseppe and Ostojic (2018).

The external drive $U(t)$ is fed to the network via input weights J_i^{In}, which are drawn from a random distribution with mean m_{In} and variance v_{In}, such that each neuron receives external input $I_i^{In}(t) = J_i^{In} U(t)$. The outgoing weights from the excitatory (resp. inhibitory) neurons J_{ij}^E (resp. J_{ij}^I) are drawn from a random distribution with mean m_E (resp. m_I) and variance v_E (resp. v_I), such that each neuron receives the following recurrent input from each of the two populations $P \in \{E, I\}$:

$$I_i^P(t) = \sum_{j=1}^{N_P} J_{ij}^P \phi\left(x_j^P(t)\right) \tag{11}$$

The mean and variance of the recurrent input affect the network dynamics differently. To account for this, we write the recurrent weights as a sum of their mean m_P and a random component $\mathscr{J}_{ij}^P$ with mean 0 and variance v_P: $J_{ij}^P = m_P + \mathscr{J}_{ij}^P$ We further introduce the shorthand $\varphi_j^P = \phi\left(x_j^P(t)\right)$, as well as the population-averaged firing rate: $\varphi^P(t) = \frac{1}{N_P}\sum_{j=1}^{N_P} \varphi_j^P(t)$. The recurrent input received by neuron i from population P is then:

$$\begin{aligned} I_i^P(t) &= \sum_{j=1}^{N_P} (m_P + \mathscr{J}_{ij}^P)\varphi_j^P(t) = m_P \sum_{j=1}^{N_P} \varphi_j^P(t) + \sum_{j=1}^{N_P} \mathscr{J}_{ij}^P \varphi_j^P(t) \\ &= m_P N_P \varphi^P(t) + \eta_i^P(t) \end{aligned} \tag{12}$$

where

$$\eta_i^P(t) = \sum_{j=1}^{N_P} \mathscr{J}_{ij}^P \varphi_j^P(t) \tag{13}$$

The total input received by each neuron i is then:

$$I_i(t) = m_E N_E \varphi^E(t) + m_I N_I \varphi^I(t) + \eta_i^E(t) + \eta_i^I(t) + \eta_i^{In}(t) \tag{14}$$

where $\eta_i^{In} = J_i^{In} U(t)$.

A.2 Statistics of the input to individual neurons

To determine the influence of the network connectivity on the statistics of the input received by each neuron, we model the input I_i received by each neuron as a Gaussian process of mean μ and variance Δ. We determine this mean and variance by averaging over the random connectivity, which we write as $[\cdot]$. We note that, when averaging over the random connectivity, the statistics of the inputs received by all neurons are identical. As a consequence, the first and second-order moments of the population-averaged firing rates of the excitatory population match those of the inhibitory population, so we can write $\left[\varphi(t)\right] = \left[\varphi_j^E(t)\right] = \left[\varphi_j^I(t)\right]$ and $\left[\varphi^2(t)\right] = \left[(\varphi_j^E t))^2\right] = \left[(\varphi_j^I t))^2\right]$.

A.2.1 Mean μ of the input to individual neurons

The mean input $\mu(t) = [I_i(t)]$ is given by:

$$\mu(t) = (m_E N_E + m_I N_I)\left[\varphi(t)\right] + \left[\eta_i^E(t)\right] + \left[\eta_i^I(t)\right] + \left[\eta_i^{In}(t)\right] \tag{15}$$

The input weights J_i^{In} are independent of $U(t)$ and have zero mean, therefore:

$$\left[\eta_i^{In}(t)\right] = \left[J_i^{In}U(t)\right] = \left[J_i^{In}\right]U(t) = 0 \tag{16}$$

In the large network limit, the neuron firing rates are assumed to be independent of their outgoing weights, therefore $\left[\mathscr{J}_{ij}^P\varphi_j^P(t)\right] = \left[\mathscr{J}_{ij}^P\right]\left[\varphi_j^P(t)\right]$. Since $\left[\mathscr{J}_{ij}^P\right] = 0$, we have:

$$\left[\eta_i^P(t)\right] = \left[\sum_{j=1}^{N_P}\mathscr{J}_{ij}^P\varphi_j^P(t)\right] = \sum_{j=1}^{N_P}\left[\mathscr{J}_{ij}^P\right]\left[\varphi_j^P(t)\right] = 0 \tag{17}$$

Therefore:

$$\mu(t) = (m_E N_E + m_I N_I)\left[\varphi(t)\right] \tag{18}$$

A.2.2 Variance Δ of the input to individual neurons

The variance Δ of the input to each neuron equals $\left[I_i(t)^2\right] - [I_i(t)]^2$. Since the terms in the right-hand side of Equation 14 are all uncorrelated, the variance of I_i equals the sum of the variances of all terms. The variance of each of the first two terms equals $(m_P N_P)^2([\phi_P^2] - [\phi_P]^2)$ We assume that, in the large network limit, the population averages are identical to averages over the random connectivity: $\varphi^E(t) = \varphi^I(t) = \left[\varphi(t)\right]$. With this assumption, these two variances will thus equal zero, and the recurrent input variance is equal to the variance of $\eta_i^E(t) + \eta_i^I(t) + \eta_i^{In}$. We decompose this into the following terms:

$$\begin{aligned}\Delta(t) =& var\left(\eta_i^E(t)\right) + var\left(\eta_i^I(t)\right) + var\left(\eta_i^{In}(t)\right)\\ &+ 2cov\left(\eta_i^E(t),\eta_i^{In}(t)\right) + 2cov\left(\eta_i^I(t),\eta_i^{In}(t)\right) + 2cov\left(\eta_i^E(t),\eta_i^I(t)\right)\end{aligned} \tag{19}$$

The variance of the external input is equal to:

$$var\left(\eta_i^{In}(t)\right) = \left[(J_i^{In})^2U(t)^2\right] - \left[J_i^{In}U(t)\right]^2 = \left[(J_i^{In})^2\right]U(t)^2 = v_{In}U(t)^2 \tag{20}$$

The external input weights J_i^{In} are assumed to be independent of the recurrent inputs, therefore:

$$cov\left(\eta_i^P(t),\eta_i^{In}(t)\right) = 0 \tag{21}$$

The variance of the recurrent inputs is equal to

$$\begin{aligned}var\left(\eta_i^P(t)\right) &= \left[\left(\sum_{j=1}^{N_P}\mathscr{J}_{ij}^P\varphi_j^P(t)\right)^2\right] - \left[\sum_{j=1}^{N_P}\mathscr{J}_{ij}^P\varphi_j^P(t)\right]^2\\ &= \left[\sum_{j=1}^{N_P}\sum_{k=1}^{N_P}\mathscr{J}_{ij}^P\mathscr{J}_{ik}^P\varphi_j^P(t)\varphi_k^P(t)\right] - \left[\sum_{j=1}^{N_P}\mathscr{J}_{ij}^P\varphi_j^P(t)\right]^2\\ &= \sum_{j=1}^{N_P}\sum_{k=1}^{N_P}\left[\mathscr{J}_{ij}^P\mathscr{J}_{ik}^P\right]\left[\varphi_j^P(t)\varphi_k^P(t)\right] - \sum_{j=1}^{N_P}\sum_{k=1}^{N_P}\left[\mathscr{J}_{ij}^P\right]\left[\varphi_j^P(t)\right]\left[\mathscr{J}_{ik}^P\right]\left[\varphi_k^P(t)\right]\end{aligned} \tag{22}$$

Since $\left[\mathscr{J}^P_{ij}\right] = 0$, the last term is zero. Moreover, since the recurrent weights are independent:

$$\left[\mathscr{J}^P_{ij}\mathscr{J}^P_{ik}\right] = \begin{cases} 0, & \text{if } j \neq k \\ v_P, & \text{if } j = k \end{cases} \tag{23}$$

Therefore:

$$var\left(\eta^P_i(t)\right) = v_P \sum_{j=1}^{N_P} \left[(\varphi^P_j(t))^2\right] = v_P N_P \left[\varphi^2(t)\right] \tag{24}$$

The covariance between the excitatory and inhibitory recurrent inputs is equal to:

$$cov\left(\eta^E_i(t), \eta^I_i(t)\right) = \left[\sum_{j=1}^{N_E} \mathscr{J}^E_{ij}\varphi^E_j(t) \sum_{k=1}^{N_I} \mathscr{J}^I_{ik}\varphi^I_k(t)\right] - \left[\sum_{j=1}^{N_E} \mathscr{J}^E_{ij}\varphi^E_j(t)\right]\left[\sum_{j=1}^{N_I} \mathscr{J}^I_{ij}\varphi^I_j(t)\right] \tag{25}$$

$$= \sum_{j=1}^{N_E}\sum_{k=1}^{N_E} \left[\mathscr{J}^E_{ij}\mathscr{J}^I_{ik}\right]\left[\varphi^E_j(t)\varphi^I_k(t)\right] - \sum_{j=1}^{N_E}\sum_{k=1}^{N_I} \left[\mathscr{J}^E_{ij}\right]\left[\varphi^E_j(t)\right]\left[\mathscr{J}^I_{ik}\right]\left[\varphi^I_k(t)\right] \tag{26}$$

The recurrent weights from the two populations are independent, therefore $\left[\mathscr{J}^E_{ij}\mathscr{J}^I_{ik}\right] = 0$ and:

$$cov\left(\eta^E_i(t), \eta^I_i(t)\right) = 0 \tag{27}$$

The input variance Δ is thus equal to:

$$\Delta(t) = (v_E N_E + v_I N_I)\left[\varphi^2(t)\right] + v_{In}U(t)^2 \tag{28}$$

A.3 Firing rate statistics

To obtain self-consistent equations for $\mu(t)$ and $\Delta(t)$, we need to express $\left[\varphi(t)\right]$ and $\left[\varphi^2(t)\right]$ as functions of $\mu(t)$ and $\Delta(t)$. For this, we assume that the external input $U(t)$ varies much more slowly than the time constant τ, and that the system is in a quasi-stationary state. We will show below (see Section A.6) that this stationary state is stable. For stationary solutions, the activation $x_i(t)$ has the same distribution as the input $I_i(t)$. We explicitly model the input x as a Gaussian process whose mean $\mu(t)$ and variance $\Delta(t)$ satisfy Equations 18 and 28: $x_i = \mu + \sqrt{\Delta}z$, where z is a Gaussian process with zero mean and unit variance, whose probability distribution is thus given by $p(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$. By modelling the input as a Gaussian, we assume that the input has zero skewness. Simulation results show that the skewness is indeed small for all investigated imbalances (Figure S.1).

We can then obtain the firing rate statistics by averaging over z:

$$\left[\varphi(t)\right] = \left[\phi(x^P_j(t))\right] = \int Dz \cdot \phi\left(\mu(t) + \sqrt{\Delta(t)}z\right) \tag{29}$$

$$\left[\varphi^2(t)\right] = \left[\phi(x^P_j(t))^2\right] = \int Dz \cdot \phi^2\left(\mu(t) + \sqrt{\Delta(t)}z\right) \tag{30}$$

where $\int Dz = \int_{-\infty}^{+\infty} p(z)\, dz = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\, dz$.

A.4 Solving the equations

Substituting equations Equations 18 and 28 into Equations 29 and 30, we obtain the following equations for the μ- and Δ- nullclines, $\mathscr{N}_\mu$ and $\mathscr{N}_\Delta$:

$$\mathscr{N}_\mu(\mu(t), \Delta(t)) = -\mu(t) + (m_E N_E + m_I N_I)\int Dz \cdot \phi\left(\mu(t) + \sqrt{\Delta(t)}z\right) = 0 \tag{31}$$

$$\mathscr{N}_\Delta(\mu(t), \Delta(t), U(t)) = -\Delta(t) + (v_E N_E + v_I N_I)\int Dz \cdot \phi^2\left(\mu(t) + \sqrt{\Delta(t)}z\right) + v_{In}U(t)^2 = 0 \tag{32}$$

We solved the system of equations numerically using a built-in solver for sets of non-linear equations, `fsolve`, in Matlab R2023a (MathWorks, Inc., Natick, Massachusetts, United States). Graphically, this corresponds to finding the intersection between the μ- and Δ- nullclines.

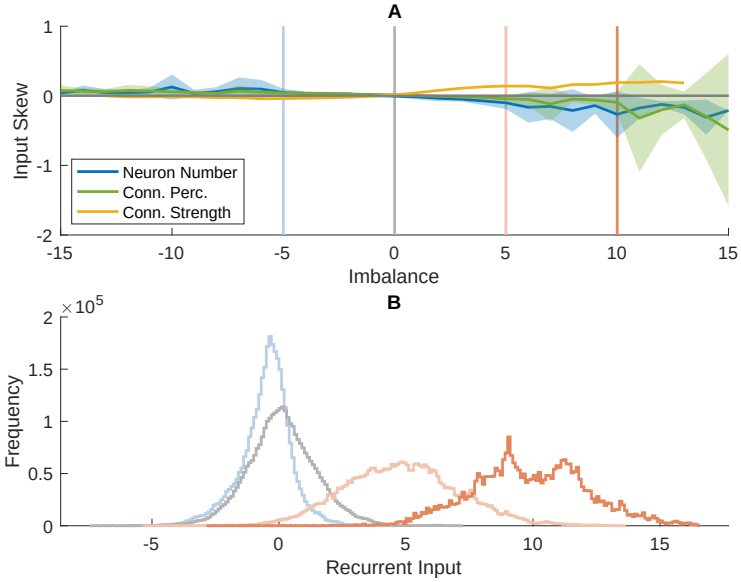

Figure S.1 – Skew of the time-averaged recurrent input distributions. *(A) shows the skew as a function of the imbalance A when varying the neuron number (blue), connection percentage (green) and connection strength (yellow), as well as the corresponding MFT predictions for $v_E N_E + v_I N_I = 2 \cdot 1.5^2$. For the neuron number and connection percentage, the average and standard deviations over all levels of N_{tot} and p_{tot} are shown. For the connection strength, only $v_E N_E + v_I N_I = 2 \cdot 1.5^2$ is shown to match MFT predictions. The coloured vertical lines indicate the imbalance levels of the four example networks shown in panel (B). (B) shows the histograms of the recurrent input in these four networks*

A.4.1 Nullclines

The μ-nullcline depends on the imbalance $m_E N_E + m_I N_I$. It is illustrated in Figure S.2 for different values of imbalance ranging from -10 to 10. For large, positive imbalance (red curves in Figure S.2 the μ-nullcline approaches a vertical line at $\mu = m_E N_E + m_I N_I$. In contrast, for large, negative imbalance (blue curves in Figure S.2), μ decreases only moderately and does not scale with $m_E N_E + m_I N_I$.

The Δ-nullcline depends on the recurrent weight variance $v_E N_E + v_I N_I$ and the external input variance $v_{In} U^2(t)$. It is illustrated in Figure S.2 for a recurrent weight variance of $v_E N_E + v_I N_I = 2 \cdot 1.5^2$ and $v_{In} U^2(t) = 1/2$ (as used in simulations throughout the paper, unless otherwise specified). Δ decreases with decreasing μ until a minimum value $\Delta_{min} = v_{In} U(t)^2$. For increasing μ, it increases up to a plateau value of $\Delta_{max} = v_{In} U^2(t) + v_E N_E + v_I N_I$.

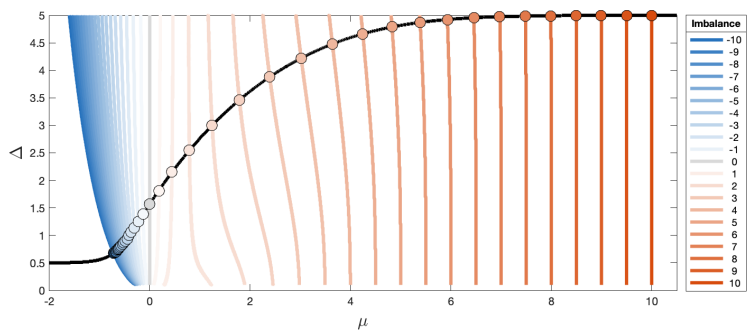

Figure S.2 – Graphical representation of the μ and Δ nullclines. *The μ nullcline (coloured) depends on the anatomical imbalance A, shown here for values ranging from -10 to 10. The Δ nullcline (black) depends on $v_E N_E + v_I N_I$, shown here in black for a single value $2 \cdot 1.5^2$, as used in the simulations. The solutions for each imbalance level, indicated by the coloured dots, are found at the intersection of the two nullclines.*

The solutions for μ (Figure S.3A) and Δ (Figure S.3B) are given by the intersection of the two nullclines. The obtained values are subsequently used to calculate the firing rate mean (Figure S.3C) and variance (Figure S.3D) using Equations 29 and 30.

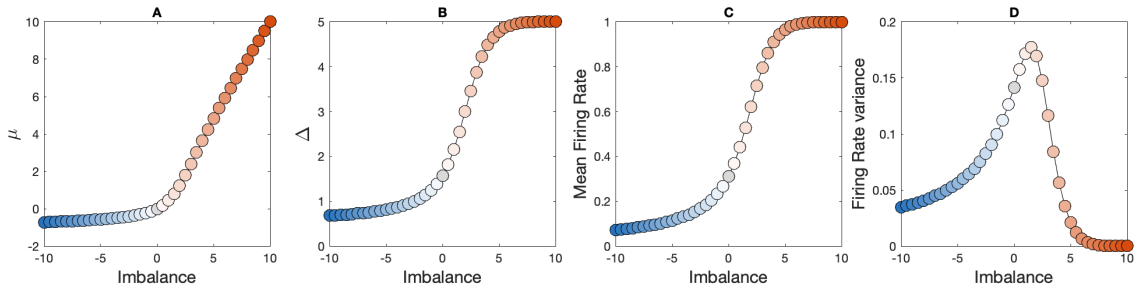

Figure S.3 – Mean-field theory solutions found via the nullcline method *for* $v_E N_E + v_I N_I = 2 \cdot 1.5^2$*, as used in the simulations. (A) shows the mean recurrent input, (B) the recurrent input variance, (C) the mean firing rate, and (D) the firing rate variance as a function of the imbalance A.*

A.5 Mean and variance of the weight distributions

In our simulations, the outgoing connection weights from population P have a probability p_P of being non-zero. The non-zero weights are drawn from a Gaussian distribution with mean 0 and variance $V_P^0 = \frac{g_P^2}{N_P p_P (1-\frac{2}{\pi} p_P)}$. This scaling is chosen such that the variance of the weight distribution scales as $v_P = g_P^2 / N_p$ (see below). For excitatory weights, the absolute value of this weight is taken; whereas for inhibitory weights, the negative of the absolute value is taken. The mean M_P and variance V_P of the non-zero weights therefore correspond to those of a half-Gaussian distribution:

$$M_P = \frac{g_P}{\sqrt{N_P \left(1 - \frac{2}{\pi} p_P\right)}} \cdot \sqrt{\frac{2}{\pi}} \tag{33}$$

$$V_P = \frac{g_P^2}{N_P p_P \left(1 - \frac{2}{\pi} p_P\right)} \cdot \left(1 - \frac{2}{\pi}\right) \tag{34}$$

The mean recurrent weight of the full connection matrix m_P is therefore equal to

$$m_P = M_P \cdot p_P = g_P \sqrt{\frac{p_P}{N_P \left(1 - \frac{2}{\pi} p_P\right)}} \sqrt{\frac{2}{\pi}} \tag{35}$$

To obtain the variance of the full matrix, we can do the same for the mean of the squared connection weight, which equals $v_P + m_P^2$, to get:

$$v_P + m_P^2 = p_P \left(V_P + M_P^2\right) \tag{36}$$

Therefore:

$$v_P = p_P \left(V_P + M_P^2\right) - m_P^2 = \frac{g_P^2}{N_P} \tag{37}$$

We then obtain the mean and variance of the sum of the recurrent weights for our networks:

$$m_E N_E + m_I N_I = \sqrt{\frac{2}{\pi}} \left(g_E \sqrt{\frac{p_E N_E}{\left(1 - \frac{2}{\pi} p_E\right)}} + g_I \sqrt{\frac{p_I N_I}{\left(1 - \frac{2}{\pi} p_I\right)}} \right) \tag{38}$$

$$v_E N_E + v_I N_I = (g_E^2 + g_I^2) \tag{39}$$

Here, $m_E N_E + m_I N_I$ corresponds to the anatomical imbalance A, and $v_E N_E + v_I N_I$ is equal to the square of the total connection strength g_{tot}, as defined in the main text.

The input weight matrix was taken from a Gaussian distribution with mean zero and variance $v_{in} = \frac{1}{3}$. The external input $U(t)$ is a sinusoid ranging from 0 to 2. As we analysed a quasi-stationary state, the MFT was performed using a single value for the external input, corresponding to the temporal mean of U: $\int U(t)^2 = 1.5$. Therefore, the term $v_{in} U(t)^2$ equalled $\frac{1}{2}$ in all included MFT analyses.

A.6 Stability Analysis

A.6.1 Local perturbation analysis

To determine whether the stationary solutions derived above are stable, we linearized the system dynamics around the steady state. To simplify notations, we introduce the $(N_E + N_I)$ by $(N_E + N_I)$ matrix $\boldsymbol{J}$ whose first N_E columns correspond to the excitatory weights J^E_{ij} and whose last N_I columns correspond to the inhibitory weights J^I_{ij}. The system dynamics are then:

$$\tau \frac{d}{dt} x_i = -x_i + \sum_j J_{ij} \phi\left(x_j\right) + J_i^{In} U(t) \tag{40}$$

The steady-state x_i^0 then satisfies:

$$x_i^0 = \sum_j J_{ij} \phi\left(x_j^0\right) + J_i^{In} U(t) \tag{41}$$

To linearize the dynamics, we consider small perturbations δx_i around the steady state x_i^0 and write $x_i(t) = x_i^0 + \delta x_i(t)$, where x_i^0 is constant. From this, it follows that $\tau \frac{d}{dt} \delta x_i = \tau \frac{d}{dt} x_i$. We can then approximate the derivative of δx_i using a first-order Taylor series expansion to get:

$$\begin{aligned}
\tau \frac{d}{dt} \delta x_i &= -x_i + \sum_j J_{ij} \phi\left(x_j\right) + J_i^{In} U(t) \\
&= -x_i^0 - \delta x_i + \sum_j J_{ij} \phi\left(x_j^0 + \delta x_j\right) + J_i^{In} U(t) \\
&\approx -x_i^0 - \delta x_i + \sum_j J_{ij} \left(\phi\left(x_j^0\right) + \delta x_j \phi'\left(x_j^0\right)\right) + J_i^{In} U(t) \\
&= -\delta x_i + \sum_j J_{ij} \delta x_j \phi'\left(x_j^0\right)
\end{aligned} \tag{42}$$

We can write this in matrix notation, using the perturbation vector $\boldsymbol{\delta x} = \begin{pmatrix} \delta x_0 \\ \vdots \\ \delta x_N \end{pmatrix}$, the $(N_E + N_I)$-dimensional identity matrix $\boldsymbol{I}$ and $S_{ij} = J_{ij} \phi'\left(x_j^0\right)$, resulting in:

$$\tau \frac{d}{dt} \boldsymbol{\delta x} = (-\boldsymbol{I} + \boldsymbol{S}) \boldsymbol{\delta x} \tag{43}$$

This system is stable if and only if the eigenvalues of $(-\boldsymbol{I} + \boldsymbol{S})$ all have a negative real part or, equivalently, if the eigenvalues of $\boldsymbol{S}$ all have real parts smaller than 1. The eigenvalue spectrum of $\boldsymbol{S}$ consists of $N-1$ eigenvalues distributed within a disk in the complex plane, and a single outlier eigenvalue (Mastrogiuseppe & Ostojic, 2018; Rajan & Abbott, 2006), as shown in Figure S.4A. The disk corresponds to the network's response to local perturbations to individual neurons, while the outlier eigenvalue corresponds to the network's response to a global perturbation of the population-averaged mean μ or variance Δ of the input.

The radius r of the disk can be calculated by summing the variances of the columns of $\boldsymbol{S}$:

$$r^2 = (\upsilon_E N_E + \upsilon_I N_I) \left[\varphi'^{\,2}\right] \tag{44}$$

When this radius is larger than unity, the stationary state loses stability and the network has chaotic, fluctuating activity (Aljadeff et al., 2015; Harish & Hansel, 2015; Mastrogiuseppe & Ostojic, 2018; Rajan & Abbott, 2006). We find that this occurs when $g_{tot} = \sqrt{\upsilon_E N_E + \upsilon_I N_I} \geq 2.5$ (see Figure S.4B), and exclude these networks from the results.

A.6.2 Global perturbation analysis

It is not known how to calculate the outlier eigenvalue directly, but following the approach of Kadmon and Sompolinsky (2015) and Mastrogiuseppe and Ostojic (2018), we derive it by assessing the response of the population-averaged mean μ and variance Δ to a global perturbation. First, we derive a dynamical system governing the temporal evolution of μ and Δ. Then, we linearize this system around the steady state μ_0, Δ_0 obtained by solving equations Equation 31 and Equation 32.

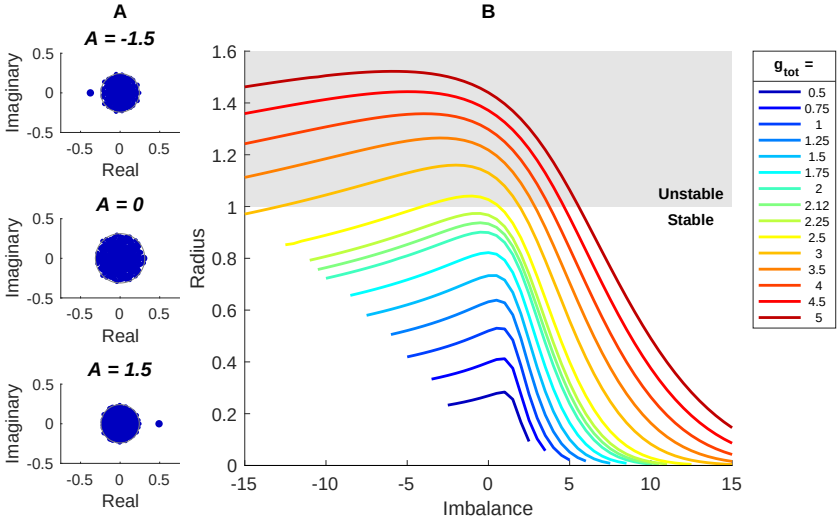

Figure S.4 – Results of the local stability analysis. *(A) shows three example eigenvalue spectra from simulated networks with total connection strength $g_{tot} = 0.5$ and imbalance $A = -1.5, 0$ and 1.5. (B) shows the MFT-derived radius of the eigenvalue spectrum of the stability matrix as a function of the imbalance A for various levels of the total connection strength ($g_{tot} = \sqrt{v_E N_E + v_I N_I}$). The grey-shaded area indicates radii for which the network is unstable.*

Time-evolution of μ and Δ We extend the previous mean-field theory analysis to consider $\mu(t) = [I_i(t)]$, i.e. the average (across different instantiations of the random connectivity) of the input received by neuron i after a global perturbation occurring at time $t = 0$. We decompose the neural activation response $x_i(t)$ as a sum of this population-averaged response $\mu(t)$ and individual neuron variation $\xi_i(t)$:

$$x_i(t) = \mu(t) + \xi_i(t) \tag{45}$$

The dynamics of x_i can therefore be written as:

$$\begin{aligned}\tau \frac{d}{dt}\mu(t) + \tau \frac{d}{dt}\xi_i(t) &= \tau \frac{d}{dt} x_i(t) \\ &= -x_i + \sum_{j=1}^{N_E}\left(\mathscr{J}_{ij}^E + m_E\right)\phi\left(x_j^E\right) + \sum_{j=1}^{N_I}\left(\mathscr{J}_{ij}^I + m_I\right)\phi\left(x_j^I\right) + J_i^{In}U \\ &= -\mu(t) - \xi_i + m_E N_E \varphi^E(t) + \sum_j \mathscr{J}_{ij}^E \phi\left(x_j^E\right) + m_I N_I \varphi^I(t) + \sum_j \mathscr{J}_{ij}^I \phi\left(x_j^I\right) + J_i^{In}U\end{aligned} \tag{46}$$

We then average over the random connectivity, substituting $[\xi_i(t)] = 0$, $\left[\frac{d}{dt}\xi_i\right] = 0$, $[x_i(t)] = \mu(t)$ and $\left[\mathscr{J}_{ij}^P\right] = 0$ in the above equation, to obtain the following differential equation for μ:

$$\tau \frac{d}{dt}\mu(t) = -\mu(t) + (m_E N_E + m_I N_I)\left[\phi(t)\right] \tag{47}$$

To obtain a differential equation of $\Delta(t) = \left[\xi_i(t)^2\right]$, we subtract equation Equation 47 from Equation 46. As we did previously, we consider that, in the large network limit, the population averages are identical to averages over the random connectivity: $\varphi^E(t) = \varphi^I(t) = \left[\varphi(t)\right]$. We then obtain:

$$\tau \frac{d}{dt}\xi_i(t) = -\xi_i(t) + \sum_{j=1}^{N_E}\mathscr{J}_{ij}^E \phi\left(x_j\right) + \sum_{j=1}^{N_I}\mathscr{J}_{ij}^I \phi\left(x_j\right) + J_i^{In}U(t) \tag{48}$$

We note that $\frac{d}{dt}\Delta = \frac{d}{dt}\left[\xi_i(t)^2\right] = \left[2\xi_i \frac{d}{dt}\xi_i\right]$ and use this to express $\tau \frac{d}{dt}\xi_i(t) + \xi_i(t)$ in the formula

above in terms of $\Delta(t)$ and $\frac{d}{dt}\Delta(t)$:

$$\begin{aligned}\left[\left(\tau\frac{d}{dt}\xi_i(t)+\xi_i(t)\right)^2\right] &= \left[\left(\tau\frac{d}{dt}\xi_i(t)\right)^2\right]+\left[\xi_i(t)^2\right]+\left[2\tau\xi_i(t)\frac{d}{dt}\xi_i(t)\right] \\ &= \left[\left(\tau\frac{d}{dt}\xi_i(t)\right)^2\right]+\Delta(t)+\tau\frac{d}{dt}\Delta(t)\end{aligned} \tag{49}$$

As temporal derivatives vanish when evaluated at fixed points, we can neglect the $\left[\left(\tau\frac{d}{dt}\xi_i(t)\right)^2\right]$ term. Substituting equation Equation 48 into Equation 49, we obtain:

$$\Delta(t)+\tau\frac{d}{dt}\Delta=\left[\left(\sum_{j=1}^{N_E}\mathscr{J}_{ij}^{E}\phi\left(x_j\right)+\sum_{j=1}^{N_I}\mathscr{J}_{ij}^{I}\phi\left(x_j\right)+J_i^{In}U(t)\right)^2\right] \tag{50}$$

The right-hand term can be evaluated as in Section A.2, resulting in:

$$\tau\frac{d}{dt}\Delta=-\Delta(t)+(v_EN_E+v_IN_I)\left[\phi^2(t)\right]+v_{In}U(t)^2 \tag{51}$$

Linearization We obtain the dynamics of the system's response to global perturbations by substituting the equations for the μ- and Δ-nullclines (Equations 31 and 32) into equations Equations 47 and 51:

$$\tau\frac{d}{dt}\mu(t)=-\mu(t)+(m_EN_E+m_IN_I)\left[\phi(t)\right]=\mathscr{N}_\mu(\mu(t),\Delta(t)) \tag{52}$$

$$\tau\frac{d}{dt}\Delta=-\Delta(t)+(v_EN_E+v_IN_I)\left[\phi^2(t)\right]+v_{In}U(t)^2=\mathscr{N}_\Delta(\mu(t),\Delta(t),U(t)) \tag{53}$$

To determine whether the system is stable, we linearize the system dynamics around the steady-state μ_0,Δ_0: $\mu(t)=\mu_0+\delta\mu(t)$,$\Delta(t)=\Delta_0+\delta\Delta(t)$.

The response of the mean input follows:

$$\begin{aligned}\tau\frac{d}{dt}\mu(t)=\mathscr{N}_\mu(\mu(t),\Delta(t)) &= \mathscr{N}_\mu(\mu_0,\Delta_0)+\delta\mu(t)\frac{\partial\mathscr{N}_\mu}{\partial\mu}+\delta\Delta(t)\frac{\partial\mathscr{N}_\mu}{\partial\Delta} \\ &= \delta\mu(t)\frac{\partial\mathscr{N}_\mu}{\partial\mu}+\delta\Delta(t)\frac{\partial\mathscr{N}_\mu}{\partial\Delta}\end{aligned} \tag{54}$$

$$\frac{\partial\mathscr{N}_\mu}{\partial\mu} = -1+(m_EN_E+m_IN_I)\frac{\partial}{\partial\mu}\left[\phi(t)\right] \tag{55}$$

$$\frac{\partial}{\partial\mu}\left[\phi(t)\right] = \frac{\partial}{\partial\mu}\int Dz\phi(\mu(t)+\sqrt{\Delta(t)}z)=\int Dz\phi'(\mu_0+\sqrt{\Delta_0}z)=\left[\phi'\right]_0 \tag{56}$$

where $[.]_0$ denotes the average evaluated at the steady-state.

$$\frac{\partial\mathscr{N}_\mu}{\partial\Delta} = (m_EN_E+m_IN_I)\frac{\partial}{\partial\Delta}\left[\phi(t)\right] \tag{57}$$

$$\frac{\partial}{\partial\Delta}\left[\phi(t)\right] = \frac{\partial}{\partial\Delta}\int Dz\phi(\mu(t)+\sqrt{\Delta(t)}z)=\int Dzz\frac{1}{2\sqrt{\Delta_0}}\phi'(\mu_0+\sqrt{\Delta_0}z) \tag{58}$$

We use $\int Dz\, zf(z)=\int Dz\, f'(z)$ (derived below in Section A.6.3). Substituting this into the equation above, we get:

$$\frac{\partial}{\partial\Delta}\left[\phi(t)\right]=\frac{1}{2\sqrt{\Delta_0}}\int Dz\sqrt{\Delta_0}\phi''(\mu_0+\sqrt{\Delta_0}z)=\frac{1}{2}\left[\phi''\right]_0 \tag{59}$$

Substituting in Equation 54, we obtain:

$$\tau\frac{d}{dt}\mu(t)=\delta\mu(t)\left(-1+(m_EN_E+m_IN_I)\left[\phi'\right]_0\right)+\delta\Delta(t)\frac{1}{2}(m_EN_E+m_IN_I)\left[\phi''\right]_0 \tag{60}$$

We similarly obtain the linearized dynamics for $\Delta(t)$:

$$\tau \frac{d}{dt}\Delta = \mathcal{N}_\Delta(\mu(t),\Delta(t),U(t)) = \mathcal{N}_\Delta(\mu_0,\Delta_0,U(t)) + \delta\mu(t)\frac{\partial \mathcal{N}_\Delta}{\partial \mu} + \delta\Delta(t)\frac{\partial \mathcal{N}_\Delta}{\partial \Delta} = \delta\mu(t)\frac{\partial \mathcal{N}_\Delta}{\partial \mu} + \delta\Delta(t)\frac{\partial \mathcal{N}_\Delta}{\partial \Delta} \quad (61)$$

$$\begin{aligned}
\frac{\partial \mathcal{N}_\Delta}{\partial \mu} &= (v_E N_E + v_I N_I)\frac{\partial}{\partial \mu}\left[\phi^2(t)\right] && (62)\\
\frac{\partial}{\partial \mu}\left[\phi^2(t)\right] &= \frac{\partial}{\partial \mu}\int Dz\phi^2(\mu(t)+\sqrt{\Delta(t)}z) = \int Dz 2\phi(\mu_0+\sqrt{\Delta_0}z)\phi'(\mu_0+\sqrt{\Delta_0}z) = 2\left[\phi\phi'\right]_0 && (63)\\
\frac{\partial \mathcal{N}_\Delta}{\partial \Delta} &= -1 + (v_E N_E + v_I N_I)\frac{\partial}{\partial \Delta}\left[\phi^2(t)\right] && (64)\\
\frac{\partial}{\partial \Delta}\left[\phi^2(t)\right] &= \frac{\partial}{\partial \Delta}\int Dz\phi^2(\mu(t)+\sqrt{\Delta(t)}z) = \int Dz 2\frac{z}{2\sqrt{\Delta_0}}\phi(\mu_0+\sqrt{\Delta_0}z)\phi'(\mu_0+\sqrt{\Delta_0}z) && (65)\\
&= \frac{1}{\sqrt{\Delta_0}}\int Dz\sqrt{\Delta_0}\phi'(\mu_0+\sqrt{\Delta_0}z)^2 + \sqrt{\Delta_0}\phi(\mu_0+\sqrt{\Delta_0}z)\phi''(\mu_0+\sqrt{\Delta_0}z) && (66)\\
&= \left[(\phi')^2\right]_0 + \left[\phi\phi''\right]_0 && (67)
\end{aligned}$$

Substituting this in Section A.6.2, we obtain:

$$\tau \frac{d}{dt}\Delta = (v_E N_E + v_I N_I)2\left[\phi\phi'\right]_0 \delta\mu(t) + \left(-1 + (v_E N_E + v_I N_I)\left(\left[(\phi')^2\right]_0 + \left[\phi\phi''\right]_0\right)\right)\delta\Delta(t) \quad (68)$$

Combining Equation 60 and Equation 68, we obtain:

$$\tau \frac{d}{dt}\begin{pmatrix}\delta\mu\\ \delta\Delta\end{pmatrix} = -\begin{pmatrix}\delta\mu\\ \delta\Delta\end{pmatrix} + \begin{pmatrix}(m_E N_E + m_I N_I)\left[\phi'\right]_0 & \frac{1}{2}(m_E N_E + m_I N_I)\left[\phi''\right]_0\\ 2(v_E N_E + v_I N_I)\left[\phi\phi'\right]_0 & (v_E N_E + v_I N_I)\left(\left[(\phi')^2\right]_0 + \left[\phi\phi''\right]_0\right)\end{pmatrix}\begin{pmatrix}\delta\mu\\ \delta\Delta\end{pmatrix} \quad (69)$$

The system is stable to global perturbations if and only if the eigenvalues of $\boldsymbol{M}$ have real part smaller than 1, where:

$$\boldsymbol{M} = \begin{pmatrix}(m_E N_E + m_I N_I)\left[\phi'\right]_0 & \frac{1}{2}(m_E N_E + m_I N_I)\left[\phi''\right]_0\\ 2(v_E N_E + v_I N_I)\left[\phi\phi'\right]_0 & (v_E N_E + v_I N_I)\left(\left[(\phi')^2\right]_0 + \left[\phi\phi''\right]_0\right)\end{pmatrix} \quad (70)$$

To simplify notations, we introduce:

$$\begin{aligned}
a &= (m_E N_E + m_I N_I)\left[\phi'\right]_0 && (71)\\
b &= \frac{1}{2}(m_E N_E + m_I N_I)\left[\phi''\right]_0 && (72)\\
c &= 2(v_E N_E + v_I N_I)\left[\phi\phi'\right]_0 && (73)\\
d &= (v_E N_E + v_I N_I)\left(\left[(\phi')^2\right]_0 + \left[\phi\phi''\right]_0\right) && (74)
\end{aligned}$$

The eigenvalues of $\boldsymbol{M}$ are the roots of the following polynomial:

$$\det(\boldsymbol{M} - \lambda \boldsymbol{I}) = \begin{vmatrix} a-\lambda & b\\ c & d-\lambda\end{vmatrix} = (a-\lambda)(d-\lambda) - bc = \lambda^2 - \lambda(a+d) + ad - bc \quad (75)$$

The discriminant of this polynomial is:

$$D = (a+d)^2 - 4(ad-bc) = (a-d)^2 + 4bc \quad (76)$$

The roots λ_+, λ_- are thus given by:

$$\lambda_\pm = \frac{1}{2}(a+d \pm \sqrt{(a-d)^2 + 4bc}) \quad (77)$$

These roots are illustrated in Figure S.5. Note that, when $m_E N_E + m_I N_I$ is close to zero, both roots are smaller than the radius of the disk of eigenvalues corresponding to the system's response to local perturbations (equation Equation 44). In this case, the matrix $\boldsymbol{S}$ does not have an outlier eigenvalue (see Figure S.4A, middle). For positive $m_E N_E + m_I N_I$ and small $v_E N_E + v_I N_I$, we have $\lambda_+ > r$ (and $|\lambda_-| < r$), and $\boldsymbol{S}$ has one outlier eigenvalue equal to λ_+ (see Figure S.4A, bottom, and Figure S.5). For negative and large enough $m_E N_E + m_I N_I$, we have $\lambda_- < -r$ (and $|\lambda_+| < r$), and $\boldsymbol{S}$ has one outlier eigenvalue equal to λ_- (see Figure S.4A, top and Figure S.5). The real parts of both λ_+ and λ_- are always smaller than 1 (see Figure S.6), so the system is always stable to global perturbations and its stability will only depend on the radius (local stability).

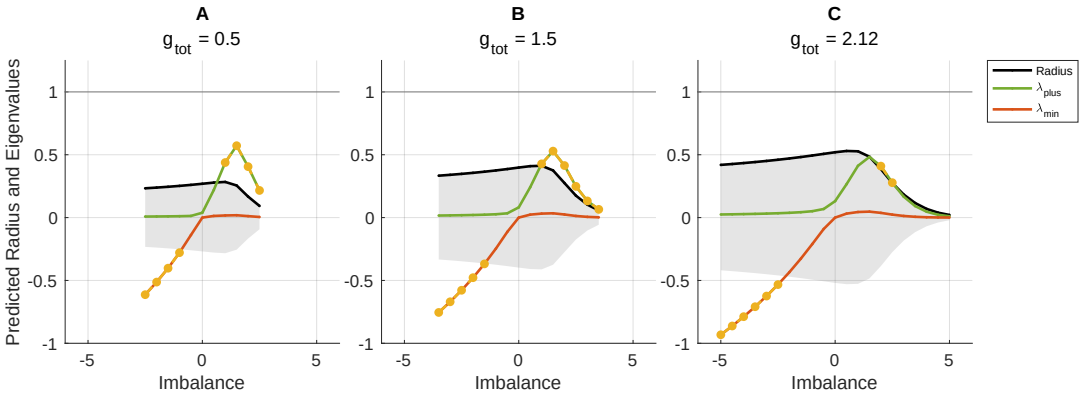

Figure S.5 – Example results of the global stability analysis. *The predictions of the radius and outliers of the global stability matrix are shown for three levels of total connection strength ($g_{tot} = \sqrt{v_E N_E + v_I N_I} = 0.5, 1, 5$, and $\sqrt{2 \cdot 1.5^2}$). Each panel shows the positive (λ_+, green line) and negative (λ_-, red line) eigenvalues of the global stability matrix* $\boldsymbol{M}$*, as well as the radius of its eigenvalue spectrum (black line). The grey area indicates where the eigenvalues are smaller than the radius of the eigenvalue spectrum. Yellow dots indicate outliers, i.e., maximum eigenvalues that are larger than the radius.*

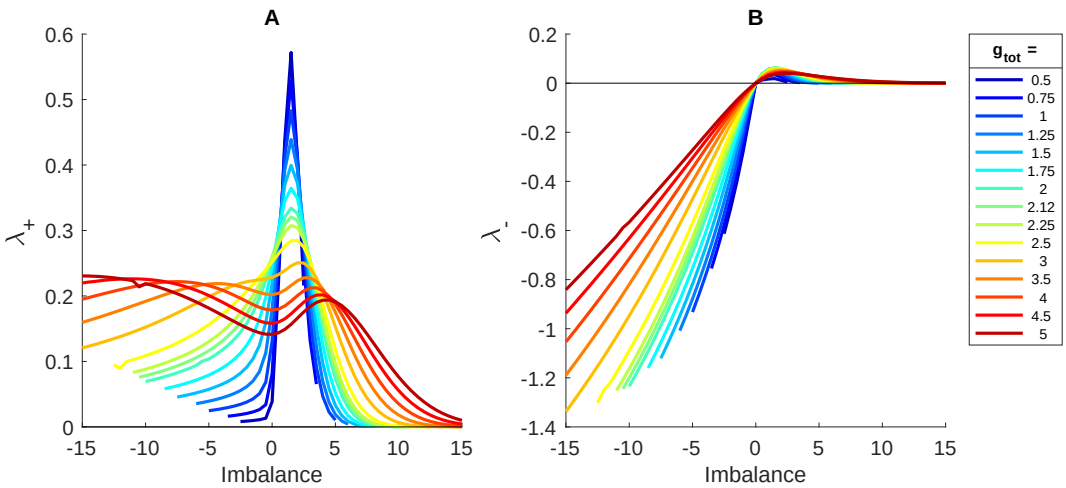

Figure S.6 – Results of the global stability analysis. *Mean-field theory derived values of the (A) positive (λ_+) and (B) negative (λ_-) eigenvalues of the global stability matrix* $\boldsymbol{M}$ *are shown as a function of the imbalance A, for various level of the total connection strength $g_{tot} = \sqrt{v_E N_E + v_I N_I}$*

A.6.3 Integration by parts

We consider any function $f(z)$ such that $\lim_{z\to\infty} f(z)\exp(-\frac{z^2}{2}) = 0$.

We introduce :

$$g(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}, \qquad g'(z) = -z\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \tag{78}$$

We have:

$$\int Dzzf(z) = \int_{-\infty}^{\infty} dz \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} zf(z) = -\int_{-\infty}^{\infty} dz g'(z) f(z) \tag{79}$$

Using integration by parts, we get:

$$-\int_{-\infty}^{\infty} dz g'(z) f(z) = -\left[f(z)g(z)\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} dz g(z) f'(z) = \int_{-\infty}^{\infty} Dz f'(z) \tag{80}$$

Thus:

$$\int Dzzf(z) = \int_{-\infty}^{\infty} Dz f'(z) \tag{81}$$

Appendix B Calculation of spinal cord efferent and afferent signals using computational musculo-skeletal model Myonardo

Myonardo is a Matlab-based, 3D computational musculo-skeletal model developed in the Simscape Multibody 3D simulation environment, consisting of 23 bone segments, 23 joints and 180 muscle-tendon-units (Predimo GmbH. Myonardo [Software]. Münster, Germany). The relative mass and size of the individual segments, as well as the muscular attachments were determined based on the work of Shippen and May (2012). The model was scaled to the subject's body mass and height (Hatze, 1980; Winter, 2009).

As a first step, *Myonardo* calculated the muscle-tendon lengths, muscle-tendon velocities and the muscular lever arms relative to the instantaneous joint centre, based on the instantaneous joint angles and joint angular velocities acquired from the 3D kinematics. The muscle line of action between the origin and the insertion of each muscle-tendon unit is separated into muscle parts between via-points, such that the muscle length was calculated as the sum of the length of each muscle part, i.e. the Euclidean distance between the attachments of the muscle part.

Based on these muscle lengths l_m and velocities v_m, the maximum possible muscle force $f_{m,max}$, could then be calculated using the force-velocity relationship (f_v), as well as the active (f_{la}) and passive (f_{lp}) force-length relationships as follows:

$$f_{m,max} = Act(t) \cdot f_v \cdot f_{la} + f_{lp} \tag{82}$$

where $Act(t)$, the activation of the muscle in the interval [0,1], was set to 1. The Hill-type force-velocity relationship (f_v) was calculated as:

$$f_v(v_m) = \begin{cases} \frac{c}{v_m+b} - a & \text{if } v_m \leq 0 \\ \frac{C}{v_m-B} + A & \text{if } v_m > 0 \end{cases} \tag{83}$$

where properties a, b, c and A, B, C were estimated from the isometric force f_{iso}, resting length l_{opt}, eccentric gain f_{ecc} and the ratio of fast-twitch to slow-twitch fibres $FT\%$, using parameter values from Thaller and Wagner (2004) and Herzog (1999). The active (f_{la}) and passive (f_{lp}) force-length relations were calculated as:

$$f_{la}(l_m) = exp\left[-(\frac{(l_m/l_{opt})^{k_1} - 1}{k_2})^{k_3}\right] \tag{84}$$

$$f_{lp}(l_m) = exp((l_m - l_{opt})/k_4) \tag{85}$$

with $k_1 = 0.96$, $k_2 = 0.35$, $k_3 = 2$, $f_{ecc} = 1.5$, and $FT\% = 0.5$. The remaining muscle properties f_{iso} and l_{opt} were taken from Rajagopal et al. (2016).

Based on the maximum force $f_{m,max}$ for each muscle and the lever arms r, the maximum torque that each muscle can generate around a joint at a certain instance of time was calculated as $T_{max} = r \times f_{m,max}$. Then, a set of muscular activations $Act(t)$, within the boundary $0 \leq Act \leq 1$, can be determined for each joint, such that

$$T_{net} = \sum_{i=1}^{n} T_{max_i} \cdot Act_i \tag{86}$$

with i indicating the n muscles that are acting at the given joint. The required net torque T_{net} is known from inverse dynamics and the maximum torque T_{max} was calculated as described above. As there are more muscles than degrees of freedom of a joint contributing to the joint-torques, the muscular activations of these redundant muscles have to be distributed via optimization. In the present study, the muscular activation was calculated by minimizing the sum of squared muscular activations for each instant of time-based on the linear least-squares solver (`lsqlin`):

$$\min_{Act} \| Act \|^2 \tag{87}$$

Appendix C Network training algorithm

To train the neural networks, we use a recursive least squares algorithm that updates the output weights at each time step, as described by Haykin (2014) and employed in the FORCE algorithm of Sussillo and Abbott (2009).

The output signals of the network were calculated by multiplying the $[k \times N]$-dimensional output weights matrix $\boldsymbol{W}(t)$ with the sigmoid-transformed neuron activations $\boldsymbol{\varphi}(t) = \phi(\boldsymbol{x}(t))$.

$$\boldsymbol{Y}(t) = \boldsymbol{W}(t) \cdot \boldsymbol{\varphi}(t) \tag{88}$$

The output error $\boldsymbol{\epsilon}$, a k-dimensional row vector, was calculated as the difference between the calculated output $\boldsymbol{Y}(t)$ and the target output $\boldsymbol{Y}_{\boldsymbol{target}}(t)$.

$$\boldsymbol{\epsilon}(t) = \boldsymbol{Y}(t) - \boldsymbol{Y}_{target}(t) \tag{89}$$

The output weights $\boldsymbol{W}(t)$ were updated based on the current activity, the output error $\boldsymbol{\epsilon}$, and the $[N \times N]$ learning rate matrix $\boldsymbol{P}(t)$. Here, $\boldsymbol{P}(t)$ denotes an estimate of the inverse of the correlation matrix of $\boldsymbol{\varphi}(t)$ with a regularization term (Sussillo & Abbott, 2009).

$$\boldsymbol{W}(t + \Delta t) = \boldsymbol{W}(t) - \epsilon(t) \frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t)}{1 + \boldsymbol{\varphi}(t)^T \boldsymbol{P}(t)\boldsymbol{\varphi}(t)} \tag{90}$$

$\boldsymbol{P}(t)$ was updated according to:

$$\boldsymbol{P}(t + \Delta t) = \boldsymbol{P}(t) - \frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t) \cdot (\boldsymbol{P}(t)\boldsymbol{\varphi}(t))^T}{1 + \boldsymbol{\varphi}(t)^T \boldsymbol{P}(t)\boldsymbol{\varphi}(t)} \tag{91}$$

Appendix D Supplementary Figures

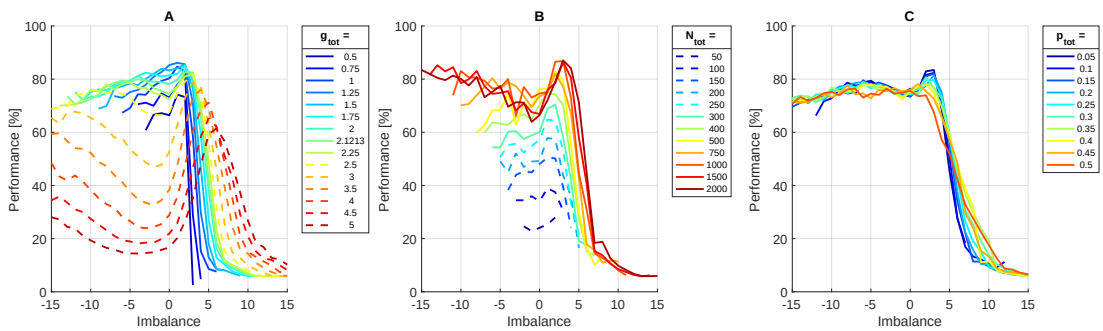

Figure S.7 – Performance as a function of the imbalance *A, for all investigated levels of (A) g_{tot}, (B) N_{tot}, and (C) p_{tot}. Dashed lines indicated networks that have been excluded from the main results (i.e., those with $N_{tot} < 300$ and $g_{tot} \geq 2.5$).*

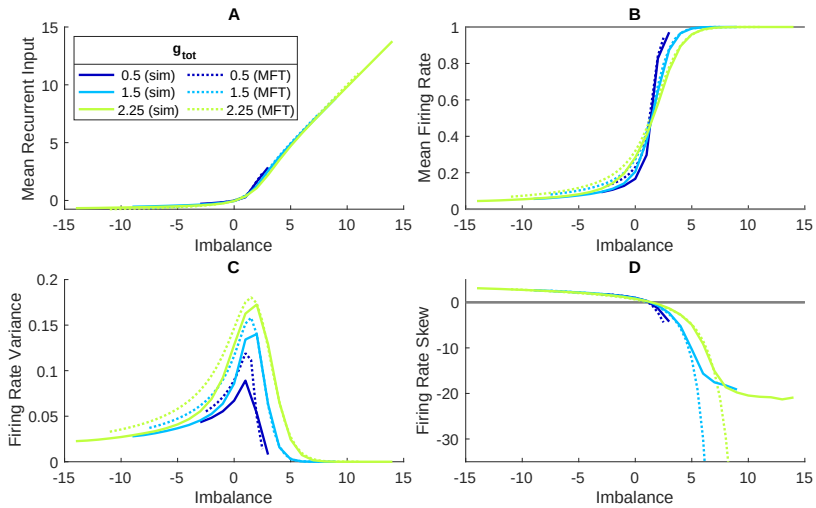

Figure S.8 – Mean-field theory predictions (dotted)versus simulation results (solid) *for three levels of total connection strength ($g_{tot} = 0.5, 1.5,$ and $\sqrt{2 \cdot 1.5^2}$). Note that, for the simulations, results without output feedback are shown, as output feedback has not been taken into account in the MFT predictions.*

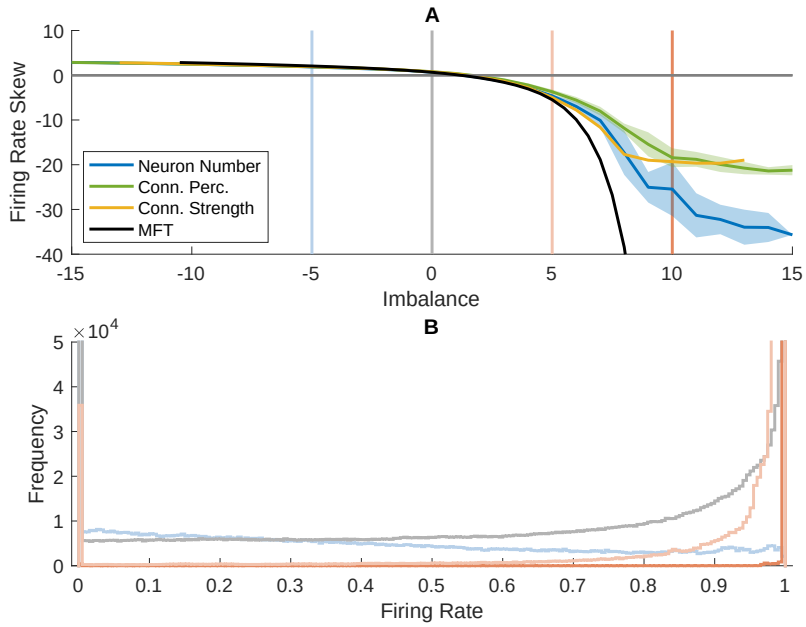

Figure S.9 – Skew of the time-averaged firing rate distributions. *(A) shows the firing rate skew as a function of the imbalance (A) when varying the neuron number (blue), connection percentage (green) and connection strength (yellow), as well as the corresponding MFT predictions for $g_{tot} = \sqrt{2 \cdot 1.5^2}$. For the neuron number and connection percentage, the average and standard deviations over all levels of N_{tot} and p_{tot} are shown. For the connection strength, only $g_{tot} = \sqrt{2 \cdot 1.5^2}$ is shown, to match MFT predictions. The coloured vertical lines indicate the imbalance levels of the four example networks shown in panel (B). (B) shows the firing rate histograms of these four networks, zoomed in to the $(0 - 5e^4)$ range.*

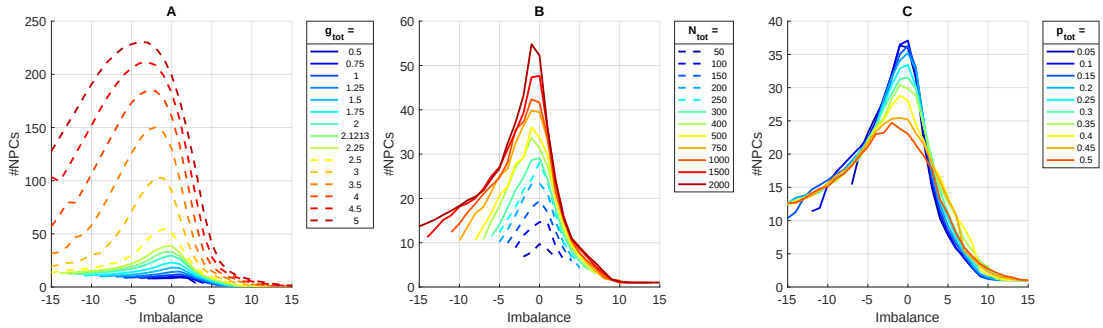

Figure S.10 – Number of principle components (NPCs) as a function of the imbalance *A, for all investigated levels of (A)* g_{tot}*, (B)* N_{tot}*, and (C)* p_{tot}*. Dashed lines indicated networks that have been excluded from the main results (i.e., those with* $N_{tot} < 300$ *and* $g_{tot} \geq 2.5$*).*