Decreased spinal inhibition leads to undiversified locomotor patterns

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

[1]Dept. of Movement Science, University of Münster, Horstmarer Landweg 62b, Münster, 48149, Germany.
[2]Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Münster, Fliednerstraße 21, Münster, 48149, Germany.
[3]Center for Nonlinear Science (CeNoS), University of Münster, Corrensstraße 2, Münster, 48149, Germany.
[4]School of Arts, Sciences and Humanities (EACH), University of São Paulo, Arlindo Bettio 1000, São Paulo, 03828-000, Brazil.
[5]Institut des Systèmes Intelligents et de Robotique (ISIR), Sorbonne Université, CNRS, Paris, F-75005, France.

*Corresponding author(s). E-mail(s): mdegraaf@uni-muenster.de;
Contributing authors: heiko.wagner@uni-muenster.de; mochi@usp.br; charlotte.lemouel@normale.fr;

Abstract

During walking and running, animals display rich and coordinated motor patterns 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 and simulations. We created networks with varying neuron numbers, connection percentages and connection strengths for the excitatory and inhibitory neuron populations, and introduced the anatomical *imbalance* that quantifies the overall effect of the two populations. We trained the networks to reproduce muscle activation patterns derived from human recordings and evaluated their performance. Our results indicate that network dynamics and performance depend critically on the anatomical imbalance in the network. Excitation-dominated networks lead to saturated firing rates, thereby reducing the firing rate heterogeneity and leading to muscle coactivation and inflexible motor patterns. Inhibition-dominated networks, on the other hand, perform well, displaying balanced input to the neurons and sufficient heterogeneity in the neuron firing rate 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

Humans and other animals create rich and coordinated motor patterns during walking and running. With increasing locomotor speed, the stepping frequency increases and the pattern of muscle contraction changes, both in terms of the amplitude of the muscle contractions as well as their timing within the step cycle (Ivanenko et al., 2004, 2006; Lacquaniti et al., 2012). The muscle activation patterns are generated by locomotor central pattern generators (CPGs) within the spinal cord (Delcomyn, 1980; Ijspeert, 2008; Grillner and Wallén, 1985; Grillner et al., 1998). The spinal cord is not only capable of generating the rhythm of locomotion, but also of transforming this simple rhythm into the complex muscle contraction patterns observed during locomotion. In this study, we investigate how this pattern formation is influenced by spinal anatomical characteristics, specifically the excitation-to-inhibition-ratio.

The role of inhibition in CPGs has already 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 (Cowley and Schmidt, 1995; Lanuza et al., 2004; Talpalar et al., 2013; Cohen and Harris-Warrick, 1984). This indicates that inhibition is essential for generating the alternating pattern between flexion and extension (Zhang et al., 2014; Britz et al., 2015). Similarly, inhibitory neurons have also been shown to play an important role in alternation between the left and right sides of the body (Lanuza et al., 2004; Talpalar et al., 2013; Kiehn, 2006; Jankowska, 2008; Quinlan and Kiehn, 2007). Inhibition thus plays a critical role in pattern formation of CPGs, but the mechanism through which it affects motor patterns is unknown. Therefore, we aim to study how the excitatory-to-inhibitory ratio affects locomotor pattern formation.

Separate spinal circuits have been identified that regulate rhythm generation on the one hand and pattern formation on the other. Rhythm-generating neurons are located medially and project locally, whereas pattern-forming neurons are located 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 modelling approaches can successfully replicate experimental phenomena such as non-resetting deletions (Lafreniere-Roula and McCrea, 2005), i.e. the finding that rhythm is maintained even when parts of the movement cycle are skipped (Rybak et al., 2006; McCrea and Rybak, 2007, 2008), as well as the effects of sensory stimulation on locomotor patterns (Rybak et al., 2006; McCrea and Rybak, 2007).

Rhythm generation by CPGs is most often conceptually modelled using the half-centre model (explained in Latash, 2012). This model assumes reciprocal inhibition between two neural pools (the half-centres), with one exciting 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, the activity in the extensor pool decreases by some self-inhibitory process, such as fatigue (Brown, 1911; McCrea and Rybak, 2008; Latash, 2012). 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 is necessary to extend the basic rhythm of alternating flexion and extension into the coordinated multi-joint muscle activation patterns that underlie locomotion. Muscle synergy analysis has identified a set of 4 to 5 invariant activation patterns – the synergies – that describe lower limb muscle activation during healthy locomotion (Dominici et al., 2011; Ivanenko et al., 2004; Chvatal and Ting, 2012; Clark et al., 2010; Gui and Zhang, 2016; Barroso et al., 2014). Such synergies can recreate locomotion in musculoskeletal models at a wide range of speeds (Neptune et al., 2009; McGowan et al., 2010; Aoi et al., 2019; Di Russo et al., 2023), but the mechanisms underlying their formation remain unclear. Conceptual half-centre models struggle to reproduce these diverse activation patterns. Some models propose coupled oscillators for each joint (unit burst generators, Grillner, 1981), or two half-centres 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 coactivation 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).

However, none of these conceptual models explain the concurrent inhibitory and excitatory neural activity observed in nature. According to the half-centre 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. In contrast, 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 and Berg, 2016). These skewed firing rates cannot be recreated by reciprocal inhibition, as reciprocal inhibition leads to disinhibition of the active half-centre and thus to run-away activity (Berg et al., 2019). Recurrent inhibition must thus be present in spinal neural circuits. Since the half-centre models do not incorporate recurrent inhibition, they are unable to investigate its effects, thus indicating the need for an alternative model.

To further investigate the role of 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 simple input signal (Funahashi and 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 (Hinaut and Dominey, 2013; Boström et al., 2013; Wyffels and 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 and Abbott, 2009) and joint angles (Wyffels and Schrauwen, 2009; Hoellinger et al., 2013).

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 and Sompolinsky, 2015) or in networks with balanced excitation and inhibition (Harish and Hansel, 2015). Recurrent neural networks thus provide a versatile and insightful framework for studying neural pattern formation (Wyffels and Schrauwen, 2009; Hoellinger et al., 2013).

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 biological and artificial neural networks has been extensively studied (reviewed in, e.g., Hennequin et al., 2017; Sadeh and Clopath, 2021; Herstel and Wierenga, 2021; Liang et al., 2024; Isaacson and Scanziani, 2011), 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 (Sussillo and Abbott, 2009; Boström et al., 2013; Jaeger, 2001), causing them to exert both excitatory and inhibitory influences. In spiking neural networks, the neurons do obey Dale's law, but the ratio between the number of excitatory and inhibitory neurons remains relatively unexplored. Most studies use the same 4:1 excitatory-to-inhibitory ratio to reflect the population sizes found in the cortex (Marom and Shahaf, 2002; Sahara et al., 2012; Marín, 2012; Wonders and Anderson, 2006; Meinecke and Peters, 1987). Furthermore, none of these studies focused on motor pattern generation, where spinal cord-specific EI ratios might be more relevant. Findings in rats indicate the anatomical EI ratio might be smaller in the spinal cord (Todd and Sullivan, 1990). This indicates that the 4:1 excitatory-to-inhibitory ratio is not omnipresent, and that inhibition might play a larger role in spinal motor pattern generation. This

indicates that a more nuanced understanding of the excitatory-to-inhibitory balance is needed for spinal motor control.

The goal of our paper is to study the role of recurrent inhibition and global anatomical EI balance in spinal motor pattern formation. We adapt an existing reservoir computing model (Sussillo and 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 the population's sizes and connectivities. We initialize the networks with different anatomical imbalances and train them to reproduce muscle activation patterns calculated from human locomotion recordings. We analyze the effects of the imbalance 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 for 17 muscles 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 from both our simulations and from a theoretical assessment using mean-field theory (Sompolinsky et al., 1988; Rajan et al., 2010; Mastrogiuseppe and Ostojic, 2017, 2018).

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 (fig. 1, red triangles in the reservoir) and an inhibitory population I with neuron number N_I (fig. 1, blue circles in the reservoir). Any given neuron i in either population receives external input, output feedback and recurrent input.

The external input represents the output of the rhythm generation layer and equals $J^{In}U(t)$, where $U(t) = 1 + \sin(\omega t)$ with ω equal to the target stride frequency (fig. 1, bottom left), and J^{In} the input weights drawn from a Gaussian distribution with mean 0 and variance $v_{in} = \frac{1}{3}$. The mean and variance of the distribution were chosen to match those of Sussillo and Abbott (2009). While Sussillo & Abbott employed a uniform distribution for their input weights, we used a Gaussian distribution to align with the mean-field theory, which assumes Gaussian-distributed inputs.

The output feedback to the network equals $J_i^{Fb}Y(t)$, where $Y(t)$ is the 17-dimensional output of the network (fig. 1, bottom right) and the feedback weights J^{Fb} are drawn from a uniform distribution between -1 and 1.

For the recurrent inputs within the network, we use a sparse connection matrix, reflecting the sparse connectivity of biological neural networks. The connection weights J_{ij}^E and J_{ij}^I have a probability of p_E, resp. p_I, of being non-zero. To ensure Dale's law, the weights for the non-zero connections are drawn from either a positive (for excitatory neurons) or negative (for inhibitory neurons) truncated 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 full set of 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 is the parameter that scales the excitatory connection strength (Sompolinsky et al., 1988; Sussillo and 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:

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

with time constant $\tau = 0.01$. The neurons' firing rate equals $\phi(x_i)$. For biological realism, the transfer function ϕ should result in (1) positive

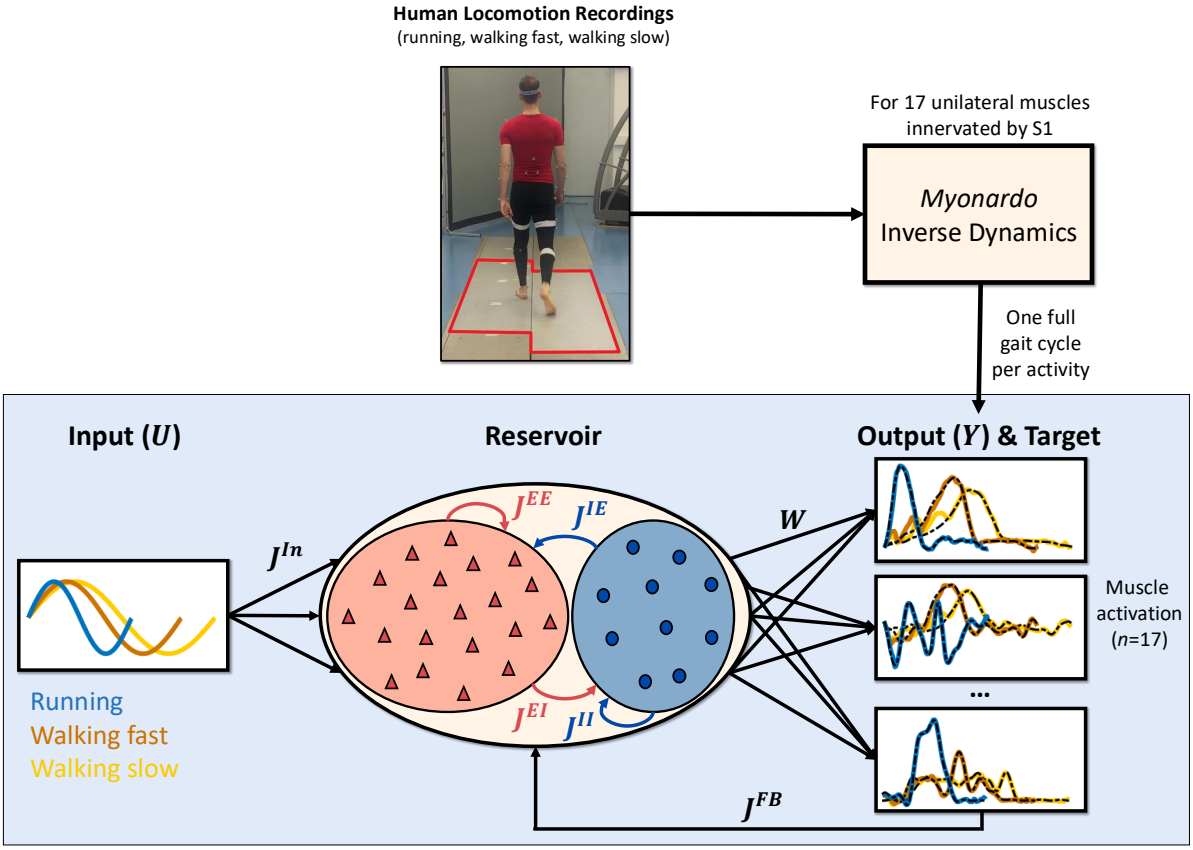

Fig. 1 Schematic overview of the methods. Human locomotor signals were recorded from one adult during slow walking, faster walking, and running (top middle). Ground reaction forces were recorded using six Kistler Force Platforms (outlined in red) and kinematics were recorded using a Qualisys measuring system. The recorded kinetics and kinematics were then input into the musculoskeletal model *Myonardo* for inverse dynamics calculations (top right), which provided the required muscle activations 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, of which the neuron number, connection percentage and connection strength have been varied separately. The current figure shows an example network that has more excitatory than inhibitory neurons. 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 (orange), and slow walking (yellow). The muscle activation outputs resulted from all-to-all connections between the neurons in both populations via trained output weights W. The network outputs were fed back into the network via feedback weights J^{Fb}. Example outputs are shown for all three locomotor activities (coloured lines) overlayed on the target outputs (dashed black lines) for three of the 17 output signals (from top to bottom: m. gastrocnemius, m. flexor digitorum longus, and m. gluteus medius)

.

firing rates, (2) zero firing rate for negative input, as biological neurons do not fire when they receive negative or zero input, and (3) saturated firing rates for large positive input, matching experimental findings of firing rate saturation (Prinz et al., 2003). Therefore, we chose the rectified hyperbolic tangent, where ϕ is $\tanh(\cdot)$ if x_i is positive, and zero otherwise. Provided requirement (1) is satisfied, the choice of transfer function is not critical. The paper's main results were duplicated when using a sigmoid transfer function (see fig. 11).

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, which is initialized as zeroes and evolves throughout training (see section 2.3.3 and section B).

Our network is an adaptation of an echo state neural network (Sussillo and 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 and Abbott, 2009). Second, the firing rates of all neurons were constrained to be positive, whereas the firing rate is $\tanh(x_i)$ in the original network, 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 fig. 1) or inhibitory (blue circles in fig. 1).

2.2 Mean-field theory analysis

We performed a theoretical analysis of the networks using the mean-field theory, a technique where a description of the network dynamics can be derived self-consistently by averaging over the random parts of the network connectivity (Sompolinsky et al., 1988; Rajan et al., 2010; Mastrogiuseppe and Ostojic, 2017, 2018). 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 section A, but we will present a short overview here.

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:

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

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

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. Note that the mean-field predictions are generally valid for any variances and transfer functions, and can easily be adapted.

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 first parameter as the imbalance A:

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

By definition, the mean weight of the excitatory connection m_E is positive, and the mean weight of the inhibitory connections m_I is negative, resulting in a total imbalance that can either be positive, in networks dominated by excitation, or negative, in networks dominated by inhibition. Our mean-field analysis investigated ranges from -15 to +15, to match the values from our simulations (described in section 2.3.1).

Next, we assumed 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. Since neural integration occurs over a much faster timescale ($\tau = 10$ ms) than locomotion ($\frac{1}{\omega} = 720$ ms for running), this assumption provided a good match between theory and simulations (see fig. 16). In the quasi-stationary state, the activation x_i of each neuron has the same mean and variance as its input. We explicitly modelled 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 neural input mean μ and variance Δ (see Appendix section A.3):

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

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

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 numerically solved this set of four equations with four unknowns to self-consistently determine the stationary solution given by $\mu_0, \Delta_0, [\phi]_0, [\phi^2]_0$.

Finally, we performed a stability analysis and showed that the stationary solution loses stability

for large values of $v_E N_E + v_I N_I$ (see Appendix section A.6). We determined the maximal value of $v_E N_E + v_I N_I$ for which networks are stable for all values of imbalance A and restricted our simulations to this stable range (i.e. $g_{tot} \leq 2.25$, see fig. 12a). 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). We find that, with this restriction, network performance does not depend on $v_E N_E + v_I N_I$, or equivalently, g_{tot} (see fig. 15a).

2.3 Simulations

The differential equation governing the change in neural activation (eq. (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:

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

We investigated networks with imbalances ranging from -15 to $+15$ in steps of 1, by globally varying 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_E and N_I, 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 fig. 15). The connection percentage was set constant at $p_E = p_I = 0.1$, corresponding to connection ratios found in *ex vivo* neural networks (Marom and Shahaf, 2002), and the connection strength at $g_E = g_I = 1.5$, as this has been shown to produce good network performance (Sussillo and Abbott, 2009).

Second, the imbalance was modified via p_E and p_I, 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_E and g_I, 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 fig. 12). 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 performed at self-chosen speeds: slow walking, fast walking, 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 fig. 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, 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 Psychology and Sports Science at 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, Münster, Germany). For more information on the musculoskeletal model and the performed calculations, see section C 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. gastrocnemius, 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 21 strides, with the three activities randomly interleaved. A different test signal was created for each of the twenty network instances, by changing the sequence. 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 the first 20 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 section B). 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 derivatives reinitialized at the start of each iteration.

2.3.4 Outcome parameters

Network Dynamics

Network dynamics were assessed by recording the mean and neuron-to-neuron (or network) variance of the recurrent input x and the firing rates φ. This network variance provides a measure for the firing rate heterogeneity, reflecting the diversity of neuronal activity across the population.

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 (coloured lines, fig. 2) and target (black lines, fig. 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 fig. 2) or a failure if RMSE > 0.05 (shaded in red in fig. 2).

Effective network dimensionality

We quantified the effective dimensionality of the internal network dynamics as the number of network principal components (NPCs) necessary to account for 99% of the variance in the neuron firing rates. The effective network dimensionality characterizes the complexity of the network dynamics by indicating the number of unique components that are needed to capture the majority of the variance in its activity. Low dimensionality indicates that the neuronal activity patterns in the network are highly correlated with each other (as in the right-hand side panels of fig. 17), whereas high dimensionality indicates diverse neuronal activity, which allows the network to capture more

complex patterns (as in the left-hand side panels of fig. 17). The same approach was applied to quantify the number of principal components in the output and target output signals.

Muscle parameters

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

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

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 coactivation 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.

3 Results

We present our findings on the impact of the global 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.

3.1 Mean-field theory

The mean-field predictions for the mean and variance of both the recurrent input and the firing rate showed a good match to simulations for all studied parameters (see figs. 3 and 16). 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 (A) and the total connection strength (section A, eqs. (38) and (39)). While the total connection strengths have quantitative differences, they do exhibit qualitatively similar behaviour (fig. 16).

3.2 Network Dynamics

As predicted by our mean-field theory analysis (fig. 3, black), we found that the network dynamics depend critically on the excitatory to inhibitory imbalance, regardless of whether it was varied in terms of neuron number (fig. 3, blue), connection percentage (fig. 3, green) or connection strength (fig. 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 (fig. 3a). Here, the incoming excitatory current (dashed lines in fig. 3b) overpowers the incoming inhibitory current (dotted lines in fig. 3b). This large excitatory current causes the mean firing rate to saturate (fig. 3c), which not only keeps the neuronal input high, but also results in a drastic drop in the firing rate variance (fig. 3d, with most or all neurons firing at their maximum constantly (see fig. 17, $A = 5, 7.5$ and 10). This is also evidenced by a strong negative skewness in the firing rate distributions (see fig. 3e, and example histograms in fig. 18).

In contrast, for inhibitory networks ($A < 0$), the mean input received by the neurons is negative but small (fig. 3a), with the recurrent inputs from both the excitatory and inhibitory populations relatively balanced and close to zero (fig. 3b). As a results, small variations in the inputs to individual neurons can drive variations in the neurons' firing rate, thus preventing the firing rates from saturating at their minimum value of zero (fig. 3c and fig. 17, $A = -5$ and 0). Consequently, the firing rate variance does not drop to zero (fig. 3d) and the firing rate distribution has a moderately positive skew (fig. 3e).

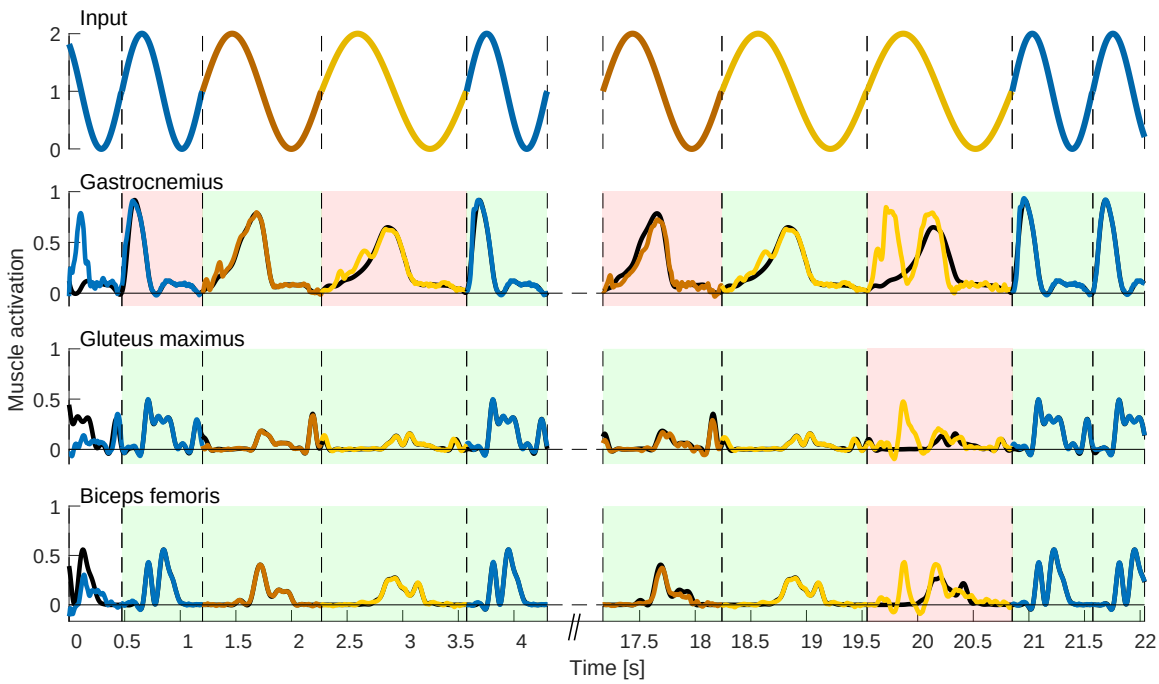

Fig. 2 Performance calculation in example output signals. The networks are tested on their ability to reproduce the motor patterns of randomly interleaved gait cycles. The input to the network (top row) and network outputs of three example muscles (bottom rows) are shown for slow walking (yellow), fast 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, the cycle is classified as a success (green background) if the root mean square difference between the target and the actual output is less than 0.05. 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. This figure shows the output of a random network with $N_E = N_I = 375$, $p_E = p_I = 0.1$ and $g_E = g_I = 1.5$ and a performance of 78.2% (66.7% if only taking into account the visible cycles)

3.3 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 firing rate heterogeneity in excitation-dominated networks, the dimensionality of the network drops drastically when excitation increases (fig. 4a and fig. 17, $A = 5, 7.5$ and 10). This is accompanied by a concurrent drop in network performance (fig. 4b). While the dimensionality in inhibitory networks is also lower than in balanced networks (fig. 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 fig. 4a,c). This means that the effective dimensionality of inhibition-dominated networks is theoretically large enough to reproduce all dimensions present in the target output, as reflected by their relatively high performance (fig. 4b).

Indeed, network performance depends critically on the effective dimensionality of the networks (fig. 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 (fig. 4c). Thus, there 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. Although the influence of the imbalance on the effective dimensionality depends on g_{tot}, N_{tot}, and p_{tot} (see fig. 19), the influence of the imbalance on performance is largely independent of these factors (fig. 15, solid lines). This difference can be explained by the relatively

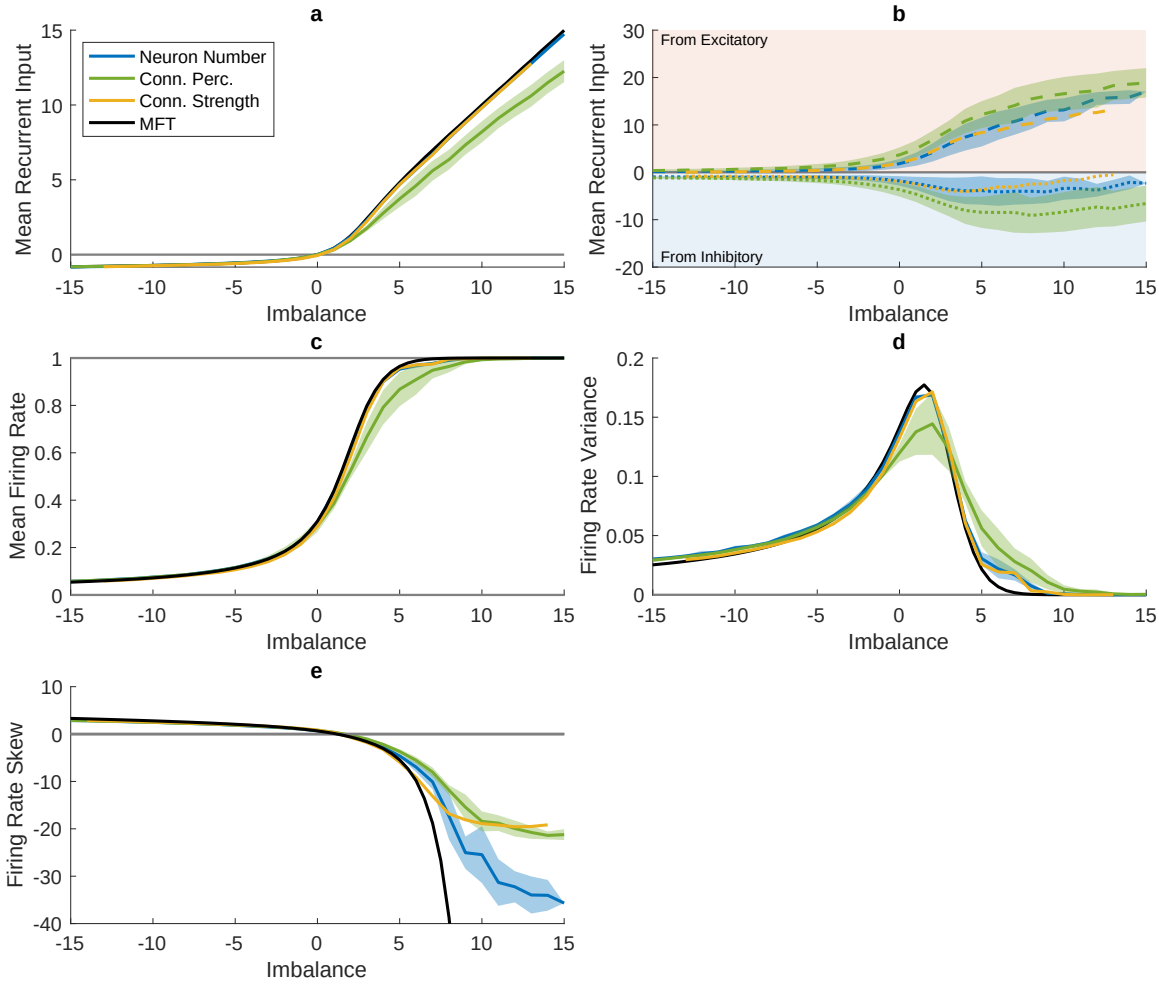

Fig. 3 Network Dynamics as a function of imbalance A, with (a) showing the mean of the input to individual neurons (x), (b) the mean input split between the input coming *from* the excitatory (positive; in the red-shaded area) and inhibitory (negative; in the blue-shaded area) populations, (c) the mean of the firing rate (φ), (d) the firing rate variance, and (E) the skew of the firing rate distribution. The averages over all 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 fig. 16). 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

wide plateau and slow drop-off in fig. 4c), indicating that, above a certain threshold, a change in the number of effective dimensions does not necessarily lead to a large change in performance.

3.4 Motor patterns

Shifting the imbalance toward inhibition also produces a markedly different effect on the motor output compared to shifting it towards excitation. fig. 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 fig. 5) can produce distinct motor outputs for the three tasks. In contrast, in excitation-dominated networks (third, fourth and fifth column in fig. 5),

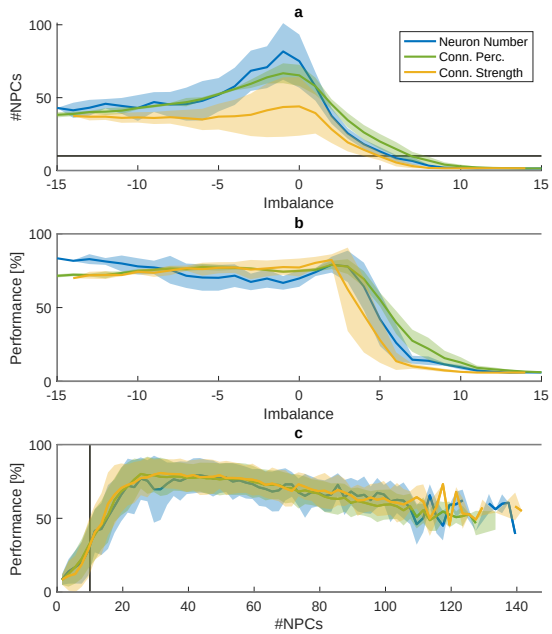

Fig. 4 Effective dimensionality and performance of the networks. The imbalance A is plotted against (a) the number of network principal components (#NPCs) and (b) the network performance. (c) depicts the network performance as a function of the number of NPCs. The average over 20 network instances and 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)

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 fig. 6b). In addition, the outputs of the different muscles become more similar. This is evidenced by an increase in the coactivation index CAI (see fig. 6a). The reduced firing rate heterogeneity and effective dimensionality in the excitatory networks are thus accompanied by a simplification of the networks' outputs.

4 Discussion

This study aimed to explore how the global 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, thus, to reduced network variance. Conversely, inhibition-dominated networks were expected to maintain higher network variance. 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 artificial neural networks (reviewed in, e.g., Hennequin et al., 2017; Sadeh and Clopath, 2021; Herstel and Wierenga, 2021; Liang et al., 2024; Isaacson and Scanziani, 2011), demonstrating that a functional EI balance is crucial for proper neural functioning. However, this functional balance does not necessarily result from 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 and Sullivan, 1990), and only around 20% of the neurons in the mammalian cerebral cortex (Marom and Shahaf, 2002; Sahara et al., 2012; Marín, 2012; Wonders and Anderson, 2006; Meinecke and Peters, 1987). 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

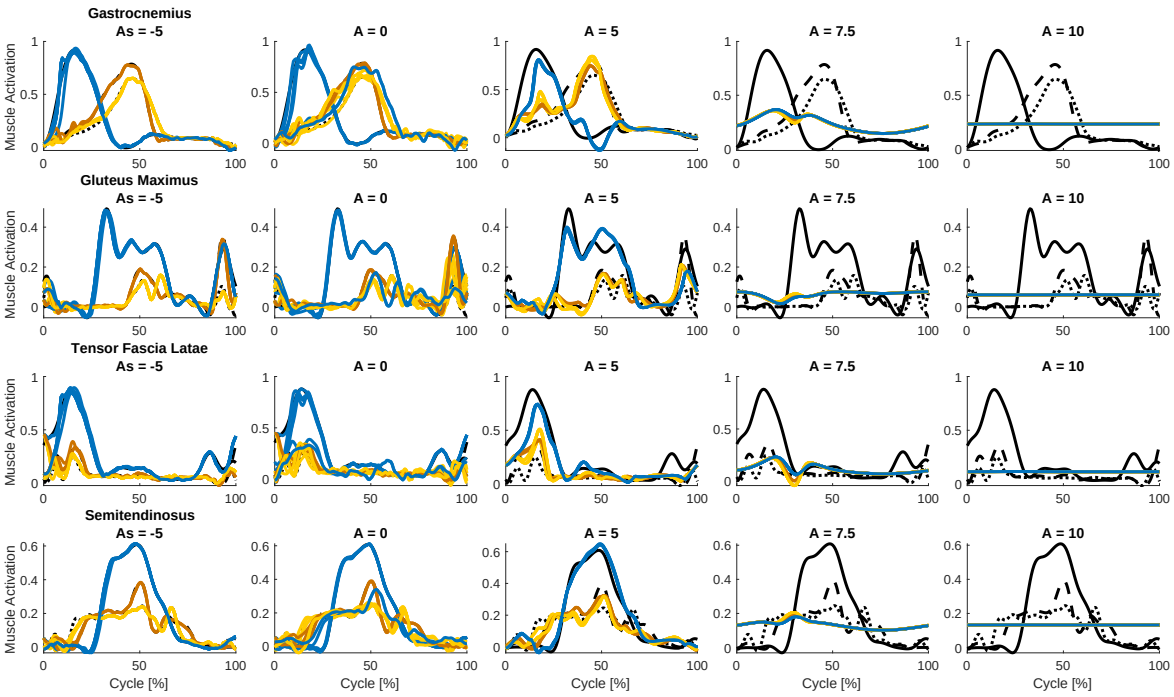

Fig. 5 Example output signals plotted as a function of the cycle percentage for various imbalances, with 0% and 100% corresponding to right heel strikes. Each column shows signals belonging to a random example network with an imbalance of $A = -5$, 0, +5, and +10 from left to right. The imbalances were varied by changing the connection strengths for the two populations (from $g_E \approx 0.92$ and $g_I \approx 1.91$ for $A = -5$, to $g_E \approx 2.12$ and $g_I \approx 0.14$ for $A = 10$), while keeping the other parameters constant ($N_E = N_I = 375$, $p_E = p_I = 0.1$, $g_{tot} = \sqrt{2 \cdot 1.5^2}$). 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 activity, 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. These example networks had a success percentage of 85%, 71%, 50%, 8% and 6% respectively

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 asymmetrical or imbalanced.

To investigate the effects of the anatomical EI balance, we introduced the anatomical *imbalance* to quantify the excitatory and inhibitory populations' influence 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 rises beyond the stable region (see fig. 15a, $g_{tot} \geq 2.5$), or when the networks become too small (fig. 15b, $N_{tot} \leq 250$). The imbalance and strength equations are expressed in terms of 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

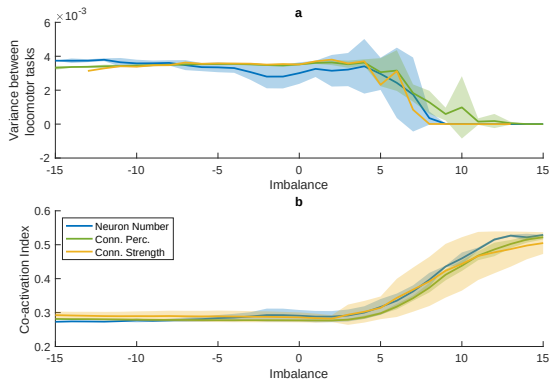

Fig. 6 Increased excitation leads to a simplification of the motor output, both between muscles as well as between locomotor tasks. (a) average coactivation index between each pair of muscles and (b) variance between the outputs for the different locomotor tasks (i.e., slow walking, fast walking, and 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 fig. 16c)

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 and Berg, 2016) and various cortical areas (Shafi et al., 2007; Hromádka et al., 2008; O'Connor et al., 2010) 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; Vogels et al., 2005; van Vreeswijk and Sompolinsky, 1996; Lindén and Berg, 2021; Roxin et al., 2011), whereas excitatory networks exhibit a distribution with a single peak at the maximum firing rate (Lindén and Berg, 2021). Similarly, Roxin et al. (2011) showed *in silico* that increasing the inhibitory neurons' firing rate via the external input transforms the firing rate distribution from negatively skewed to a more lognormal distribution, 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 and 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 network dimensionality is crucial for network performance (Susman et al., 2021; Carroll and Pecora, 2019; Carroll, 2020). In our study, networks with too few principal components could not replicate the target output signals. This matches prior 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 and Pecora, 2019). In excitatory networks, the effective reservoir dimensionality is reduced below the target output dimensionality and 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 (Sadeh and Clopath, 2021; Herstel and Wierenga, 2021). These results highlight the importance of recurrent inhibition in maintaining appropriate effective network dimensionality.

Third, our results indicate that inhibition plays a crucial role in generating diverse and variable motor patterns. High excitation networks could not produce different outputs for different muscles, as evidenced by the high coactivation 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 variance. These findings align with *in vivo* studies, where blocking inhibition led to synchronous activation of flexor and extensor motoneurons (Beato and Nistri, 1999; Cazalets et al., 1998; Cowley and 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 observed in various motor disorders. 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; Bekius et al., 2020), 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 (Clark et al., 2010; Cheung et al., 2012), 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 and Stagg, 2021). The early neuroprotective effect of inhibition is 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 (Rothwell et al., 1983; Cohen and Hallett, 1988), 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 and Ayala, 1987; Dudek and Sutula, 2007), autism (Casanova et al., 2003; Rubenstein, 2010; Rubenstein and Merzenich, 2003; Markicevic et al., 2020) and schizophrenia (Yizhar et al., 2011; Murray et al., 2014). 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 (Zampella et al., 2021; Jansiewicz et al., 2006; de Jong et al., 2011). In schizophrenia, motor abnormalities such as catatonia have been linked to EI imbalance (Walther and 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 locomotor activities, measured in 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), we do not know how well our results generalize to untested frequencies. It has, however, been shown that reservoir computing networks are capable of 'morphing' to untrained signals (Boström et al., 2013). This is also the case for our network (see fig. 20), although we could not quantify its morphing ability as we did not have target data for the intermediate frequencies. Therefore, we performed an additional experiment in which we trained networks with various imbalances to recreate motor primitives, a set of temporal basis functions, and generalize them to various frequencies. Previous experimental work in humans has shown that muscle activities over a range of locomotor speeds can be accounted for by five such primitives

(schematically represented in fig. 7b), each peaking at a different part of the locomotor cycle. Confirming our main results, balanced and inhibition-dominated networks were able to generalize to the untrained frequencies, while excitation-dominated networks could not (see fig. 7). When used as inputs in neuromechanical models such as those of Aoi et al. (2019) or Di Russo et al. (2023), these primitives could thus generate locomotion over a wide range of speeds. This finding underscores our network's potential to create locomotor patterns for unseen speeds.

Second, sensory feedback, which is known to contribute to stability, perturbation responses, and even steady-state muscle activation during locomotion (Frigon et al., 2021), was not included in our model. Nonetheless, we argue that our findings most likely generalize, as feed-forward pattern formation plays a crucial role in shaping locomotor output, regardless of sensory feedback. Experimental work in deafferented animals has shown that patterns more complex than simple flexion-extension alteration could be produced in these animals without sensory input, albeit not as robustly (Grillner and Zangger, 1984; Taub et al., 1975; Grillner et al., 1995). Moreover, studies have employed neuromechanical models to explore the roles of both sensory feedback and pattern formation. These works have shown that neuromechanical models cannot rely on sensory feedback alone: a pattern formation layer in some form was needed to generate stable locomotion over a range of speeds (Geyer and Herr, 2010; Di Russo et al., 2023). Our findings demonstrate that excessive excitation impairs this intrinsic pattern formation (fig. 4b). According to the aforementioned modelling studies, adding sensory feedback is not sufficient to restore locomotion in the excitation-dominated networks. Additionally, our generalization experiment (described in the paragraph above, and in fig. 7) shows that our networks can produce the temporal primitives that are able to generate stable locomotion over a range of speeds, when used as input to a neuromechanical model with sensory feedback. We, therefore, expect our findings to generalise to a model with sensory feedback: balanced and inhibition-dominated networks will be able to produce stable locomotion, whereas excitation-dominated networks will not.

Third, 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.

Fourth, 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 (Vogels et al., 2011; Froemke, 2015; Field et al., 2020). 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.

Finally, our networks had a randomly initialized connectivity, that does not necessarily reflect the structure of the spinal neural circuits. Structured connections are often modelled in neural circuit models, which are low-dimensional and focus on the properties of specific neural subgroups. However, these models typically fail to capture the heterogeneity present within the subgroups they model. The higher-dimensional neural network models that we employ can account for this heterogeneity, but the complexity of the mathematical analysis grows with the number of subpopulations that are included. Studying the role of specific connectivities in high-dimensional network models is therefore challenging, and currently the subject of active research. Shao and Ostojic, for example, investigated the effect of various neural motifs on network dynamics and showed that adding reciprocal motifs (i.e., modeling the effect of renshaw cells) only has a small effect on network dynamics (Shao and Ostojic, 2023), whereas chain motifs have a much larger effect: increasing the strength of chain motifs can induce instability even in inhibition-dominated networks (Shao et al., 2024). However, the presence of chain motifs in the spinal cord has not been documented.

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

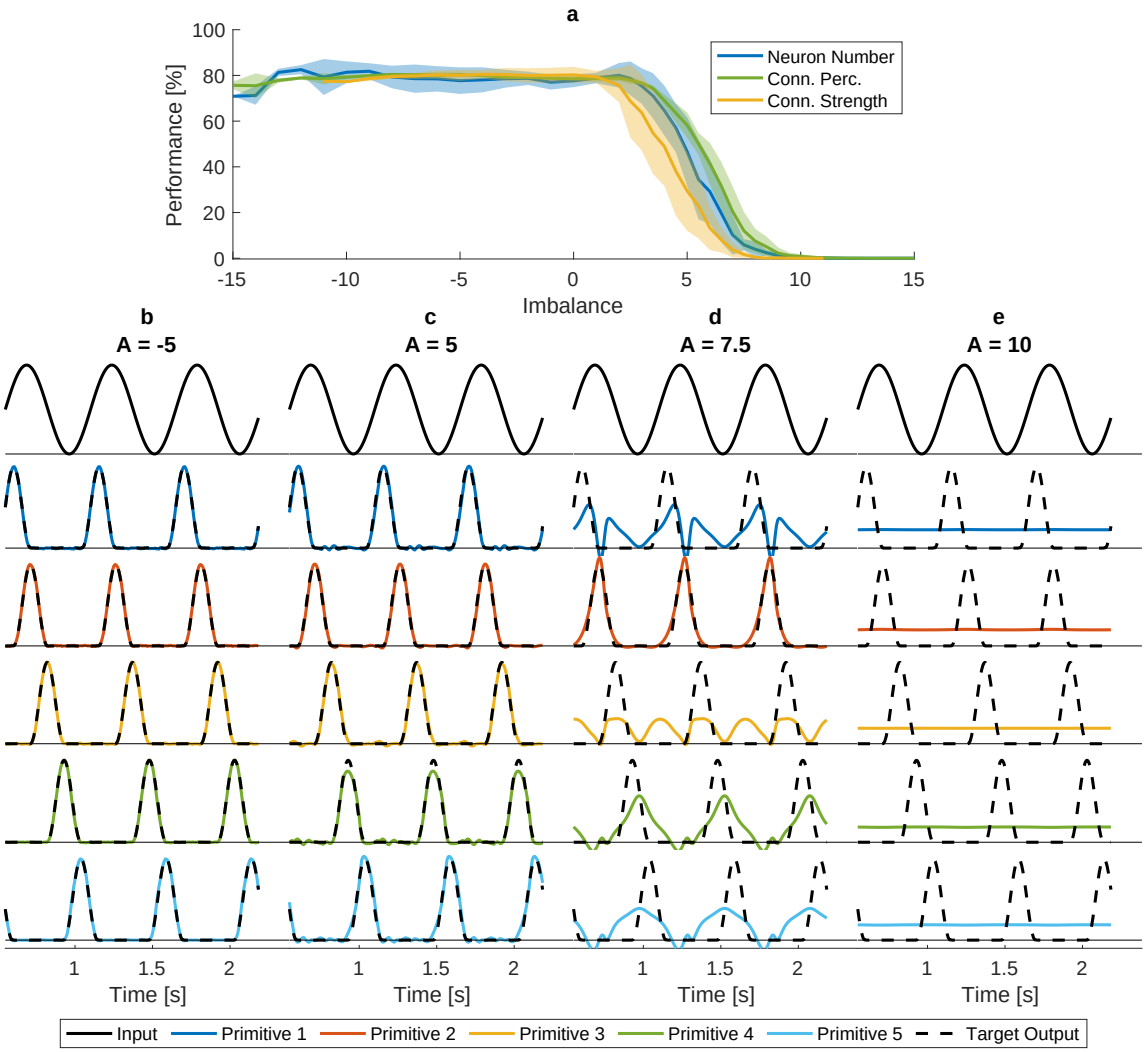

Fig. 7 Performance on a generalization task. Networks with various imbalances A and total connections strengths g_{tot} were trained to reproduce 5 bell-shaped patterns representing motor primitives (based on Di Russo et al., 2023). Networks were trained on patterns corresponding to three speeds (0.6, 1.2, and 1.6 m/s) and tested on speeds ranging from 0.3 to 2.0 m/s in steps of 0.1 m/s. (a) Performance of the networks on untrained speeds, as a function of imbalance. 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. (b-e) Inputs (top row), outputs (coloured lines) and target outputs (dashed black lines) for the five motor primitive patterns for networks with $N_E = N_I = 375$, $p_E = p_I = 0.1$, and $g_{tot} = \sqrt{2 \cdot 1.5^2}$ and imbalances of (b) $A = -5$, (c) $A = 5$, (d) $A = 7.5$, (e) $A = 10$ that were varied via g_E and g_I

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.

Supplementary information. Not applicable.

Acknowledgements

We thank Predimo GmbH for providing the Myonardo software.

Declarations

Funding

This work was supported by the *Interdisciplinary teaching programme on Machine Learning and Artificial intelligence* (InterKI), via the Federal ministry of Education and Research (BMBF). Grant number: 16DHBKI049

Competing interests

H. Wagner is shareholder in Predimo GmbH. The other authors declare that they have no competing interests that could have influenced the results of this paper.

Ethics approval and consent to participate

The study has been approved by the Ethics Committee of the Faculty of Psychology and Sports Science at the University of Münster (#2019-10-RD). The participant signed the informed consent form before the start of the measurements.

Consent for publication

The authors affirm that human research participants provided informed consent for publication of the images in fig. 1.

Data and code availability

Data and code are made available at https://doi.org/10.5281/zenodo.13981334

Materials availability

Not applicable.

Author contribution

Conceptualization: Myriam de Graaf, Heiko Wagner; Methodology: Myriam de Graaf, Heiko Wagner, Luis Mochizuki, Charlotte Le Mouel; Formal analysis and investigation: Myriam de Graaf, Charlotte Le Mouel; Writing - original draft preparation: Myriam de Graaf, Charlotte Le Mouel; Writing - review and editing: Myriam de Graaf, Heiko Wagner, Luis Mochizuki, Charlotte Le Mouel; Funding acquisition: Heiko Wagner; Supervision: Heiko Wagner, Luis Mochizuki, Charlotte Le Mouel.

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 (Sompolinsky et al., 1988; Rajan et al., 2010; Mastrogiuseppe and Ostojic, 2017, 2018). 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 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 $\mathcal{J}_{ij}^P$ with mean 0 and variance v_P: $J_{ij}^P = m_P + \mathcal{J}_{ij}^P$ 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:

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

where

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

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 $[\varphi(t)] = \left[\varphi_j^E(t)\right] = \left[\varphi_j^I(t)\right]$ 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:

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

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 $\left[\mathcal{J}_{ij}^P \varphi_j^P(t)\right] = \left[\mathcal{J}_{ij}^P\right]\left[\varphi_j^P(t)\right]$. Since $\left[\mathcal{J}_{ij}^P\right] = 0$, we have:

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

Therefore:

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

A.2.2 Variance Δ of the input to individual neurons

The variance Δ of the input to each neuron equals $\left[I_i(t)^2\right] - \left[I_i(t)\right]^2$. Since the terms in the right-hand side of eq. (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: $\varphi^E(t) = \varphi^I(t) = [\varphi(t)]$. 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:

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

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

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

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

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

Therefore:

$$
var\left(\eta_i^P(t)\right) = v_P \sum_{j=1}^{N_P} \left[(\varphi_j^P(t))^2\right] = v_P N_P \left[\varphi^2(t)\right]
\tag{24}
$$

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

$$
cov\left(\eta_i^E(t), \eta_i^I(t)\right) = \left[\sum_{j=1}^{N_E} \mathcal{J}_{ij}^E \varphi_j^E(t) \sum_{k=1}^{N_I} \mathcal{J}_{ik}^I \varphi_k^I(t)\right] - \left[\sum_{j=1}^{N_E} \mathcal{J}_{ij}^E \varphi_j^E(t)\right]\left[\sum_{j=1}^{N_I} \mathcal{J}_{ij}^I \varphi_j^I(t)\right]
\tag{25}
$$

$$
= \sum_{j=1}^{N_E}\sum_{k=1}^{N_I} \left[\mathcal{J}_{ij}^E \mathcal{J}_{ik}^I\right]\left[\varphi_j^E(t)\varphi_k^I(t)\right] - \sum_{j=1}^{N_E}\sum_{k=1}^{N_I} \left[\mathcal{J}_{ij}^E\right]\left[\varphi_j^E(t)\right]\left[\mathcal{J}_{ik}^I\right]\left[\varphi_k^I(t)\right]
\tag{26}
$$

The recurrent weights from the two populations are independent, therefore $\left[\mathcal{J}_{ij}^E \mathcal{J}_{ik}^I\right] = 0$ and:

$$
cov\left(\eta_i^E(t), \eta_i^I(t)\right) = 0
\tag{27}
$$

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 $[\varphi(t)]$ 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 eqs. (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 (fig. 8).

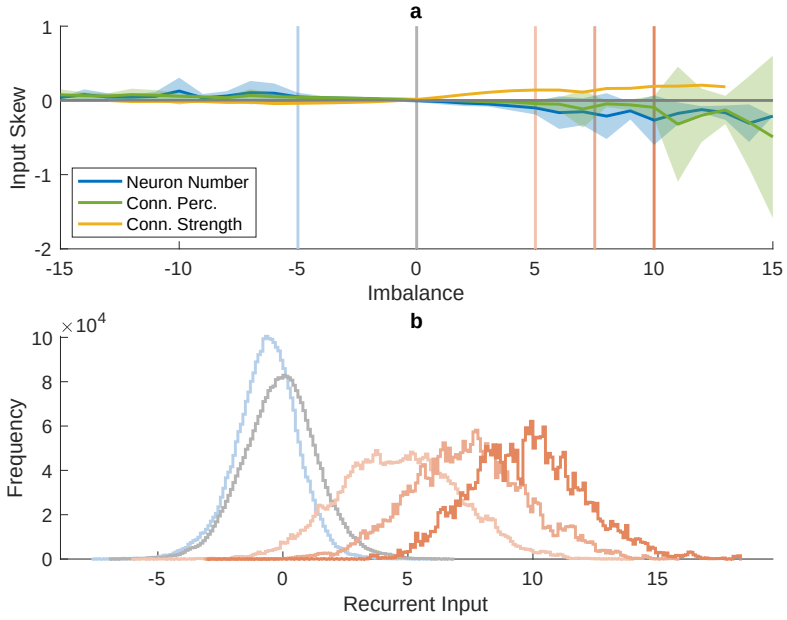

Fig. 8 Skew of the time-averaged recurrent input distributions. (a) shows the input skew as a function of the imbalance A when varying the neuron number (blue), connection percentage (green) and connection strength (yellow), as well as the corresponding MFT predictions (black) for $v_E N_E + v_I N_I = 2 \cdot 1.5^2$. 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

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

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

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

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 eqs. (18) and (28) into eqs. (29) and (30), we obtain the following equations for the μ- and Δ- nullclines, $\mathcal{N}_\mu$ and $\mathcal{N}_\Delta$:

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

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

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.

A.4.1 Nullclines

The μ-nullcline depends on the imbalance $m_E N_E + m_I N_I$. It is illustrated in fig. 9 for different values of imbalance ranging from -10 to 10. For large, positive imbalance (red curves in fig. 9) 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 fig. 9), 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 fig. 9 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). The variance 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 $\Delta_{max} = v_{In} U^2(t) + v_E N_E + v_I N_I$.

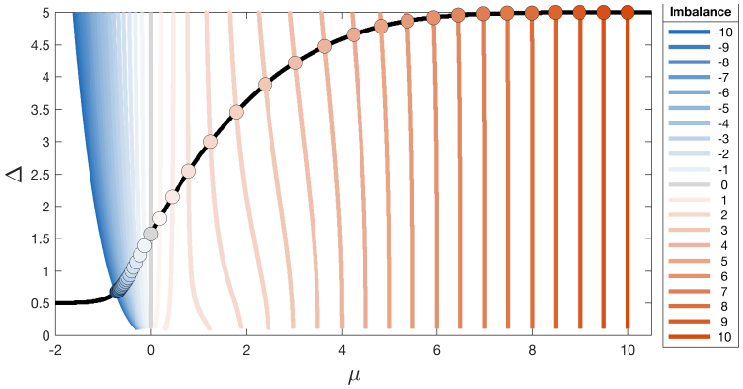

Fig. 9 Graphical representation of the μ and Δ nullclines. The μ nullcline (coloured) depends on the anatomical imbalance A, shown here for values ranging from -10 to 10. The Δ nullcline (black) depends on $v_E N_E + v_I N_I$, shown here in black for a single value $2 \cdot 1.5^2$, as used in the simulations. The solutions for each imbalance level, indicated by the coloured dots, are found at the intersection of the two nullclines

The solutions for μ (fig. 10a) and Δ (fig. 10b) are given by the intersection of the two nullclines. The obtained values are subsequently used to calculate the firing rate mean (fig. 10c) and variance (fig. 10d) using eqs. (29) and (30).

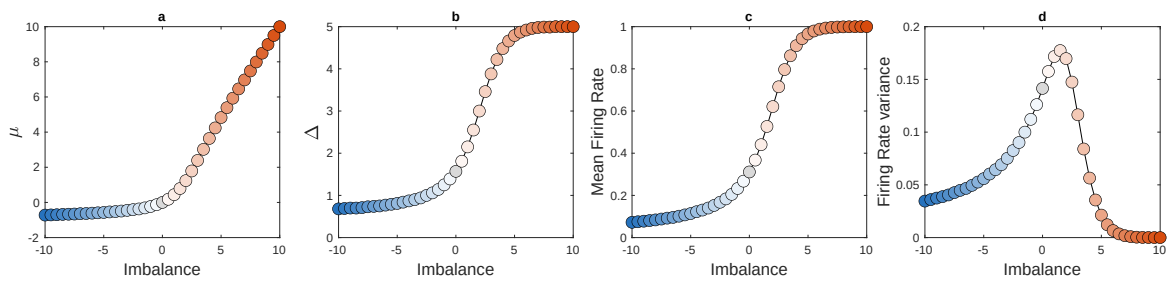

Fig. 10 Mean-field theory solutions found via the nullcline method for $v_E N_E + v_I N_I = 2 \cdot 1.5^2$ and a rectified hyperbolic tangent as activation function, as used in the simulations. (a) shows the mean recurrent input, (b) the recurrent input variance, (c) the mean firing rate, and (d) the firing rate variance as a function of the imbalance A

The mean-field theory equations we derived are valid for any transfer function, so our results can easily be applied to different networks. Using a sigmoid transfer function instead of the rectified hyperbolic

tangent, would have led to similar results (see fig. 11) The largest difference can be seen in the firing rate variance, which is lower when using a sigmoid (cf. panel D in figs. 10 and 11).

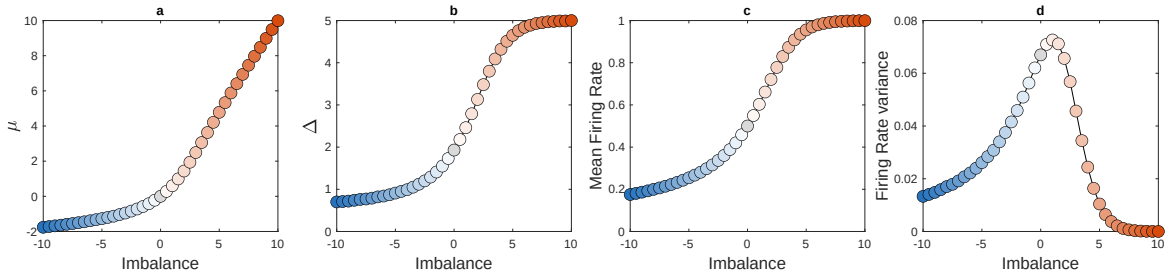

Fig. 11 Mean-field theory solutions found via the nullcline method for networks using a sigmoid activation function with $v_E N_E + v_I N_I = 2 \cdot 1.5^2$. (a) shows the mean recurrent input, (b) the recurrent input variance, (c) the mean firing rate, and (d) the firing rate variance as a function 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 $V_P^0 = \frac{g_P^2}{N_P p_P \left(1 - \frac{2}{\pi} p_P\right)}$. 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_{tot} by N_{tot} (with $N_{tot} = N_E + N_I$) matrix $\boldsymbol{J}$ whose first N_E columns correspond to the excitatory weights J_{ij}^E and whose last N_I columns correspond to the inhibitory weights J_{ij}^I. The system dynamics are then:

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

and the steady-state x_i^0 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_{tot}-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 $N_{tot} - 1$ eigenvalues distributed within a disk in the complex plane, and a single outlier

eigenvalue (Rajan and Abbott, 2006; Mastrogiuseppe and Ostojic, 2018), as shown in fig. 12a. 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 $\boldsymbol{S}$:

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

When this radius is larger than unity, the stationary state loses stability and the network has chaotic, fluctuating activity (Rajan and Abbott, 2006; Aljadeff et al., 2015; Harish and Hansel, 2015; Mastrogiuseppe and Ostojic, 2018). We find that this occurs when $g_{tot} = \sqrt{v_E N_E + v_I N_I} \geq 2.5$ (see fig. 12b), and exclude these networks from the results.

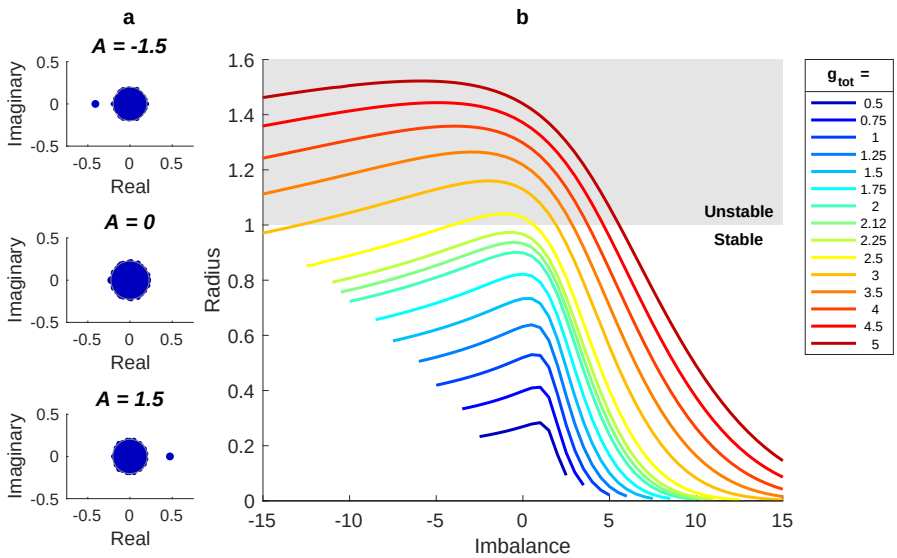

Fig. 12 Results of the local stability analysis. (a) shows three example eigenvalue spectra from simulated networks with total connection strength $g_{tot} = 0.5$ and imbalance $A = -1.5$, 0 and 1.5. (b) shows the MFT-derived radius of the eigenvalue spectrum of the stability matrix as a function of 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

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 eqs. (31) and (32).

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:

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

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 $\left[\mathcal{J}_{ij}^P\right] = 0$ in the above equation, to obtain the following differential equation for μ:

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

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

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

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 eq. (48) into eq. (49), we obtain:

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

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

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

Linearization

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

$$\tau \frac{d}{dt}\mu(t) = -\mu(t) + (m_E N_E + m_I N_I)\left[\phi(t)\right] = \mathcal{N}_\mu(\mu(t), \Delta(t)) \tag{52}$$

$$\tau \frac{d}{dt}\Delta = -\Delta(t) + (v_E N_E + v_I N_I)\left[\phi^2(t)\right] + v_{In} U(t)^2 = \mathcal{N}_\Delta(\mu(t), \Delta(t), U(t)) \tag{53}$$

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:

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

$$\frac{\partial \mathcal{N}_\mu}{\partial \mu} = -1 + (m_E N_E + m_I N_I)\frac{\partial}{\partial \mu}\left[\phi(t)\right] \tag{55}$$

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

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

For the variance, we get:

$$\frac{\partial \mathcal{N}_\mu}{\partial \Delta} = (m_E N_E + m_I N_I)\frac{\partial}{\partial \Delta}\left[\phi(t)\right] \tag{57}$$

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

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 into eq. (54), we obtain:

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

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

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

$$\frac{\partial \mathcal{N}_\Delta}{\partial \mu} = (v_E N_E + v_I N_I)\frac{\partial}{\partial \mu}\left[\phi^2(t)\right] \tag{62}$$

$$\frac{\partial}{\partial\mu}\left[\phi^2(t)\right] = \frac{\partial}{\partial\mu}\int Dz\phi^2(\mu(t)+\sqrt{\Delta(t)}z) = \int Dz2\phi(\mu_0+\sqrt{\Delta_0}z)\phi'(\mu_0+\sqrt{\Delta_0}z) = 2\left[\phi\phi'\right]_0 \tag{63}$$

$$\frac{\partial\mathcal{N}_\Delta}{\partial\Delta} = -1+(v_EN_E+v_IN_I)\frac{\partial}{\partial\Delta}\left[\phi^2(t)\right] \tag{64}$$

$$\frac{\partial}{\partial\Delta}\left[\phi^2(t)\right] = \frac{\partial}{\partial\Delta}\int Dz\phi^2(\mu(t)+\sqrt{\Delta(t)}z) = \int Dz2\frac{z}{2\sqrt{\Delta_0}}\phi(\mu_0+\sqrt{\Delta_0}z)\phi'(\mu_0+\sqrt{\Delta_0}z) \tag{65}$$

$$= \frac{1}{\sqrt{\Delta_0}}\int Dz\sqrt{\Delta_0}\phi'(\mu_0+\sqrt{\Delta_0}z)^2+\sqrt{\Delta_0}\phi(\mu_0+\sqrt{\Delta_0}z)\phi''(\mu_0+\sqrt{\Delta_0}z) \tag{66}$$

$$= \left[(\phi')^2\right]_0+\left[\phi\phi''\right]_0 \tag{67}$$

Substituting this into eq. (61), we obtain:

$$\tau\frac{d}{dt}\Delta = (v_EN_E+v_IN_I)2\left[\phi\phi'\right]_0\delta\mu(t)+\left(-1+(v_EN_E+v_IN_I)\left(\left[(\phi')^2\right]_0+\left[\phi\phi''\right]_0)\right)\right)\delta\Delta(t) \tag{68}$$

Combining eq. (60) and eq. (68), we obtain:

$$\tau\frac{d}{dt}\begin{pmatrix}\delta\mu\\ \delta\Delta\end{pmatrix} = -\begin{pmatrix}\delta\mu\\ \delta\Delta\end{pmatrix}+\begin{pmatrix}(m_EN_E+m_IN_I)\left[\phi'\right]_0 & \frac{1}{2}(m_EN_E+m_IN_I)\left[\phi''\right]_0\\ 2(v_EN_E+v_IN_I)\left[\phi\phi'\right]_0 & (v_EN_E+v_IN_I)\left(\left[(\phi')^2\right]_0+\left[\phi\phi''\right]_0)\right)\end{pmatrix}\begin{pmatrix}\delta\mu\\ \delta\Delta\end{pmatrix} \tag{69}$$

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

$$\boldsymbol{M} = \begin{pmatrix}(m_EN_E+m_IN_I)\left[\phi'\right]_0 & \frac{1}{2}(m_EN_E+m_IN_I)\left[\phi''\right]_0\\ 2(v_EN_E+v_IN_I)\left[\phi\phi'\right]_0 & (v_EN_E+v_IN_I)\left(\left[(\phi')^2\right]_0+\left[\phi\phi''\right]_0)\right)\end{pmatrix} \tag{70}$$

To simplify notations, we introduce:

$$a = (m_EN_E+m_IN_I)\left[\phi'\right]_0 \tag{71}$$

$$b = \frac{1}{2}(m_EN_E+m_IN_I)\left[\phi''\right]_0 \tag{72}$$

$$c = 2(v_EN_E+v_IN_I)\left[\phi\phi'\right]_0 \tag{73}$$

$$d = (v_EN_E+v_IN_I)\left(\left[(\phi')^2\right]_0+\left[\phi\phi''\right]_0)\right) \tag{74}$$

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

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

The discriminant of this polynomial is:

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

The roots λ_+,λ_- are thus given by:

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

These roots are illustrated in fig. 13. Note that, when $m_EN_E+m_IN_I$ 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 eq. (44)). In this case, the matrix $\boldsymbol{S}$ does not have an outlier eigenvalue (see

fig. 12a, 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 fig. 12a, bottom, and fig. 13). 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 fig. 12a, top and fig. 13). The real parts of both λ_+ and λ_- are always smaller than 1 (see fig. 14), so the system is always stable to global perturbations and its stability will only depend on the radius (local stability).

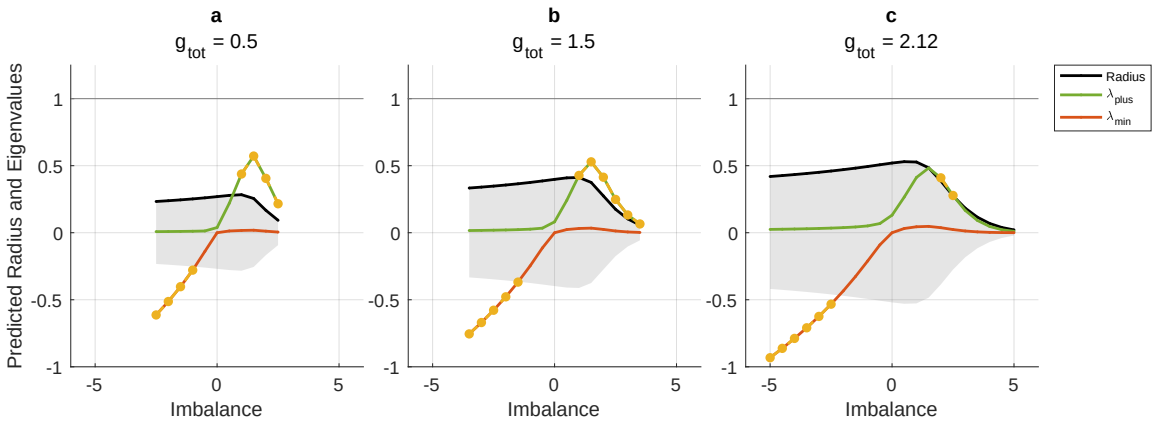

Fig. 13 Example results of the global stability analysis. The predictions of the radius and outliers of the global stability matrix are shown for three levels of total connection strength ($g_{tot} = \sqrt{v_E N_E + v_I N_I} = 0.5,\ 1, 5$, and $\sqrt{2 \cdot 1.5^2}$). Each panel shows the positive (λ_+, 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

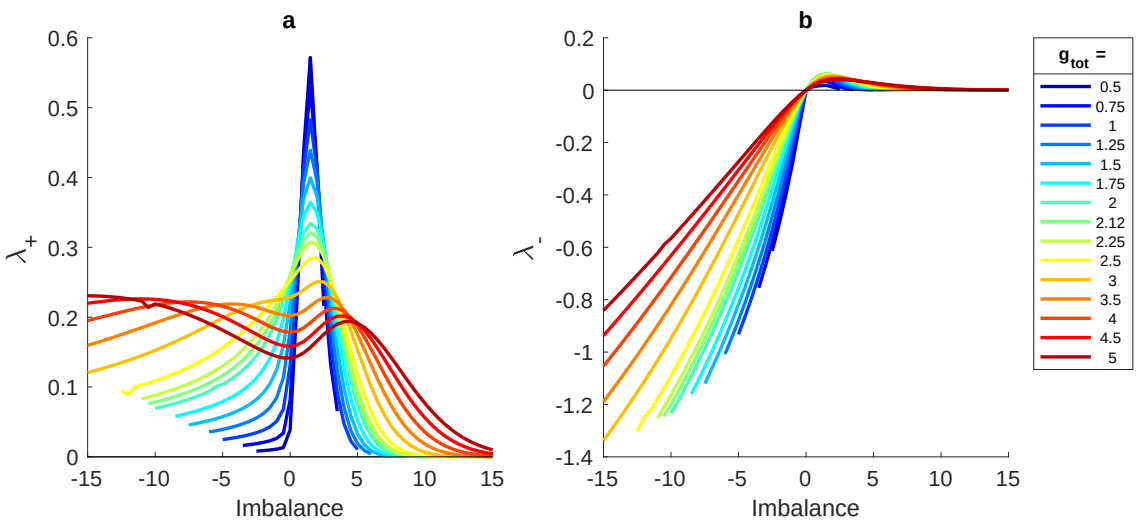

Fig. 14 Results of the global stability analysis. Mean-field theory derived values of the (a) positive (λ_+) and (b) negative (λ_-) eigenvalues of the global stability matrix $\boldsymbol{M}$ are shown as a function of the imbalance A, for various level of the total connection strength $g_{tot} = \sqrt{v_E N_E + v_I N_I}$

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:

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

Using integration by parts, we get:

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

Thus:

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

Appendix B 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 $\boldsymbol{\varphi}(t) = \phi(\boldsymbol{x}(t))$.

$$\boldsymbol{Y}(t) = \boldsymbol{W}(t) \cdot \boldsymbol{\varphi}(t) \tag{82}$$

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 $\boldsymbol{Y_{target}}(t)$.

$$\boldsymbol{\epsilon}(t) = \boldsymbol{Y}(t) - \boldsymbol{Y}_{target}(t) \tag{83}$$

The output weights $\boldsymbol{W}(t)$ were updated based on the current activity, the output error $\boldsymbol{\epsilon}$, and the $[N_{tot} \times N_{tot}]$ 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 and Abbott, 2009).

$$\boldsymbol{W}(t+\Delta t) = \boldsymbol{W}(t) - \epsilon(t)\frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t)}{1+\boldsymbol{\varphi}(t)^T\boldsymbol{P}(t)\boldsymbol{\varphi}(t)} \tag{84}$$

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

$$\boldsymbol{P}(t+\Delta t) = \boldsymbol{P}(t) - \frac{\boldsymbol{P}(t)\boldsymbol{\varphi}(t)\cdot(\boldsymbol{P}(t)\boldsymbol{\varphi}(t))^T}{1+\boldsymbol{\varphi}(t)^T\boldsymbol{P}(t)\boldsymbol{\varphi}(t)} \tag{85}$$

Appendix C 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, 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 (Winter, 2009; Hatze, 1980).

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:

$$f_{m,max} = Act(t) \cdot f_v \cdot f_{la} + f_{lp} \tag{86}$$

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:

$$f_v(v_m) = \begin{cases} \frac{c}{v_m+b} - a & \text{if } v_m \leq 0 \\ \frac{C}{v_m-B} + A & \text{if } v_m > 0 \end{cases} \tag{87}$$

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:

$$f_{la}(l_m) = exp\left[-(\frac{(l_m/l_{opt})^{k_1} - 1}{k_2})^{k_3}\right] \tag{88}$$

$$f_{lp}(l_m) = exp((l_m - l_{opt})/k_4) \tag{89}$$

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

$$T_{net} = \sum_{i=1}^{n} T_{max_i} \cdot Act_i \tag{90}$$

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`):

$$\min_{Act} \parallel Act \parallel^2 \tag{91}$$

Appendix D Supplementary Figures

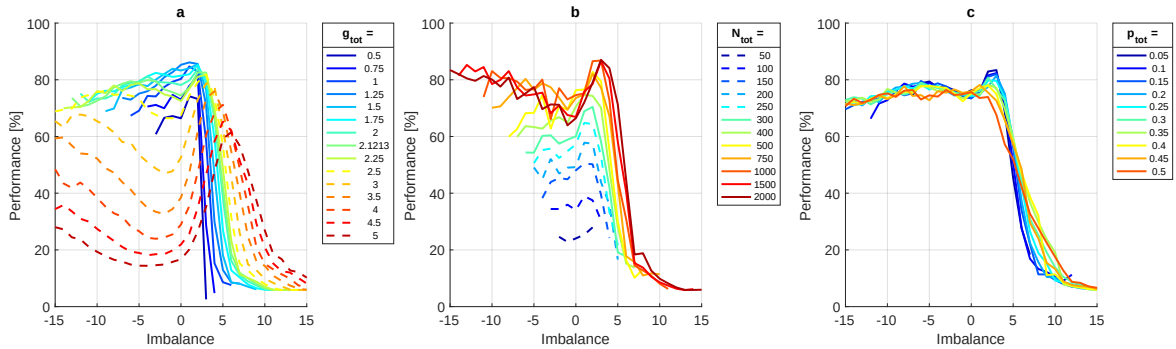

Fig. 15 Performance as a function of the imbalance A, for all investigated levels of (a) g_{tot}, (b) N_{tot}, and (c) p_{tot}. Dashed lines indicated networks that have been excluded from the main results (i.e., those with $N_{tot} \leq 250$ and $g_{tot} \geq 2.5$)

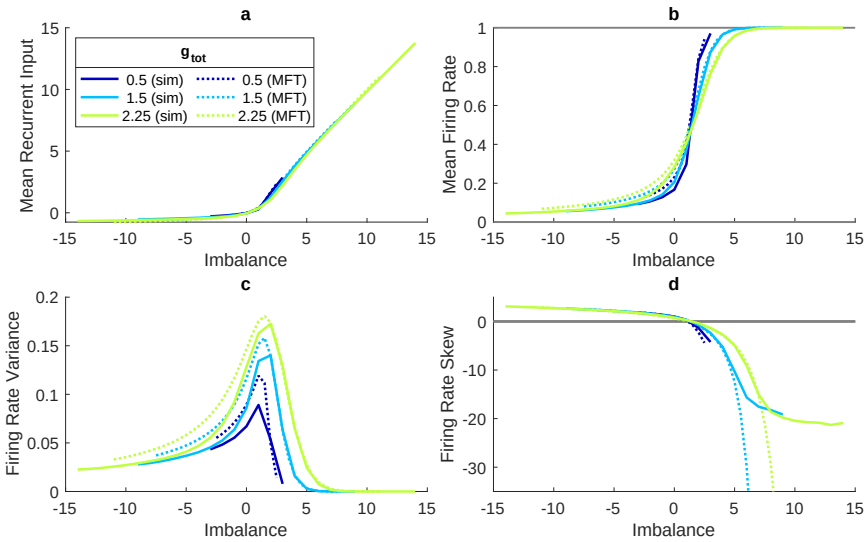

Fig. 16 Mean-field theory predictions (dotted)versus simulation results (solid) for three levels of total connection strength ($g_{tot} = 0.5$, 1.5, and $\sqrt{2 \cdot 1.5^2}$). Note that, for the simulations, results without output feedback are shown, as output feedback has not been taken into account in the MFT predictions

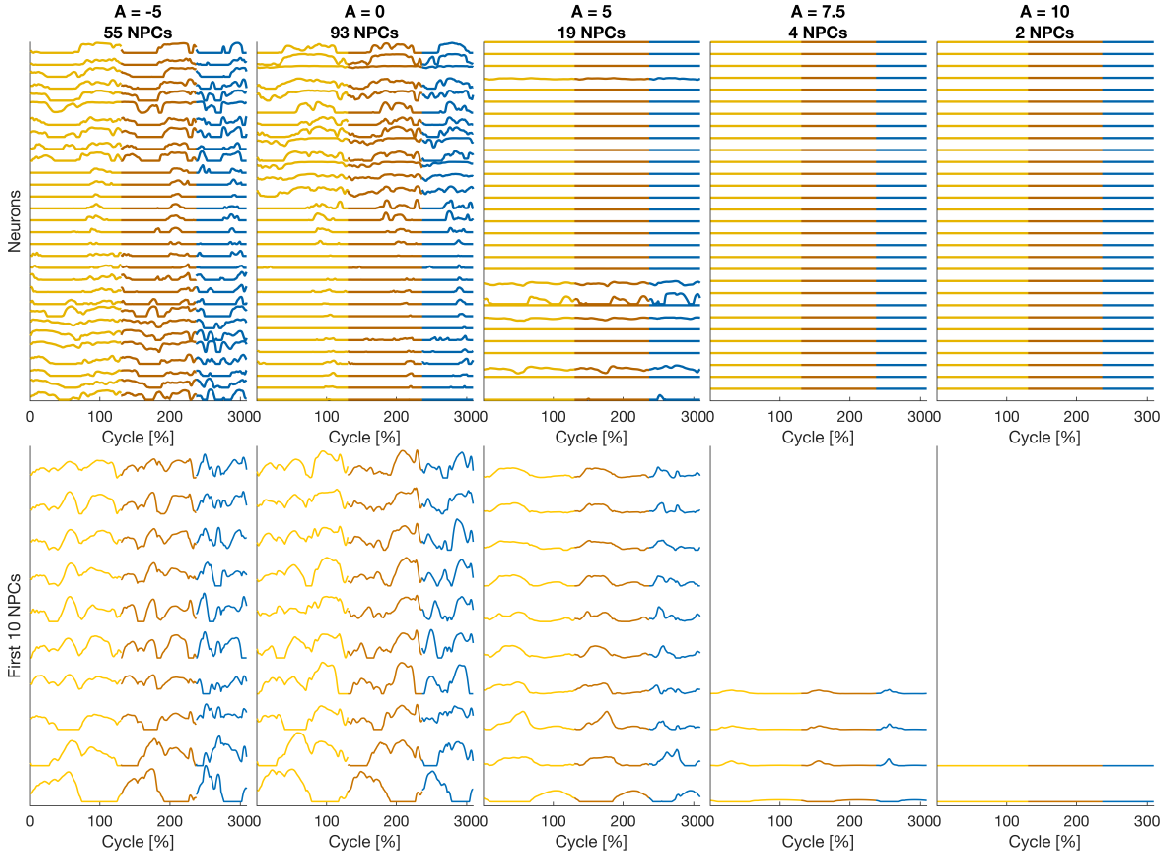

Fig. 17 Example neural firing rates (φ) and resulting network principal components (NPCs) for five example networks with imbalances of $A = -5, 0, 5, 7.5$ and 10 from left to right. The top row displays the firing rates for a selection of 30 out of 750 neurons. The bottom row shows up to ten NPCs resulting from the full set of neural activity patterns. The NPCs are ordered by their contribution to the overall variance, with those explaining the largest variance at the bottom. One example cycle for each of the three locomotor activities is illustrated, with, from left to right, walking slowly (yellow), walking fast (orange) and running (blue)

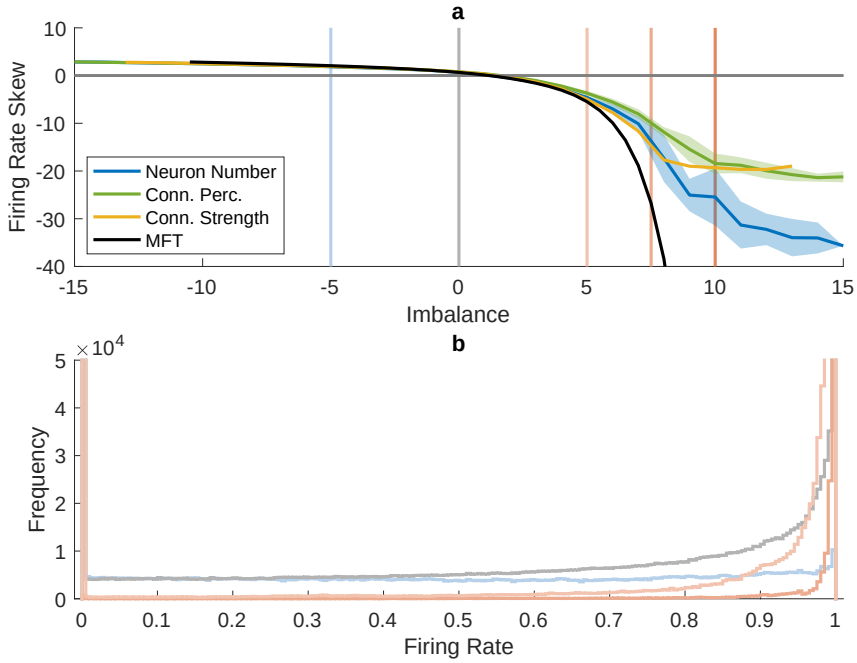

Fig. 18 Skew of the time-averaged firing rate distributions. (a) shows the firing rate skew as a function of the imbalance (A) when varying the neuron number (blue), connection percentage (green) and connection strength (yellow), as well as the corresponding MFT predictions for $g_{tot} = \sqrt{2 \cdot 1.5^2}$. 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

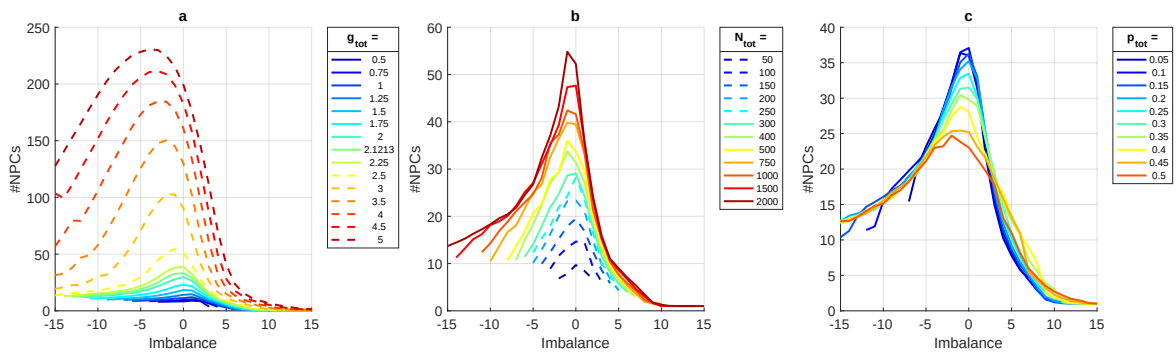

Fig. 19 Number of principle components (NPCs) as a function of the imbalance A, for all investigated levels of (a) g_{tot}, (b) N_{tot}, and (c) p_{tot}. Dashed lines indicated networks that have been excluded from the main results (i.e., those with $N_{tot} \leq 250$ and $g_{tot} \geq 2.5$)

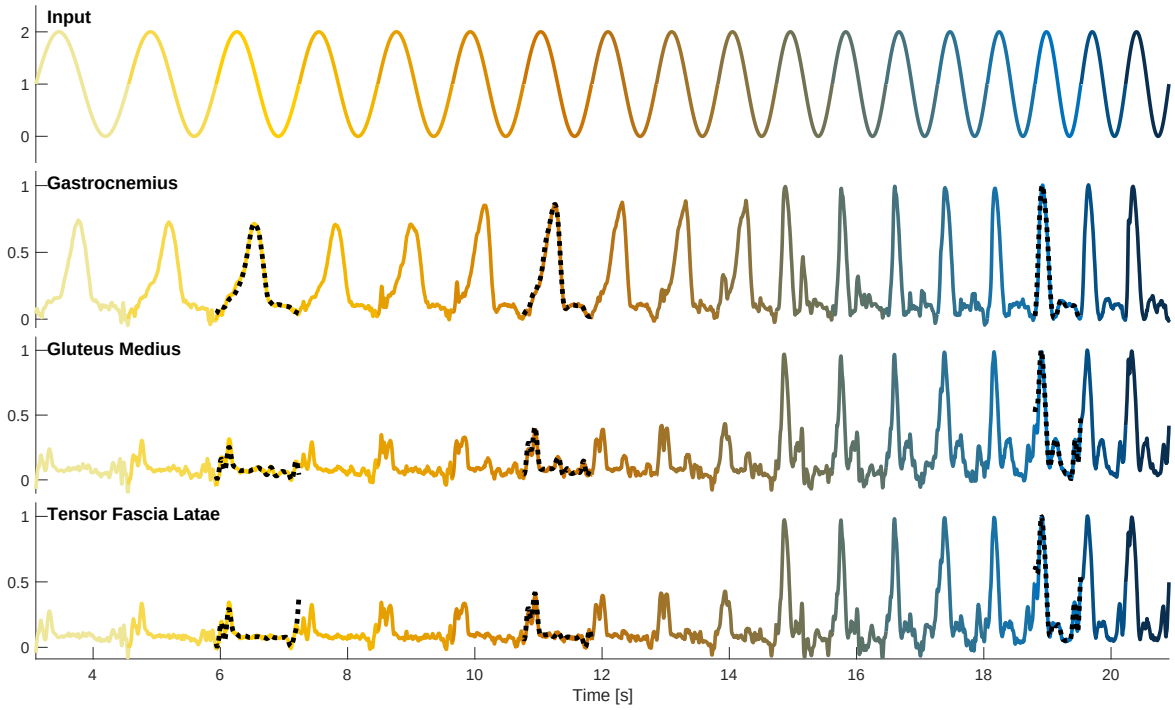

Fig. 20 Example outputs for a morphing task. An example network with imbalance $A = 0$ was trained normally and tested on a morphed signal, where each cycle had a higher frequency than the one before. The top row shows the input signal, the other rows show network outputs for three example muscles: m. gastrocnemius, m.gluteus medius and m. tensor fascia latae. The colour of the signals reflects the frequency of each cycle (with yellow for slow walking, orange for fast walking, and blue for running). Output targets for the trained cycles are overlaid as dotted black lines. Output targets for the morphed cycles were not available. The example network had $N_E = N_I = 375$, $p_E = p_I = 0.1$ and $g_E = g_I = 1.5$. Its performance on the standard test signal was 80.5%

References

Aoi, S., Ohashi, T., Bamba, R., Fujiki, S., Tamura, D., Funato, T., Senda, K., Ivanenko, Y., Tsuchiya, K.: Neuromusculoskeletal model that walks and runs across a speed range with a few motor control parameter changes based on the muscle synergy hypothesis. Sci. Rep. **9**(1), 369 (2019) https://doi.org/10.1038/s41598-018-37460-3

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

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

Bekius, A., Bach, M.M., Krogt, M.M., Vries, R., Buizer, A.I., Dominici, N.: Muscle Synergies During Walking in Children With Cerebral Palsy: A Systematic Review. Front. Physiol. **11**, 632 (2020) https://doi.org/10.3389/fphys.2020.00632

Beaulieu, C., Kisvarday, Z., Somogyi, P., Cynader, M., Cowey, A.: Quantitative Distribution of GABA-immunopositive and-immunonegative Neurons and Synapses in the Monkey Striate Cortex (Area 17).

Cereb. Cortex. **2**(4), 295–309 (1992) https://doi.org/10.1093/cercor/2.4.295

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

Brown, T.G.: The intrinsic factors in the act of progression in the mammal. Proc. R. Soc. Lond. Ser. B **84**(572), 308–319 (1911) https://doi.org/10.1098/rspb.1911.0077

Barroso, F.O., Torricelli, D., Moreno, J.C., Taylor, J., Gomez-Soriano, J., Bravo-Esteban, E., Piazza, S., Santos, C., Pons, J.L.: Shared muscle synergies in human walking and cycling. J. Neurophysiol. **112**(8), 1984–1998 (2014) https://doi.org/10.1152/jn.00220.2014

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

Boström, K.J., Wagner, H., Prieske, M., De Lussanet, M.: Model for a flexible motor memory based on a self-active recurrent neural network. Hum. Mov. Sci. **32**(5), 880–898 (2013) 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.: A genetically defined asymmetry underlies the inhibitory control of flexor-extensor locomotor movements. eLife **4**, 04718 (2015) https://doi.org/10.7554/eLife.04718

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

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

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

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

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

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

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

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

Chvatal, S.A., Ting, L.H.: Voluntary and Reactive Recruitment of Locomotor Muscle Synergies during Perturbed Walking. J. Neurosci. **32**(35), 12237–12250 (2012) https://doi.org/10.1523/JNEUROSCI.6344-11.2012

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

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

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

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

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

Dominici, N., Ivanenko, Y.P., Cappellini, G., d'Avella, A., Mondì, V., Cicchese, M., Fabiano, A., Silei, T., Di Paolo, A., Giannini, C., Poppele, R.E., Lacquaniti, F.: Locomotor Primitives in Newborn Babies and Their Development. Science **334**(6058), 997–999 (2011) https://doi.org/10.1126/science.1210617

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

Di Russo, A., Stanev, D., Sabnis, A., Danner, S.M., Ausborn, J., Armand, S., Ijspeert, A.: Investigating the roles of reflexes and central pattern generators in the control and modulation of human locomotion using a physiologically plausible neuromechanical model. J. Neural Eng. **20**(6), 066006 (2023) https://doi.org/10.1088/1741-2552/acfdcc

Dudek, F.E., Sutula, T.P.: Epileptogenesis in the dentate gyrus: a critical perspective. In: Prog. Brain Res. vol. 163, pp. 755–773. Elsevier, Amsterdam, The Netherlands (2007). https://doi.org/10.1016/S0079-6123(07)63041-6

Eccles, J.: From electrical to chemical transmission in the central nervous system. Notes Rec. R. Soc. **30**(2), 219–230 (1976) https://doi.org/10.1098/rsnr.1976.0015

Falconer, K.: Quantitative assessment of co-contraction at the ankle joint in walking. Electromyogr. Clin. Neurophysiol. **25**(2-3), 135–149 (1985)

Frigon, A., Akay, T., Prilutsky, B.I.: Control of Mammalian Locomotion by Somatosensory Feedback. Compr. Physiol. **12**(1), 2877–2947 (2021) https://doi.org/10.1002/cphy.c210020

Field, R.E., D'amour, J.A., Tremblay, R., Miehl, C., Rudy, B., Gjorgjieva, J., Froemke, R.C.: Heterosynaptic Plasticity Determines the Set Point for Cortical Excitatory-Inhibitory Balance. Neuron **106**(5), 842–8544 (2020) https://doi.org/10.1016/j.neuron.2020.03.002

Funahashi, K.-i., Nakamura, Y.: Approximation of dynamical systems by continuous time recurrent neural

networks. Neural Netw. **6**(6), 801–806 (1993) https://doi.org/10.1016/S0893-6080(05)80125-X

Fogarty, M.J.: Inhibitory Synaptic Influences on Developmental Motor Disorders. Int. J. Mol. Sci. **24**(8), 6962 (2023) https://doi.org/10.3390/ijms24086962

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

Grillner, S., Deliagina, T., El Manira, A., Hill, R.H., Orlovsky, G.N., Wallén, P., Ekeberg, Ö., Lansner, A.: Neural networks that co-ordinate locomotion and body orientation in lamprey. Trends Neurosci. **18**(6), 270–279 (1995) https://doi.org/10.1016/0166-2236(95)80008-P

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

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

Geyer, H., Herr, H.: A Muscle-Reflex Model That Encodes Principles of Legged Mechanics Produces Human Walking Dynamics and Muscle Activities. IEEE Trans. Neural Syst. Rehabil. Eng **18**(3), 263–273 (2010) https://doi.org/10.1109/TNSRE.2010.2047592

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

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

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

Grillner, S., Zangger, P.: The effect of dorsal root transection on the efferent motor pattern in the cat's hindlimb during locomotion. Acta Physiol. Scand. **120**(3), 393–405 (1984) https://doi.org/10.1111/j.1748-1716.1984.tb07400.x

Gui, K., Zhang, D.: Influence of locomotion speed on biomechanical subtask and muscle synergy. J. Electromyogr. Kinesiol. **30**, 209–215 (2016) https://doi.org/10.1016/j.jelekin.2016.07.010

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

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

Haykin, S.S.: Adaptive Filter Theory, Fifth, international edition edn. Pearson, Harlow, England (2014)

Hinaut, X., Dominey, P.F.: Real-Time Parallel Processing of Grammatical Structure in the Fronto-Striatal System: A Recurrent Network Simulation Study Using Reservoir Computing. PLoS One **8**(2), 52946

(2013) https://doi.org/10.1371/journal.pone.0052946

Hromádka, T., DeWeese, M.R., Zador, A.M.: Sparse Representation of Sounds in the Unanesthetized Auditory Cortex. PLOS Biol. **6**(1), 16 (2008) https://doi.org/10.1371/journal.pbio.0060016

Herzog, W.: Muscle. In: Nigg, B.M., Herzog, W. (eds.) Biomechanics of the Musculo-skeletal System. Wiley, Chichester, England (1999)

Harish, O., Hansel, D.: Asynchronous Rate Chaos in Spiking Neuronal Circuits. PLoS Comput. Biol. **11**(7), 1004266 (2015) https://doi.org/10.1371/journal.pcbi.1004266

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

Herstel, L.J., Wierenga, C.J.: Network control through coordinated inhibition. Curr. Opin. Neurobiol. **67**, 34–41 (2021) https://doi.org/10.1016/j.conb.2020.08.001

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

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

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

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

Jaeger, H.: The "echo state" approach to analysing and training recurrent neural networks – with an Erratum note. GMD Report 148, GMD - German National Research Institute for Computer Science (2001). http://www.faculty.jacobs-university.de/hjaeger/pubs/EchoStatesTechRep.pdf

Jankowska, E.: Spinal interneuronal networks in the cat: Elementary components. Brain Res. Rev. **57**(1), 46–55 (2008) 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.: Motor Signs Distinguish Children with High Functioning Autism and Asperger's Syndrome from Controls. J. Autism Dev. Disord. **36**(5), 613–621 (2006) https://doi.org/10.1007/s10803-006-0109-y

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

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

Kadmon, J., Sompolinsky, H.: Transition to Chaos in Random Neuronal Networks. Phys. Rev. X **5**(4),

041030 (2015) https://doi.org/10.1103/PhysRevX.5.041030

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

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

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

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

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

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

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

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

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

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

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

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

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

Markin, S.N., Lemay, M.A., Prilutsky, B.I., Rybak, I.A.: Motoneuronal and muscle synergies involved in

cat hindlimb control during fictive and real locomotion: a comparison study. J. Neurophysiol. **107**(8), 2057–2071 (2012) https://doi.org/10.1152/jn.00865.2011

McGowan, C.P., Neptune, R.R., Clark, D.J., Kautz, S.A.: Modular control of human walking: Adaptations to altered mechanical demands. J. Biomech. **43**(3), 412 (2010) https://doi.org/10.1016/j.jbiomech.2009.10.009

Mastrogiuseppe, F., Ostojic, S.: Intrinsically-generated fluctuating activity in excitatory-inhibitory networks. PLoS Comput. Biol. **13**(4), 1005498 (2017) https://doi.org/10.1371/journal.pcbi.1005498

Mastrogiuseppe, F., Ostojic, S.: Linking Connectivity, Dynamics, and Computations in Low-Rank Recurrent Neural Networks. Neuron **99**(3), 609–62329 (2018) https://doi.org/10.1016/j.neuron.2018.07.003

Meinecke, D.L., Peters, A.: GABA immunoreactive neurons in rat visual cortex. J. Comp. Neurol. **261**(3), 388–404 (1987) https://doi.org/10.1002/cne.902610305

McCrea, D.A., Rybak, I.A.: Modeling the mammalian locomotor CPG: insights from mistakes and perturbations. In: Cisek, P., Drew, T., Kalaska, J.F. (eds.) Prog. Brain Res. Computational Neurosci.: Theoretical Insights into Brain Function, vol. 165, pp. 235–253. Elsevier, Amsterdam, The Netlands (2007). https://doi.org/10.1016/S0079-6123(06)65015-2

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

Marom, S., Shahaf, G.: Development, learning and memory in large random networks of cortical neurons: lessons beyond anatomy. Q. Rev. Biophys. **35**(1), 63–87 (2002) https://doi.org/10.1017/S0033583501003742

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

Neptune, R.R., Clark, D.J., Kautz, S.A.: Modular Control of Human Walking: A Simulation Study. J. Biomech. **42**(9), 1282–1287 (2009) https://doi.org/10.1016/j.jbiomech.2009.03.009

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

O'Connor, D.H., Peron, S.P., Huber, D., Svoboda, K.: Neural Activity in Barrel Cortex Underlying Vibrissa-Based Object Localization in Mice. Neuron **67**(6), 1048–1061 (2010) 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., Brin, M.F., Fahn, S., Corey, D.P., Breakefield, X.O.: The TOR1A (DYT1) gene family and its role in early onset torsion dystonia. Genomics **62**(3), 377–384 (1999) https://doi.org/10.1006/geno.1999.6039

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

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

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

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.: Pathophysiology of Dyt1-Tor1a dystonia in mice is mediated by spinal neural circuit dysfunction. Sci. Transl. Med. **15**(694), 3904 (2023) https://doi.org/10.1126/scitranslmed.adg3904

Prinz, A.A., Thirumalai, V., Marder, E.: The Functional Consequences of Changes in the Strength and Duration of Synaptic Inputs to Oscillatory Neurons. J. Neurosci. **23**(3), 943–954 (2003) https://doi.org/10.1523/JNEUROSCI.23-03-00943.2003

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

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

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

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

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

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

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

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

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

Rothwell, J.C., Obeso, J.A., Day, B.L., Marsden, C.D.: Pathophysiology of dystonias. Adv. Neurol. **39**, 851–863 (1983)

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

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

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

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

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

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

Shao, Y., Dahmen, D., Recanatesi, S., Shea-Brown, E., Ostojic, S.: Identifying the impact of local connectivity patterns on dynamics in excitatory-inhibitory networks. arXiv (2024). https://doi.org/10.48550/arXiv.2411.06802

Sharrard, W.J.W.: The Segmental Innervation of the Lower Limb Muscles in Man. Ann. R. Coll. Surg. Engl. **35**(2), 106–122 (1964)

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

Shippen, J., May, B.: A kinematic approach to calculating ground reaction forces in dance. J. Dance Med. Sci. **16**(1), 39–43 (2012)

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

Shao, Y., Ostojic, S.: Relating local connectivity and global dynamics in recurrent excitatory-inhibitory networks. PLoS Comput. Biol. **19**(1), 1010855 (2023) https://doi.org/10.1371/journal.pcbi.1010855

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

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

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

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

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

Taub, E., Goldberg, I.A., Taub, P.: Deafferentation in monkeys: Pointing at a target without visual feedback. Exp. Neurol. **46**(1), 178–186 (1975) https://doi.org/10.1016/0014-4886(75)90040-0

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

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

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

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

Vogels, T.P., Rajan, K., Abbott, L.F.: Neural Network Dynamics. Annu. Rev. Neurosci. **28**(1), 357–376 (2005) https://doi.org/10.1146/annurev.neuro.28.061604.135637

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

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

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

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

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

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

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

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

Yaksh, T.L.: 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 (1989) 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.: Neocortical excitation/inhibition balance in information processing and social dysfunction. Nature **477**(7363), 171–178 (2011) https://doi.org/10.1038/nature10360

Zhang, J., Lanuza, G.M., Britz, O., Wang, Z., Siembab, V.C., Zhang, Y., Velasquez, T., Alvarez, F.J., Frank, E., Goulding, M.: V1 and v2b interneurons secure the alternating flexor-extensor motor activity mice require for limbed locomotion. Neuron **82**(1), 138–150 (2014) https://doi.org/10.1016/j.neuron.2014.02.013

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