



City Research Online

City St George's, University of London

Citation: Hauke, D. J., Rodriguez-Sanchez, J., Oloye, H., Berndt, L. C. S., Pinotsis, D., Friston, K. J., Mathalon, D. H. & Adams, R. A. (2026). A Canonical Microcircuit for Estimating Excitation/Inhibition (E/I) Balance. *Translational Psychiatry*,

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/37366/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

A Canonical Microcircuit for Estimating Excitation/Inhibition (E/I) Balance

Daniel J. Hauke¹, Julia Rodriguez-Sanchez^{1,2}, Hope Oloye¹, Lioba C. S. Berndt^{1,3,4},
Dimitris Pinotsis², Karl J. Friston⁵, Daniel H. Mathalon^{6,7}, & Rick A. Adams^{1,8}

Affiliations:

¹Hawkes Institute, Department of Computer Science, University College London, London, UK

²Department of Psychology and Neuroscience, City St George's, University of London, London, UK

³Department of Psychology, Faculty of Health and Life Sciences, University of Exeter, Exeter, UK

⁴School of Physiology, Pharmacology and Neuroscience, University of Bristol, Bristol, UK

⁵Department of Imaging Neuroscience, Queen Square Institute of Neurology, University College London, London, UK

⁶Department of Psychiatry & Behavioral Sciences, University of California, San Francisco, CA, USA

⁷Mental Health Service, San Francisco Veterans Affairs Health Care System, San Francisco, CA, USA

⁸Institute of Cognitive Neuroscience, University College London, London, UK

Corresponding author:

Daniel J. Hauke

90 High Holborn

London WC1V 6LJ, UK

d.hauke@ucl.ac.uk

Abstract word count: 192

Main text word count: 4976

Figure/Table count: 4/1

Abstract

Excitation/inhibition (E/I) balance is crucial for maintaining healthy brain function and can be disrupted in various neurological and psychiatric disorders. Despite its importance, there are few tools to study E/I balance non-invasively in humans. Here, we propose a canonical microcircuit model to estimate E/I balance from non-invasive magnetoencephalography (MEG), electroencephalography (EEG) or optically pumped magnetometers (OPM) recordings by parameterising global pyramidal and inhibitory cell excitability. We first establish that E/I parameters are identifiable and recoverable. We then explore the effects of these new parameters and their interaction with other parameters in a series of simulations. To highlight the clinical relevance of this new model, we simulate changes in E/I balance and their impact on event-related potentials (ERPs) derived from paired-click, passive and active oddball paradigms, which are among the most robust clinical biomarkers of schizophrenia. Our simulations show that a loss of pyramidal cell excitability can explain reduced ERP amplitudes across all three paradigms, mirroring empirical findings in schizophrenia. This method may serve as a computational assay for estimating synaptopathy and E/I balance from non-invasive electrophysiological recordings across various clinical conditions thereby advancing efforts to develop personalised interventions to restore E/I balance.

Keywords: E/I balance, dynamic causal modelling, predictive coding, biophysical modelling, EEG, MEG

Introduction

Our brain relies on a fine-tuned balance between excitation (E) and inhibition (I), referred to as E/I balance, to maintain healthy function. Even small increases in E/I can lead to runaway excitation, as observed in epileptic seizures¹. Conversely, excessively increased inhibition can lead to sedation - for example, after alcohol intake² - or even to loss of consciousness³. E/I balance can be manipulated pharmacologically in various ways through drugs acting on glutamatergic and gamma-aminobutyric-acid (GABA)-ergic targets, but also through psychotomimetics like ketamine⁴⁻⁶ and possibly psychedelics like psilocybin⁷ and lysergic acid diethylamide^{7,8}.

Disturbances of E/I balance have featured in aetiological theories of numerous neurological and psychiatric disorders like epilepsy⁹, autism¹⁰, depression¹¹, and schizophrenia¹²⁻¹⁴. In the case of schizophrenia, for instance, multiple research lines point towards E/I imbalance, including genetic evidence implicating α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) and N-methyl-D-aspartate (NMDA) receptors^{15,16}, and postmortem studies that suggest reduced dendritic spine density and interneuron deficits^{17,18}. Additionally, electrophysiological signals measured with electroencephalography (EEG) and magnetoencephalography (MEG) are changed in individuals at clinical high risk for psychosis and patients with schizophrenia. This includes reduced gamma-band activity¹⁹, mismatch negativity²⁰⁻²⁴ and P300 in oddball paradigms^{22,23,25,26}, and reduced event-related potentials (ERPs) in paired-click paradigms²⁷⁻²⁹. There are pharmacological studies that reproduce ERP reductions observed in schizophrenia with ketamine and GABAergic drugs suggesting potential links with E/I balance^{5,6,30-37}. However, the field needs new methods to estimate E/I balance *in vivo* from non-invasive recordings to identify the precise nature of circuit dysfunction underlying these M/EEG biomarkers.

Here, we propose a canonical microcircuit model specifically designed to estimate E and I cell function from non-invasive electrophysiological recordings with dynamic causal modelling (DCM). Unlike fMRI-based models of E/I balance (e.g.,³⁸) that are limited by the hemodynamic response function acting as a low-pass filter focusing on M/EEG allows us to exploit rich temporal information M/EEG provides. We simulate how changes in E/I parameters can explain ERP alterations observed in schizophrenia to highlight their predictive validity, and we systematically assess how these parameters interact with each other through a series of simulations. The aims of this paper are threefold: 1) to introduce this new method for studying E/I balance, 2) to demonstrate the model's utility for studying neuropsychiatric disorders using schizophrenia as an example, and 3) to provide a comprehensive resource for the scientific community by making the modelling code publicly available along with empirically-informed priors, for future investigations of E/I balance across various clinical conditions.

Methods

Canonical Microcircuit for estimating E/I balance

Our model is grounded in the DCM framework. This state-space modelling framework was originally introduced to model fMRI data^{39,40} and later extended to M/EEG⁴¹⁻⁴⁴ based on seminal work from Jansen and Rit^{45,46}. This modelling approach can be concisely summarised by two sets of equations (in standard form of dynamical systems):

$$\frac{dx}{dt} = f(x(t), u(t), \theta) \quad (1)$$

$$y(t) = g(x(t), \theta) + \epsilon \quad (2)$$

The state equation (Eq. 1) describes how system states x often simply referred to as “states” or “variables” in dynamic-systems notation (here voltage and current of neuronal populations) evolve in time as a function of states x at time t , a thalamic input into the system u at time t and model parameters θ (e.g., expressing synaptic connectivity between neuronal populations). The observer equation (Eq. 2) describes how changes in the hidden states x are mapped onto the measured signal y at time t (e.g., an EEG recording) through an observation model with parameters θ and a normally distributed observation error ϵ .

DCM for EEG comes in two forms: conductance-based and convolution-based models. The former models the membrane conductance of each neuronal population and specific ion channels (GABA, AMPA and NMDA receptors) based on Hodgkin and Huxley’s model of the squid giant axon⁴⁷ scaled from single cells to neuronal populations with the Morris and Lecar reduction^{48–50}. Conversely, convolution-based DCM models are based on Jansen and Rit^{45,46} and model the conversion of incoming spikes to postsynaptic potentials (PSPs) through convolution with a synaptic response kernel^{41–44,51,52}. The E/I model we propose generalises across both convolution-based and conductance-based models, but we focus on the convolution-based model to introduce the E/I parameters. Next, we will unpack the modelling approach across different spatial scales.

Microscale Model: Neuronal Population

At the microscale, we model a neuronal population through two basic operations (Figure 1). First, at dendritic spines, incoming spikes are converted into PSPs through convolution with a synaptic response function:

$$h(t) = \begin{cases} \frac{1}{\tau} t e^{-\frac{1}{\tau}t}, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (3)$$

where τ is the synaptic time constant, with larger τ leading to more prolonged responses in time (Figure 1).

Secondly, PSPs V_{in} are converted into an average spike rate σ at the axon hillock using a sigmoid function:

$$\sigma = \frac{1}{1 + e^{-S \cdot V_{in}}} - V_0 \quad (4)$$

where V_0 indicates the baseline firing rate, V_{in} is average membrane potential of the neuronal population, and S parametrizes the slope of the sigmoid function and captures the dispersion of firing thresholds within this neuronal population, with smaller values of S resulting in a shallower slope (i.e., more dispersed thresholds in the population, Figure 1). Each population can be modelled through the convolution of these two functions:

$$V = h \otimes \sigma \quad (5)$$

which results in a system of two coupled ordinary differential equations for each neuronal population, the first describing the change in voltage and the second describing the change in current:

$$\begin{aligned} \dot{v} &= i \\ \dot{i} &= f(\text{Input}) - \frac{2}{\tau} \dot{v} - \frac{1}{\tau^2} v \end{aligned} \quad (6)$$

The input is determined by intrinsic (within-source) and extrinsic (between-source) connectivity.

Mesoscale Model: Canonical Microcircuit

At the mesoscale, we model a cortical column using the canonical microcircuit for predictive coding comprising inhibitory interneurons, spiny stellate cells, as well as superficial and deep pyramidal cells, based on previous work⁵³⁻⁵⁶. We connect these neuronal populations in the same cortical source based on Shaw et al.⁵⁷ (Figure 1), which was previously applied to model data from patients with schizophrenia by Adams et al.⁵⁸ (for model comparisons with the previous model and model recovery analyses, please see Supplement).

Importantly, we parameterise the self-connection of E (superficial pyramidal cells) and I cells (inhibitory interneurons), which we refer to as g_{ee} and g_{ii} , where the first subscript indicates the origin of the connection (e for the excitatory and i for the inhibitory population) and the second subscript specifies the target population. We focus on superficial pyramidal cells based on evidence suggesting that the excitability of superficial pyramidal cells is reduced in schizophrenia⁵⁸. These parameters determine the excitability of E and I cells and will change their average output firing rate. In this study, we operationalise E/I ratio as the ratio of average output firing rate of E and I cells that can be manipulated through changing these two parameters. Increasing g_{ee} , for example, corresponds to increasing the connection strength of the (inhibitory) self-connection, which models a loss of E cell excitability (or reduced E/I ratio). Conversely, increasing g_{ii} models a loss of I cell excitability resulting in net disinhibition of the microcircuit. Both self-connections are always inhibitory to ensure that the model generates stable dynamics. We can generalise this approach to the conductance-based model by parameterising the (self-inhibitory) GABAergic connections of the same neuronal populations.

We make two important modelling choices regarding these E/I parameters: 1) we constrain these parameters to be the same across all modelled regions in order to ensure an efficient global estimate of E and I cell excitability and 2) we allow them to be both general and condition-specific, expressing the average E/I balance across all conditions and the difference in E/I balance between conditions, respectively.

First, whilst there may be regional variation in E and I, we assume that genetic predisposition to altered E/I is generalised across cortex, rather than confined to specific areas (e.g., based on⁵⁹⁻⁶² in the case of schizophrenia-related changes). This is reflected in our cortex-wide parameter with the added potential benefit of using all the EEG data to estimate it. Note, however, that this assumption can be relaxed if one wishes to model gradients of E/I balance across the brain either by relaxing the global equality constraint or by adding a multiplicative weight of the global parameter (e.g., based on a receptor density map).

Second, we introduce these E and I parameters both as condition-specific parameters (capturing condition-specific rebalancing of E/I ratio in response to experimental inputs), referred to as $B^{g_{ee}}$ and $B^{g_{ii}}$ in the remainder of the paper, and condition-independent parameters expressing the average E/I balance, referred to as g_{ee} and g_{ii} . This is motivated by the notion that E/I changes are adaptive and may well have more impact on some EEG conditions, e.g. those that typically generate a much larger response (like oddballs). To summarise, the key differences between this model and Shaw et al.'s model⁵⁷ are 1) imposing equality constraints on intrinsic self-connections of E and I cells g_{ee} and g_{ii} and 2) introducing condition-specific modulation of self-connections of E and I cells $B^{g_{ee}}$ and $B^{g_{ii}}$.

Macroscale Model: Network

Lastly, we connect different sources with excitatory forward connections A^F that originate from superficial pyramidal cells in one source and project to spiny stellate cells and deep pyramidal cells of another source, and inhibitory backward connections A^B which originate from deep pyramidal cell populations and project to superficial pyramidal cells and inhibitory interneurons^{55,56}. This results in the following system of coupled differential equations for each source. Note that these coupled equations have the form of Eq. (6), but for consistency with previous work and the provided code, we denote all hidden states (voltage and current) with x . For spiny stellate cells, we have:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{1}{\tau_{ss}} \left(A^F \cdot \sigma(x_3) - g_2 \cdot \sigma(x_5) - g_1 \cdot \sigma(x_1) + U - 2\dot{x}_2 - \frac{x_1}{\tau_{ss}} \right) \end{aligned} \quad (7)$$

For E cells (superficial pyramidal cells), we have:

$$\begin{aligned} \dot{x}_3 &= x_4 \\ \dot{x}_4 &= \frac{1}{\tau_{sp}} \left(-A^B \cdot \sigma(x_7) - g_{11} \cdot \sigma(x_5) + g_7 \cdot \sigma(x_1) - g_6 \cdot \sigma(x_3) - 2\dot{x}_4 - \frac{x_3}{\tau_{sp}} \right) \end{aligned} \quad (8)$$

For I cells (inhibitory interneurons), we have:

$$\begin{aligned} \dot{x}_5 &= x_6 \\ \dot{x}_6 &= \frac{1}{\tau_{ii}} \left(-A^B \cdot \sigma(x_7) - g_3 \cdot \sigma(x_5) + g_4 \cdot \sigma(x_1) + g_{10} \cdot \sigma(x_3) + g_5 \cdot \sigma(x_7) - 2\dot{x}_6 - \frac{x_5}{\tau_{ii}} \right) \end{aligned} \quad (9)$$

And lastly for deep pyramidal cells, we have:

$$\begin{aligned} \dot{x}_7 &= x_8 \\ \dot{x}_8 &= \frac{1}{\tau_{dp}} \left(A^F \cdot \sigma(x_3) - g_8 \cdot \sigma(x_5) + g_{12} \cdot \sigma(x_3) - g_9 \cdot \sigma(x_7) - 2\dot{x}_8 - \frac{x_7}{\tau_{dp}} \right) \end{aligned} \quad (10)$$

We implemented this model using SPM12 (version 7771, Wellcome Trust Centre for Neuroimaging, London, UK, <https://github.com/spm/spm12/releases/tag/r7771>) in MATLAB (version R2020b; The MathWorks Inc., Natick, MA, USA; <https://www.mathworks.com>). The code for both the convolution- and conductance-based E/I models can be found on GitHub (https://github.com/daniel-hauke/dcm_ei).

Simulations

To illustrate the effects of the E and I parameters, we conduct a series of simulations based on parameter estimates from data collected from a large number of healthy volunteers during three different paradigms as part of the Bipolar and Schizophrenia Network for Intermediate Phenotypes (B-SNIP)^{62,63} and the second North American Prodrome Longitudinal Study (NAPLS2)^{64–66}. We briefly recapitulate the most pertinent information about the paradigms and preprocessing below. For more detailed information, please see the referenced studies and the Supplement.

Participants

We analysed EEG data from $n=266$ (paired-click paradigm), $n=241$ (passive oddball paradigm), and $n=237$ (active oddball paradigm) healthy participants. The EEG setup and preprocessing is reported in the Supplement. To provide priors for future analyses – that can be used beyond investigations focused on schizophrenia – we computed the grand mean data from healthy control participants. We ground our simulations in these denoised grand mean ‘participants’ to illustrate the effects of changing E and I parameters.

Tasks

Paired-click paradigm

Participants were presented with 150 binaural broadband auditory click pairs (4ms, 75dB, inter-click interval: 500ms, average interstimulus interval: 9.5s)⁶⁷.

Passive oddball paradigm

Participants performed a visual oddball primary task while being presented with 1794 tones that were either standard tones (85%, 633Hz, 50ms), duration deviants (5%, 633Hz, 100ms), pitch deviants (5%, 1000Hz, 50ms) or duration+pitch double-deviants (5%, 1000Hz, 100ms). Tone length included 5ms rise/fall time. Subjects were told to ignore the tones and focus on the visual task. Here, we focused on modelling the double-deviant tone as it showed the largest clinical effects in terms of predicting conversion to psychosis in individuals at clinical high risk for psychosis⁶⁶.

Active oddball paradigm

In the active oddball paradigm, participants were presented with 450 stimuli including 80% standard tones (500Hz, 50ms with 5ms rise/fall time), 10% target tones (1000Hz, 50ms, 5ms rise/fall time) and 10% novel sounds with a stimulus onset asynchrony of 1250ms⁶⁵. Novel sounds were various natural and man-made sounds⁶⁸. Participants were asked to respond to target tones by pressing a button.

Networks

We defined three task-specific networks to be modelled with our E/I model.

Paired-click paradigm

Since we are unaware of previous studies that used DCM to model paired-click responses, we defined a task network based on previous literature^{69,70} comprising left (MNI: [-53 -20 13]) and right superior temporal gyrus (STG; [49 -10 10]) as auditory input regions, left ([-36 23 3]) and right insular cortex (Ins; [36 23 3]), and left ([-1 -20 67]) and right medial frontal gyrus (MFG; [1 -20 67]). Note that primary auditory cortex (A1) is modelled implicitly through the exogenous input to limit the number of model parameters. To confirm that these sources explain sufficient variance in the data, we conducted a multiple sparse prior (MSP) source analysis with 16mm spheres drawn around these coordinates and compared it to an (unconstrained) independent and identically distributed (IID) source analysis in SPM. These six sources explained 95.06% of the variance versus 97.18% in the IID case. Formal model comparison based on the model evidence suggested that there was strong evidence that the 6-region source reconstruction was the better model (log Bayes factor: 23.6)⁷¹.

Passive oddball paradigm

We modelled the passive oddball paradigm with a well-established 6-region network comprising bilateral A1 (left: [-42 -22 7], right: [46 14 8]) as input regions, STG (left: [-61 -32 8], right: [59 -25 8]) and inferior frontal gyrus (IFG; left: [-46 20 8], right: [46 20 8]) based on extensive previous work^{21,72-77} including various DCM studies^{58,78-80}.

Active oddball paradigm

We defined a 6-region network for the active oddball paradigm based on a previous meta-analysis of 75 oddball paradigm PET/fMRI studies⁸¹. We chose STG as input regions (left: [-61 -32 8], right: [59 -25 8]) using the same coordinates as for the passive oddball task network and confirmed that they fell into the brain region activated during the active oddball task⁸¹. Importantly, in contrast to the passive oddball paradigm, the active oddball paradigm requires participants to maintain top-down attentional control towards the deviant tones to respond with button presses⁸². We thus modelled bilateral inferior frontal junction (IFJ; left: [-56 7 29], right: [50 8 30]) and intraparietal sulcus (IPS; left: [-33 -42 64], right: [33 -42 64]) regions of the dorsal attention network based on Kim⁸¹. We found that this 6-region model explained 94.31% of the variance versus 95.14% of an IID source reconstruction and was strongly favoured in model comparison (log Bayes factor: 49.2).

Integration

We integrated the convolution-based model with a Euler integrator⁸³, which we found to have the best accuracy-runtime trade-off for our purposes (see ⁸³ for a runtime comparison). We also provide code that works with the default SPM12 integrators and an implementation using MATLABs dde23 integrator for benchmarking the conductance-based model (Supplement).

Priors

We used the default SPM12 priors with the following changes. In the source model, we implemented equality constraints on transmission delays (extrinsic and intrinsic) and intrinsic connectivity parameters across regions to drastically reduce the number of free parameters by 40 (20→2 delay and 24→2 intrinsic connectivity parameters) for a six-region model. We used the equivalent current dipole (ECD) forward model and assumed that only E cells (superficial pyramidal cells) contributed to the EEG signal (fixing an additional 2 parameters), although note that this assumption can be easily relaxed.

The neuronal time constants τ of the original canonical microcircuit model were selected to model gamma oscillations⁵⁶. We found that, under the new integration schemes, the default priors on τ were not suitable for generating the oscillations that we wished to model – especially for the active oddball task, where the characteristic responses are on slower (theta) time scales. In simulations, we also observed that the slope of the sigmoid output firing rate function S — parametrising the dispersion of firing thresholds or depolarisation of a neuronal population⁸⁴ — was crucial in ensuring that the model generated dampened as opposed to ongoing or increasing oscillations. We thus performed a grid search over different combinations of the four time constants and S spanning the following grids τ_{ss} : [2 16 32 64] ms, τ_{ii} : [2 16 32] ms, τ_{sp} : [2 16 32 64 128] ms, τ_{dp} : [2 16 32 64 128] ms and S : [-2 -1 0 1 2]. We identified the best model based on the free energy approximation to the log model evidence, while constraining the solutions to physiological plausible parameter ranges, which we defined as [1 60] ms for spiny stellate cells^{85–88}, [1 30] ms for inhibitory interneurons^{89,90} and [1 200] ms for pyramidal cells. The τ range for pyramidal cells was motivated by evidence that NMDA receptors on these neurons can have time constants of around 40-200 ms^{91,92}, and these receptors may be captured implicitly in the convolution-based model. All priors are summarised in Table 1.

Model diagnostics

To assess model fit (i.e., accuracy), we computed the Pearson correlation between the responses predicted by the model and the measured EEG responses. To assess parameter identifiability, we fit the model to 100 randomly selected healthy participants using the posteriors from the grand mean as empirical priors. We report the Pearson correlation between E/I parameters across participants to test whether parameters are identifiable (i.e., their posterior dependencies or correlations are small). In line with previous work⁹³, we consider $r < |0.6|$ to be an acceptable correlation, suggesting that the parameters mediate distinct processes. To measure parameter recovery, we (i) simulated synthetic data using the posterior parameters from models fit to 100 randomly selected participants, (ii) added empirical levels of noise, estimated from the data (i.e., using the posterior estimate saved in DCM.Ce), and (iii) reinverted the models on the simulated data. We report intraclass correlation coefficients (ICC), as recommended previously^{94,95}, and consider recovery to be poor (<0.4), fair (0.4–0.59), good (0.6–0.74), or excellent (>0.75) as suggested by Fleiss⁹⁶. We also report Pearson correlations between simulated and recovered parameters in line with previous work^{93,97}.

Parameter sensitivity analysis

To illustrate the effects of the E and I parameters, we simulated ERPs starting with their posterior and incrementally increased or decreased the parameters over the interval [-0.5 0.5] in steps of 0.125, which corresponds to scaling the connectivity strength of the self-connection by $\pm 50\%$, while keeping all other parameters fixed to their posterior estimate. We also varied multiple parameters together. To

simplify the visualisation for multidimensional simulations, we display the amplitude of characteristic difference wave ERP components. Specifically, we defined the P2 component as the most positive deflection in Cz between 120-250ms in the paired-click paradigm, the MMN, as the most negative deflection in Fz between 150-250ms in the passive oddball paradigm, and the P3b component as the most positive deflection in Pz between 250-600ms in the active oddball paradigm.

Results

Model diagnostics

The model fit all paradigms remarkably well. Specifically, it predicted responses in the paired stimulus paradigm with an average $r=0.95$ (tone 1: $r=0.95$; tone 2: $r=0.96$; Figure 2a), in the passive oddball paradigm with $r=0.89$ (standard: $r=0.85$; deviant: $r=0.94$; Figure 2d) and the active oddball paradigm with $r=0.84$ (standard: $r=0.71$; target: $r=0.97$; Figure 2g). Note that ERPs tend to be much larger in the deviant and target conditions of the oddball paradigm compared to the standard tone condition, while the noise is comparable between conditions. This results in reduced SNR in the standard condition and likely explains why the model shows reduced fits for this condition. We observed good parameter identifiability (all $r<|0.6|$) across all paradigms; Figure 2b, 2e and 2h). Parameter recovery of B^{gee} was excellent, while B^{gii} recovery was fair across all three paradigms (Figure 2c, 2f, 2i) and comparable to other findings⁹⁸.

Posterior parameter estimates

Posterior parameter estimates are reported in Supplementary Tables 1-3 and provided in SPM12 DCM format on GitHub (https://github.com/daniel-hauke/dcm_ei). We recommend using these parameters as empirical priors for future DCM analyses with this model and these paradigms.

Simulating changes in E and I cell excitability

To illustrate the validity of the model, we simulated a loss of condition-specific E or I cell excitability across all three paradigms and compared those results to known effects in the schizophrenia literature (Figure 3). We found that increasing B^{gee} to simulate a loss of pyramidal cell excitability reduced N1 and P2 components of the difference waveform in the paired click paradigm by reducing the response to the first tone and increasing the response to the second tone (Figure 3a), which may explain reduced S2/S1 ratios observed in patients²⁷⁻²⁹. Another finding in the schizophrenia literature is that patients show reduced responses to both tones²⁷, which can be simulated with our model by increasing g_{ee} (reducing average excitability across both conditions, Figure S1). Interestingly, simulating a loss of inhibitory cell excitability by increasing B^{gii} had no effect on the waveform, suggesting that this parameter alone cannot explain the changes observed in schizophrenia (Figure 1b). Increasing I-to-E coupling (g_{11}) did not change this result. Note, however, that the effect of changing B^{gii} depends on the other parameters (see Figure S3 for illustrations of single participants).

In the passive oddball paradigm, we found that a loss of E cell excitability ($\uparrow B^{gee}$) resulted in reduced ERP amplitudes in responses to deviant tones and an overall reduced difference waveform with reduced mismatch negativity and P3a responses (Figure 3c), both observed in schizophrenia^{20,24-26,65,66}. A similar change in I cell excitability resulted in no comparable change in the ERP (Figure 3d) nor did it when I-to-E coupling (g_{11}) was increased.

For the active oddball paradigm, we observed distinct effects of changing E and I cell excitability. While reduced E/I ratio - either through a loss of E cell excitability ($\uparrow B^{gee}$) or hyperexcitable I cells ($\downarrow B^{gii}$) - resulted in reduced P3b in the target-standard difference waveform (Figure 3e), changing E cell excitability affected the ERPs in response to both standard and target tones, whereas changing I cell excitability only affected the response to target tones (Figure 3f). To investigate whether this qualitative distinction would also hold in the passive oddball paradigm, we increased the simulation range for I cell excitability to $\pm 200\%$ (not shown) and indeed observed similar

changes (i.e., much stronger effects on the deviant- compared to the standard-ERP). This finding is in line with P3b reductions observed in schizophrenia patients^{25,26} and suggests that the model can potentially distinguish E and I cell dysfunction because they relate to different data features.

Interactions between E and I cell excitability

Next, we investigated whether E and I cell excitability interacted, by systematically varying both parameters together. We found that the effect of changing E cell excitability on the ERP amplitude was largely independent of varying I cell excitability in the paired-click paradigm, as indicated by the straight vertical line in Figure 4a, which suggests that changing E cell excitability has the same effect on the P2 component irrespective of assumed I cell excitability. We observed a similar result for the MMN component in the passive oddball paradigm (Figure 4b). However, we found that this was primarily driven by the effect of changing E cell excitability on the deviant response for all levels of I cell excitability. There was no effect on the standard-ERP. Lastly, we found non-trivial interactions for the P3b component: reducing the E/I ratio through simulating both a loss of E cell excitability ($\uparrow B^{gee}$) and an increase in I cell excitability ($\downarrow B^{gii}$) resulted in the most pronounced effects on P3b components of the target-ERP and the difference waveform, whereas the effects of changing E cell excitability on the standard-ERP were largely independent of I cell excitability (Figure 4c). This suggests that P3b amplitudes could reflect changes in both parameters rather than E or I cell function alone (see Supplement for more detailed investigations of parameter interactions for this paradigm).

Discussion

This study set out to introduce a canonical microcircuit model to estimate E/I balance from non-invasive M/EEG recordings by parameterising global excitatory and inhibitory cell function. We illustrated how this model can be used to study condition-specific changes in E/I balance by applying it to three different auditory paradigms: paired-click, active oddball and passive oddball. We found that E/I parameters can be recovered efficiently using a variety of empirical paradigms in healthy participants. We then illustrated— with numerical studies—that a loss of pyramidal cell excitability can account for many known empirical findings in the schizophrenia M/EEG literature, including reduced ERP amplitudes across all three paradigms. Lastly, we provide the requisite code, integration schemes and empirical priors for future analyses of data from clinical cohorts.

We found excellent parameter recovery for the parameter capturing pyramidal cell excitability, rendering it a promising and most importantly, biologically interpretable biomarker. This parameter could capture multiple biological processes that change pyramidal cell excitability, for example, NMDA receptor function on pyramidal cells as suggested previously⁵⁸, or neuromodulatory effects or loss of synapses within the superficial pyramidal cell pool. This parameter should thus be interpreted with caution and future studies are needed to carefully evaluate the contributions of these processes to the E/I parameters. Possible avenues for validation include pharmacological studies administering GABAergic and glutamatergic agents, magnetic resonance spectroscopy studies in humans, human-derived induced pluripotent stem cell (iPSC) models and animal studies using opto- and chemogenetic manipulations. Alternatively, conductance-based DCMs can be used, which contain explicit parameterisations of NMDA receptors and thus may assess their function more directly; these models can be more challenging to fit, however. Interestingly, our simulations suggest that a loss of pyramidal cell excitability may be sufficient to explain the effects observed in patients with schizophrenia across three different auditory paradigms, namely reduced responses in the paired-click, active oddball and passive oddball paradigms^{20,24–29,65,66}.

Secondly, our simulation (Figure 2c) and model comparison (Figure S4d) analyses showed that changing a single model parameter is sufficient to explain changes in the MMN. This provides a very parsimonious explanation without sacrificing biological plausibility compared to previous models, which entail changes in many different parameters including local excitability as well as extrinsic forward, backward and cross-hemispheric connectivity⁸⁰. Our model is still in line with predictive coding accounts of the MMN^{21,99}, which view the MMN

as a precision-weighted prediction error signal that is changed in psychosis^{100,101}. A key feature of our model is that it enables the characterisation of E/I balance, predictive coding and their relationship within the same model, unlike other models that have focused on one view or the other, emphasizing the predictive coding view like classic canonical microcircuit models^{55–58} or the E/I balance view by modelling only E and I populations^{4,38} including Wilson-Cowan-type models¹⁰². In classic predictive coding accounts¹⁰³, forward connections are thought to convey prediction errors and backward connections to convey predictions. Synaptic gain can be viewed as a mechanism instantiating precision-weighting of prediction errors and predictions²¹, and in our model, determining E/I balance. Reducing E/I balance through the parameters we introduced can thus be viewed as reducing the precision of prediction errors or lowering the learning rate in canonical Bayesian models of belief updating (e.g.,^{104,105}) - a mechanism that would result in increased weighting prior information, which has been considered as an explanation for hallucination in schizophrenia^{106,107}. Note, however, that other biological mechanisms can implement precision-weighting, for example sharper population tuning. Future research is needed to determine when this mechanism should be included in the model.

Our model also includes deep pyramidal cells, which are thought to send predictions through extrinsic backward connections⁵⁵. The excitability of these cells can be parameterised by adding another global E parameter to their self-inhibition and relaxing the prior on their contribution to the M/EEG signal in the future. This would allow testing whether the excitability of superficial and deep pyramidal cells differently relates to precision-weighted prediction errors and predictions, respectively.

According to the predictive coding view, changing pyramidal cell excitability corresponds to changing the precision-weighting or learning rate used to scale prediction errors elicited by deviant tones. Thus, our results suggest that a failure of precision-weighting can explain reduced ERPs across all three paradigms. Importantly, this scaling likely involves NMDA receptor-related processing⁵⁸, because cortical gain parameters scale or modulate synaptic connections multiplicatively^{108,109} as opposed to additively unlike other cellular processes such as spatial and temporal summation in the cell soma. The contributions of AMPA, GABA and NMDA receptors can be further explored in future studies with the conductance-based version of the canonical microcircuit we provide here.

Beyond schizophrenia, we believe this model can be used to study E/I balance alterations across various other clinical conditions and the effect of pharmacological interventions on E/I. For example, it is well suited for understanding changes in E/I balance in depression¹¹ and the effects of fast-acting antidepressant effects observed following ketamine administration^{4,110} by modelling suitable M/EEG biomarkers like the MMN. Another interesting avenue for future research will be to apply this model to understand how E/I balance is altered under the influence of psychedelics. Leptourgos et al.⁷ recently proposed that psychedelic-induced hallucinations are likely caused by different circuit changes compared to hallucinations associated with schizophrenia. It has been suggested that 5-HT_{2A} receptor activation on pyramidal cells may lead to disinhibition of these cells^{111,112}. Our model could be used to test this theory empirically and one may expect to find increased excitability of pyramidal cells following psychedelic administration, which would capture 5-HT_{2A}-mediated disinhibition. Additionally, both predictive coding accounts¹¹³ and E/I balance accounts of psychedelics^{7,8} have been proposed and can be investigated together. Our model can be used to determine whether the excitability of superficial pyramidal cells associated with prediction errors or deep pyramidal cells sending back predictions are changed under psychedelics as proposed previously^{7,113}. In the future, it will also be important to study E/I imbalance in other disorders including epilepsy^{9,114–116}, autism¹⁰ and Alzheimer's^{117,118}.

There are several limitations of this study that merit attention. First, the recovery for the inhibitory cell excitability parameter was only fair. While this suggests the parameters can in principle be recovered from these paradigms, its recovery will need to be improved, especially for clinical applications. Future work should investigate whether other paradigms like the 40 Hz auditory steady state paradigm¹⁹ or resting-state are better suited for probing inhibitory cell function. Secondly, it will be important to investigate the test-retest reliability of E/I parameters, because our parameter recovery results only provide an upper bound on the reliability of this computational assay of E/I balance⁹⁴. Third, since we parameterise E cell excitability through an (inhibitory) self-connection, it is

also possible that this parameter reflects a local inhibitory loop perhaps involving parvalbumin-positive (PV+) interneurons. This will need to be clarified in animal studies or iPSC models. iPSCs will also be important to inform priors and constrain parameters (e.g., synaptic time constants) to biologically plausible values in the future.

In conclusion, we introduced a canonical microcircuit for estimating E/I balance. Our simulations suggest that a loss of pyramidal cell excitability may account for various ERP changes observed in schizophrenia. We hope that this tool can serve as a computational assay to assess cell and receptor function based on non-invasive M/EEG recordings^{119,120}, enabling the study of E/I balance across various clinical conditions and ultimately permitting allocation of different treatments to individual patients to restore E/I balance and alleviate symptoms.

Code availability

The posterior parameter estimates (empirical priors for future analyses) and the analysis code for both convolution- and conductance-based models for different integration schemes used in SPM12, TAPAS and the MATLAB default integrator (dde23) along with recommended settings is publicly available on GitHub (https://github.com/daniel-hauke/dcm_ei).

Data availability

Data for the paired-click paradigm is freely available through the NIMH data archive (https://nda.nih.gov/edit_collection.html?id=2274). Access to NAPLS2 EEG data can be arranged through a formal collaboration with the NAPLS2 PIs. Contact Dr. Daniel Mathalon to inquire about such collaborations.

Competing Interests

The authors declare no competing interests.

Funding

DJH, JRS, DAP and RAA are funded by RAA's Future Leaders Fellowship (MR/W011751/1). KF is supported by funding from the Wellcome Trust (226793/Z/22/Z).

Acknowledgments

We thank the B-SNIP and NAPLS2 consortia for generously providing access to the healthy control data used in this study. The B-SNIP and NAPLS data collection was funded by the NIMH. A previous version of this manuscript was made available on bioRxiv¹²¹.

References

- 1 Chagnac-Amitai Y, Connors B. Horizontal spread of synchronized activity in neocortex and its control by GABA-mediated inhibition. *J Neurophysiol* 1989; **61**: 747–758.
- 2 Ticku MK. Alcohol and GABA-benzodiazepine receptor function. *Annals of Medicine* 1990; **22**: 241–246.
- 3 Eisen AJ, Kozachkov L, Bastos AM, Donoghue JA, Mahnke MK, Brincat SL *et al*. Propofol anesthesia destabilizes neural dynamics across cortex. *Neuron* 2024; **112**: 2799–2813. e9.
- 4 Fagerholm ED, Leech R, Williams S, Zarate Jr CA, Moran RJ, Gilbert JR. Fine-tuning neural excitation/inhibition for tailored ketamine use in treatment-resistant depression. *Translational Psychiatry* 2021; **11**: 335.
- 5 Umbricht D, Koller R, Vollenweider FX, Schmid L. Mismatch negativity predicts psychotic experiences induced by NMDA receptor antagonist in healthy volunteers. *Biological Psychiatry* 2002; **51**: 400–406.
- 6 Javitt DC, Steinschneider M, Schroeder CE, Arezzo JC. Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia. *Proceedings of the National Academy of Sciences* 1996; **93**: 11962–11967.
- 7 Leptourgos P, Bouttier V, Denève S, Jardri R. From hallucinations to synaesthesia: A circular inference account of unimodal and multimodal erroneous percepts in clinical and drug-induced psychosis. *Neuroscience & Biobehavioral Reviews* 2022; **135**: 104593.
- 8 Bedford P, Hauke DJ, Wang Z, Roth V, Nagy-Huber M, Holze F *et al*. The effect of lysergic acid diethylamide (LSD) on whole-brain functional and effective connectivity. *Neuropsychopharmacology* 2023; **48**: 1175–1183.
- 9 Fritschy J-M. Epilepsy, E/I balance and GABA A receptor plasticity. *Front Mol Neurosci* 2008; **1**: 201.
- 10 Rubenstein JL, Merzenich MM. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior* 2003; **2**: 255–267.
- 11 Hu Y-T, Tan Z-L, Hirjak D, Northoff G. Brain-wide changes in excitation-inhibition balance of major depressive disorder: a systematic review of topographic patterns of GABA- and glutamatergic alterations. *Mol Psychiatry* 2023; **28**: 3257–3266.
- 12 Krystal JH, Anticevic A, Yang GJ, Dragoi G, Driesen NR, Wang X-J *et al*. Impaired Tuning of Neural Ensembles and the Pathophysiology of Schizophrenia: A Translational and Computational Neuroscience Perspective. *Biological Psychiatry* 2017; **81**: 874–885.
- 13 Sohal VS, Rubenstein JLR. Excitation-inhibition balance as a framework for investigating mechanisms in neuropsychiatric disorders. *Mol Psychiatry* 2019; **24**: 1248–1257.
- 14 Yizhar O, Fenno LE, Prigge M, Schneider F, Davidson TJ, O'Shea DJ *et al*. Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* 2011; **477**: 171–178.
- 15 Singh T, Poterba T, Curtis D, Akil H, Al Eissa M, Barchas JD *et al*. Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* 2022; **604**: 509–516.
- 16 Trubetsky V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB *et al*. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* 2022; **604**: 502–508.

- 17 Glantz LA, Lewis DA. Decreased Dendritic Spine Density on Prefrontal Cortical Pyramidal Neurons in Schizophrenia. *Arch Gen Psychiatry* 2000; **57**: 65–73.
- 18 Lewis DA, Hashimoto T, Volk DW. Cortical inhibitory neurons and schizophrenia. *Nat Rev Neurosci* 2005; **6**: 312–324.
- 19 Grent-'t-Jong T, Brickwedde M, Metzner C, Uhlhaas PJ. 40-Hz Auditory Steady-State Responses in Schizophrenia: Toward a Mechanistic Biomarker for Circuit Dysfunctions and Early Detection and Diagnosis. *Biological Psychiatry* 2023; **94**: 550–560.
- 20 Erickson MA, Ruffle A, Gold JM. A Meta-Analysis of Mismatch Negativity in Schizophrenia: From Clinical Risk to Disease Specificity and Progression. *Biological Psychiatry* 2016; **79**: 980–987.
- 21 Garrido MI, Kilner JM, Stephan KE, Friston KJ. The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology* 2009; **120**: 453–463.
- 22 Hamilton HK, Boos AK, Mathalon DH. Electroencephalography and Event-Related Potential Biomarkers in Individuals at Clinical High Risk for Psychosis. *Biological Psychiatry* 2020; **88**: 294–303.
- 23 Hamilton HK, Mathalon DH. Neurophysiological Models in Individuals at Clinical High Risk for Psychosis: Using Translational EEG Paradigms to Forecast Psychosis Risk and Resilience. In: Javitt DC, McPartland JC (eds). *Neurophysiologic Biomarkers in Neuropsychiatric Disorders: Etiologic and Treatment Considerations*. Springer Nature Switzerland: Cham, 2024, pp 385–410.
- 24 Umbricht D, Krljes S. Mismatch negativity in schizophrenia: a meta-analysis. *Schizophrenia Research* 2005; **76**: 1–23.
- 25 Bramon E, Rabe-Hesketh S, Sham P, Murray RM, Frangou S. Meta-analysis of the P300 and P50 waveforms in schizophrenia. *Schizophrenia Research* 2004; **70**: 315–329.
- 26 Jeon Y-W, Polich J. Meta-analysis of P300 and schizophrenia: Patients, paradigms, and practical implications. *Psychophysiology* 2003; **40**: 684–701.
- 27 Turetsky BI, Greenwood TA, Olincy A, Radant AD, Braff DL, Cadenhead KS *et al*. Abnormal Auditory N100 Amplitude: A Heritable Endophenotype in First-Degree Relatives of Schizophrenia Probands. *Biological Psychiatry* 2008; **64**: 1051–1059.
- 28 Shen C-L, Chou T-L, Lai W-S, Hsieh MH, Liu C-C, Liu C-M *et al*. P50, N100, and P200 Auditory Sensory Gating Deficits in Schizophrenia Patients. *Front Psychiatry* 2020; **11**: 868.
- 29 Rosburg T. Auditory N100 gating in patients with schizophrenia: A systematic meta-analysis. *Clinical Neurophysiology* 2018; **129**: 2099–2111.
- 30 Gunduz-Bruce H, Reinhart RMG, Roach BJ, Gueorguieva R, Oliver S, D'Souza DC *et al*. Glutamatergic Modulation of Auditory Information Processing in the Human Brain. *Biological Psychiatry* 2012; **71**: 969–977.
- 31 Hamilton HK, D'Souza DC, Ford JM, Roach BJ, Kort NS, Ahn K-H *et al*. Interactive effects of an *N*-methyl-d-aspartate receptor antagonist and a nicotinic acetylcholine receptor agonist on mismatch negativity: Implications for schizophrenia. *Schizophrenia Research* 2018; **191**: 87–94.
- 32 Mathalon DH, Ahn K-H, Perry E, Cho H-S, Roach BJ, Blais R *et al*. Effects of Nicotine on the Neurophysiological and Behavioral Effects of Ketamine in Humans. *Front Psychiatry* 2014; **5**: 3.
- 33 Oranje B, van Berckel BNM, Kemner C, van Ree JM, Kahn RS, Verbaten MN. The Effects of a Sub-Anaesthetic Dose of Ketamine on Human Selective Attention. *Neuropsychopharmacology* 2000; **22**: 293–302.

- 34 Plourde G, Baribeau J, Bonhomme V. Ketamine increases the amplitude of the 40-Hz auditory steady-state response in humans. *Br J Anaesth* 1997; **78**: 524–529.
- 35 Rosburg T, Kreitschmann-Andermahr I. The effects of ketamine on the mismatch negativity (MMN) in humans – A meta-analysis. *Clinical Neurophysiology* 2016; **127**: 1387–1394.
- 36 Schwertner A, Zortea M, Torres FV, Caumo W. Effects of Subanesthetic Ketamine Administration on Visual and Auditory Event-Related Potentials (ERP) in Humans: A Systematic Review. *Front Behav Neurosci* 2018; **12**: 70.
- 37 Watson TD, Petrakis IL, Edgecombe J, Perrino A, Krystal JH, Mathalon DH. Modulation of the cortical processing of novel and target stimuli by drugs affecting glutamate and GABA neurotransmission. *Int J Neuropsychopharmacol* 2009; **12**: 357–70.
- 38 Deco G, Ponce-Alvarez A, Hagmann P, Romani GL, Mantini D, Corbetta M. How Local Excitation–Inhibition Ratio Impacts the Whole Brain Dynamics. *J Neurosci* 2014; **34**: 7886–7898.
- 39 Friston KJ. Bayesian Estimation of Dynamical Systems: An Application to fMRI. *NeuroImage* 2002; **16**: 513–530.
- 40 Friston KJ, Harrison L, Penny W. Dynamic causal modelling. *NeuroImage* 2003; **19**: 1273–1302.
- 41 David O, Friston KJ. A neural mass model for MEG/EEG: coupling and neuronal dynamics. *NeuroImage* 2003; **20**: 1743–1755.
- 42 David O, Kiebel SJ, Harrison LM, Mattout J, Kilner JM, Friston KJ. Dynamic causal modeling of evoked responses in EEG and MEG. *NeuroImage* 2006; **30**: 1255–1272.
- 43 Kiebel SJ, David O, Friston KJ. Dynamic causal modelling of evoked responses in EEG/MEG with lead field parameterization. *NeuroImage* 2006; **30**: 1273–1284.
- 44 Kiebel SJ, Garrido MI, Moran RJ, Friston KJ. Dynamic causal modelling for EEG and MEG. *Cogn Neurodyn* 2008; **2**: 121–136.
- 45 Jansen BH, Zouridakis G, Brandt ME. A neurophysiologically-based mathematical model of flash visual evoked potentials. *Biol Cybern* 1993; **68**: 275–283.
- 46 Jansen BH, Rit VG. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol Cybern* 1995; **73**: 357–366.
- 47 Hodgkin A, Huxley A. A quantitative description of membrane current. *J Physiol* 1952.
- 48 Morris C, Lecar H. Voltage oscillations in the barnacle giant muscle fiber. *Biophys J* 1981; **35**: 193–213.
- 49 Pinotsis D, Leite M, Friston K. On conductance-based neural field models. *Front Comput Neurosci* 2013; **7**: 158.
- 50 Pereira I, Frässle S, Heinzle J, Schöbi D, Do CT, Gruber M *et al*. Conductance-based dynamic causal modeling: A mathematical review of its application to cross-power spectral densities. *NeuroImage* 2021; **245**: 118662.
- 51 Moran RJ, Pinotsis DA, Friston KJ. Neural masses and fields in dynamic causal modeling. *Front Comput Neurosci* 2013; **7**: 57.
- 52 Pinotsis DA, Loonis R, Bastos AM, Miller EK, Friston KJ. Bayesian Modelling of Induced Responses and Neuronal Rhythms. *Brain Topogr* 2019; **32**: 569–582.

- 53 Douglas RJ, Martin KA. A functional microcircuit for cat visual cortex. *J Physiol* 1991; **440**: 735–769.
- 54 Haeusler S, Maass W. A Statistical Analysis of Information-Processing Properties of Lamina-Specific Cortical Microcircuit Models. *Cereb Cortex* 2007; **17**: 149–162.
- 55 Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. Canonical Microcircuits for Predictive Coding. *Neuron* 2012; **76**: 695–711.
- 56 Pinotsis DA, Schwarzkopf DS, Litvak V, Rees G, Barnes G, Friston KJ. Dynamic causal modelling of lateral interactions in the visual cortex. *NeuroImage* 2013; **66**: 563–576.
- 57 Shaw AD, Moran RJ, Muthukumaraswamy SD, Brealy J, Linden DE, Friston KJ *et al.* Neurophysiologically-informed markers of individual variability and pharmacological manipulation of human cortical gamma. *NeuroImage* 2017; **161**: 19–31.
- 58 Adams RA, Pinotsis D, Tsirlis K, Unruh L, Mahajan A, Horas AM *et al.* Computational Modeling of Electroencephalography and Functional Magnetic Resonance Imaging Paradigms Indicates a Consistent Loss of Pyramidal Cell Synaptic Gain in Schizophrenia. *Biological Psychiatry* 2022; **91**: 202–215.
- 59 Fish KN, Rocco BR, DeDionisio AM, Dienel SJ, Sweet RA, Lewis DA. Altered Parvalbumin Basket Cell Terminals in the Cortical Visuospatial Working Memory Network in Schizophrenia. *Biological Psychiatry* 2021; **90**: 47–57.
- 60 Osimo EF, Beck K, Reis Marques T, Howes OD. Synaptic loss in schizophrenia: a meta-analysis and systematic review of synaptic protein and mRNA measures. *Mol Psychiatry* 2019; **24**: 549–561.
- 61 Shenton ME, Dickey CC, Frumin M, McCarley RW. A review of MRI findings in schizophrenia. *Schizophrenia Research* 2001; **49**: 1–52.
- 62 Tamminga CA, Ivleva EI, Keshavan MS, Pearlson GD, Clementz BA, Witte B *et al.* Clinical phenotypes of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *American Journal of Psychiatry* 2013; **170**: 1263–1274.
- 63 Clementz BA, Sweeney JA, Hamm JP, Ivleva EI, Ethridge LE, Pearlson GD *et al.* Identification of distinct psychosis biotypes using brain-based biomarkers. *American Journal of Psychiatry* 2016; **173**: 373–384.
- 64 Addington J, Cadenhead KS, Cornblatt BA, Mathalon DH, McGlashan TH, Perkins DO *et al.* North American Prodrome Longitudinal Study (NAPLS 2): Overview and recruitment. *Schizophrenia Research* 2012; **142**: 77–82.
- 65 Hamilton HK, Roach BJ, Bachman PM, Belger A, Carrion RE, Duncan E *et al.* Association Between P300 Responses to Auditory Oddball Stimuli and Clinical Outcomes in the Psychosis Risk Syndrome. *JAMA Psychiatry* 2019; **76**: 1187–1197.
- 66 Hamilton HK, Roach BJ, Bachman PM, Belger A, Carrión RE, Duncan E *et al.* Mismatch Negativity in Response to Auditory Deviance and Risk for Future Psychosis in Youth at Clinical High Risk for Psychosis. *JAMA Psychiatry* 2022; **79**: 780–789.
- 67 Hamm JP, Ethridge LE, Boutros NN, Keshavan MS, Sweeney JA, Pearlson GD *et al.* Diagnostic specificity and familiarity of early versus late evoked potentials to auditory paired stimuli across the schizophrenia-bipolar psychosis spectrum. *Psychophysiology* 2014; **51**: 348–357.
- 68 Friedman D, Simpson G, Hamberger M. Age-related changes in scalp topography to novel and target stimuli. *Psychophysiology* 1993; **30**: 383–396.

- 69 Bak N, Glenthøj BY, Rostrup E, Larsson HB, Oranje B. Source localization of sensory gating: A combined EEG and fMRI study in healthy volunteers. *NeuroImage* 2011; **54**: 2711–2718.
- 70 Nakagawa K, Otsuru N, Inui K, Kakigi R. Change-related auditory P50: A MEG study. *NeuroImage* 2014; **86**: 131–137.
- 71 Kass RE, Raftery AE. Bayes Factors. *Journal of the American Statistical Association* 1995; **90**: 773–795.
- 72 Doeller CF, Opitz B, Mecklinger A, Krick C, Reith W, Schröger E. Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. *NeuroImage* 2003; **20**: 1270–1282.
- 73 Garrido MI, Kilner JM, Kiebel SJ, Friston KJ. Evoked brain responses are generated by feedback loops. *Proceedings of the National Academy of Sciences* 2007; **104**: 20961–20966.
- 74 Molholm S, Martinez A, Ritter W, Javitt DC, Foxe JJ. The Neural Circuitry of Pre-attentive Auditory Change-detection: An fMRI Study of Pitch and Duration Mismatch Negativity generators. *Cereb Cortex* 2005; **15**: 545–551.
- 75 Opitz B, Rinne T, Mecklinger A, von Cramon DY, Schröger E. Differential Contribution of Frontal and Temporal Cortices to Auditory Change Detection: fMRI and ERP Results. *NeuroImage* 2002; **15**: 167–174.
- 76 Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund H-J *et al.* Probabilistic Mapping and Volume Measurement of Human Primary Auditory Cortex. *NeuroImage* 2001; **13**: 669–683.
- 77 Rinne T, Alho K, Ilmoniemi RJ, Virtanen J, Näätänen R. Separate Time Behaviors of the Temporal and Frontal Mismatch Negativity Sources. *NeuroImage* 2000; **12**: 14–19.
- 78 Garrido MI, Kilner JM, Kiebel SJ, Friston KJ. Dynamic Causal Modeling of the Response to Frequency Deviants. *Journal of Neurophysiology* 2009; **101**: 2620–2631.
- 79 Schmidt A, Diaconescu AO, Kommer M, Friston KJ, Stephan KE, Vollenweider FX. Modeling Ketamine Effects on Synaptic Plasticity During the Mismatch Negativity. *Cereb Cortex* 2013; **23**: 2394–2406.
- 80 Gütlin DC, McDermott HH, Grunzei M, Auksztulewicz R. Model-Based Approaches to Investigating Mismatch Responses in Schizophrenia. *Clin EEG Neurosci* 2025; **56**: 8–21.
- 81 Kim H. Involvement of the dorsal and ventral attention networks in oddball stimulus processing: A meta-analysis. *Human Brain Mapping* 2014; **35**: 2265–2284.
- 82 Debener S, Kranczioch C, Herrmann CS, Engel AK. Auditory novelty oddball allows reliable distinction of top-down and bottom-up processes of attention. *International Journal of Psychophysiology* 2002; **46**: 77–84.
- 83 Schöbi D, Do C-T, Frässle S, Tittgemeyer M, Heinzle J, Stephan KE. Technical note: A fast and robust integrator of delay differential equations in DCM for electrophysiological data. *NeuroImage* 2021; **244**: 118567.
- 84 Marreiros AC, Daunizeau J, Kiebel SJ, Friston KJ. Population dynamics: Variance and the sigmoid activation function. *NeuroImage* 2008; **42**: 147–157.
- 85 Markram H, Müller E, Ramaswamy S, Reimann MW, Abdellah M, Sanchez CA *et al.* Reconstruction and Simulation of Neocortical Microcircuitry. *Cell* 2015; **163**: 456–492.

- 86 Ramaswamy S, Courcol J-D, Abdellah M, Adaszewski SR, Antille N, Arsever S *et al.* The neocortical microcircuit collaboration portal: a resource for rat somatosensory cortex. *Front Neural Circuits* 2015; **9**: 44.
- 87 Stern P, Edwards FA, Sakmann B. Fast and slow components of unitary EPSCs on stellate cells elicited by focal stimulation in slices of rat visual cortex. *The Journal of Physiology* 1992; **449**: 247–278.
- 88 Sun Q-Q, Huguenard JR, Prince DA. Barrel Cortex Microcircuits: Thalamocortical Feedforward Inhibition in Spiny Stellate Cells Is Mediated by a Small Number of Fast-Spiking Interneurons. *J Neurosci* 2006; **26**: 1219–1230.
- 89 Bartos M, Vida I, Frotscher M, Geiger JRP, Jonas P. Rapid Signaling at Inhibitory Synapses in a Dentate Gyrus Interneuron Network. *J Neurosci* 2001; **21**: 2687–2698.
- 90 Hájos N, Mody I. Synaptic Communication among Hippocampal Interneurons: Properties of Spontaneous IPSCs in Morphologically Identified Cells. *J Neurosci* 1997; **17**: 8427–8442.
- 91 Wang H, Stradtman GG, Wang X-J, Gao W-J. A specialized NMDA receptor function in layer 5 recurrent microcircuitry of the adult rat prefrontal cortex. *Proceedings of the National Academy of Sciences* 2008; **105**: 16791–16796.
- 92 Zhang W-B, Ross PJ, Tu Y, Wang Y, Beggs S, Sengar AS *et al.* Fyn Kinase regulates GluN2B subunit-dominant NMDA receptors in human induced pluripotent stem cell-derived neurons. *Sci Rep* 2016; **6**: 23837.
- 93 Hauke DJ, Wobmann M, Andreou C, Mackintosh AJ, de Bock R, Karvelis P *et al.* Altered Perception of Environmental Volatility During Social Learning in Emerging Psychosis. *Comput Psychiatr* 2024; **8**: 1–22.
- 94 Karvelis P, Paulus MP, Diaconescu AO. Individual differences in computational psychiatry: A review of current challenges. *Neuroscience & Biobehavioral Reviews* 2023; **148**: 105137.
- 95 Karvelis P, Hauke DJ, Wobmann M, Andreou C, Mackintosh A, Bock R de *et al.* Test-retest reliability of behavioral and computational measures of advice taking under volatility. *PLOS ONE* 2024; **19**: e0312255.
- 96 Fleiss JL. *Design and Analysis of Clinical Experiments*. John Wiley & Sons, 2011.
- 97 Hauke DJ, Roth V, Karvelis P, Adams RA, Moritz S, Borgwardt S *et al.* Increased Belief Instability in Psychotic Disorders Predicts Treatment Response to Metacognitive Training. *Schizophr Bull* 2022; **48**: 826–838.
- 98 Berndt LCS, Singh KD, Shaw AD. Restoring Synaptic Balance in Schizophrenia: Insights From a Thalamo-Cortical Conductance-Based Model. *Schizophr Bull* 2025; : sbaf149.
- 99 Charlton CE, Hauke DJ, Wobmann M, de Bock R, Andreou C, Borgwardt S *et al.* Localizing hierarchical prediction errors and precisions during an oddball task with volatility: Computational insights and relationship with psychosocial functioning in healthy individuals. *Imaging Neuroscience* 2025; **3**: imag_a_00461.
- 100 Charlton CE, Lepock JR, Hauke DJ, Mizrahi R, Kiang M, Diaconescu AO. Atypical prediction error learning is associated with prodromal symptoms in individuals at clinical high risk for psychosis. *Schizophr* 2022; **8**: 105.
- 101 Hauke DJ, Charlton CE, Schmidt A, Griffiths JD, Woods SW, Ford JM *et al.* Aberrant Hierarchical Prediction Errors Are Associated With Transition to Psychosis: A Computational Single-Trial Analysis of the Mismatch Negativity. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 2023; **8**: 1176–1185.

- 102 Wilson HR, Cowan JD. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 1973; **13**: 55–80.
- 103 Rao RPN, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 1999; **2**: 79–87.
- 104 Mathys C, Daunizeau J, Friston KJ, Stephan KE. A Bayesian Foundation for Individual Learning Under Uncertainty. *Front Hum Neurosci* 2011; **5**: 39.
- 105 Mathys CD, Lomakina EI, Daunizeau J, Iglesias S, Brodersen KH, Friston KJ *et al*. Uncertainty in perception and the Hierarchical Gaussian Filter. *Front Hum Neurosci* 2014; **8**: 825.
- 106 Powers AR, Mathys C, Corlett PR. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science* 2017; **357**: 596–600.
- 107 Corlett PR, Horga G, Fletcher PC, Alderson-Day B, Schmack K, Powers AR. Hallucinations and Strong Priors. *Trends in Cognitive Sciences* 2019; **23**: 114–127.
- 108 Friston K, Brown HR, Siemerikus J, Stephan KE. The dysconnection hypothesis (2016). *Schizophrenia Research* 2016; **176**: 83–94.
- 109 Stephan KE, Baldeweg T, Friston KJ. Synaptic Plasticity and Dysconnection in Schizophrenia. *Biological Psychiatry* 2006; **59**: 929–939.
- 110 Hashimoto K. Rapid-acting antidepressant ketamine, its metabolites and other candidates: A historical overview and future perspective. *Psychiatry and Clinical Neurosciences* 2019; **73**: 613–627.
- 111 Aghajanian GK, Marek GJ. Serotonin, via 5-HT_{2A} receptors, increases EPSCs in layer V pyramidal cells of prefrontal cortex by an asynchronous mode of glutamate release. *Brain Research* 1999; **825**: 161–171.
- 112 Nichols DE. Psychedelics. *Pharmacological Reviews* 2016; **68**: 264–355.
- 113 Carhart-Harris RL, Friston KJ. REBUS and the Anarchic Brain: Toward a Unified Model of the Brain Action of Psychedelics. *Pharmacological Reviews* 2019; **71**: 316–344.
- 114 Rosch RE, Wright S, Cooray G, Papadopoulou M, Goyal S, Lim M *et al*. NMDA-receptor antibodies alter cortical microcircuit dynamics. *Proceedings of the National Academy of Sciences* 2018; **115**: E9916–E9925.
- 115 Rosch RE, Hunter PR, Baldeweg T, Friston KJ, Meyer MP. Calcium imaging and dynamic causal modelling reveal brain-wide changes in effective connectivity and synaptic dynamics during epileptic seizures. *PLOS Computational Biology* 2018; **14**: e1006375.
- 116 Rosch R, Burrows DRW, Jones LB, Peters CH, Ruben P, Samarut É. Functional Genomics of Epilepsy and Associated Neurodevelopmental Disorders Using Simple Animal Models: From Genes, Molecules to Brain Networks. *Front Cell Neurosci* 2019; **13**: 556.
- 117 Maestú F, de Haan W, Busche MA, DeFelipe J. Neuronal excitation/inhibition imbalance: core element of a translational perspective on Alzheimer pathophysiology. *Ageing Research Reviews* 2021; **69**: 101372.
- 118 Bi D, Wen L, Wu Z, Shen Y. GABAergic dysfunction in excitatory and inhibitory (E/I) imbalance drives the pathogenesis of Alzheimer's disease. *Alzheimer's & Dementia* 2020; **16**: 1312–1329.

- 119 Karvelis P, Charlton CE, Allohverdi SG, Bedford P, Hauke DJ, Diaconescu AO. Computational approaches to treatment response prediction in major depression using brain activity and behavioral data: A systematic review. *Network Neuroscience* 2022; **6**: 1066–1103.
- 120 Frässle S, Yao Y, Schöbi D, Aponte EA, Heinzle J, Stephan KE. Generative models for clinical applications in computational psychiatry. *WIREs Cognitive Science* 2018; **9**: e1460.
- 121 Hauke DJ, Rodriguez-Sanchez J, Oloye H, Berndt LCS, Pinotsis D, Friston KJ *et al.* A Canonical Microcircuit for Estimating Excitation/Inhibition (E/I) Balance. 2025; : 2025.07.10.664116.
- 122 Hauke DJ. Model overview. 2025<https://biorender.com/tysnd6e>.

Figure Legends

Figure 1. **Model Overview.** At the macroscale, the model requires defining a network of relevant sources that receive a thalamic input U and are connected via extrinsic forward A^F and backward A^B . At the mesoscale, this model extends the canonical microcircuit for predictive coding^{55,56}, modelling each source as a microcircuit with four neuronal populations: E cells or superficial pyramidal cells (**SP**), spiny stellate cells (**SS**), I cells or inhibitory interneurons (**II**) and deep pyramidal cells (**DP**). These populations are connected via intrinsic connections g according to Shaw et al.⁵⁷. To estimate E/I balance, we parameterise the global (across-region) excitability of excitatory pyramidal g_{ee} and inhibitory cells g_{ii} modelled as (inhibitory) self-connections. At the microscale each neuronal population is modelled through two basic operations 1) a convolution with a synaptic response kernel and 2) a sigmoid function to map postsynaptic potentials onto an output firing rate. The figure was created in BioRender¹²².

Figure 2. **Model diagnostics.** Shown are observed responses (lighter colours) and responses predicted by the model (darker colours) for the paired-click (**a**), passive oddball (**d**) or active oddball paradigm (**g**). Correlation matrices with between-subject parameter correlations for 100 randomly selected participants are displayed to illustrate parameter identifiability for the paired-click (**b**), passive oddball (**e**) or active oddball paradigm (**h**). We interpret parameter correlations $r < |0.6|$ as indicative of acceptable parameter identifiability (i.e., limited conditional dependency) in line with Hauke et al.⁹⁷. Parameter recovery is shown for the paired-click (**c**), passive oddball (**f**) or active oddball paradigm (**i**). To measure parameter recovery, we simulated synthetic data using the posterior parameters (x-axis) from models fit to 100 randomly selected participants adding empirical levels of noise, which were estimated from the data, and reinverted the models on the simulated data to obtain recovered parameters (y-axis). We consider intraclass correlation coefficients (**ICC**) to be indicative of poor (<0.4), fair (0.4–0.59), good (0.6–0.74), or excellent (>0.75) recovery in line with Fleiss⁹⁶.

Figure 3. **Simulating changes in condition-specific E and I cell excitability.** Shown are simulated responses in the paired-click paradigm (**a**, **b**), the passive oddball paradigm (**c**, **d**) or the active oddball paradigm (**e**, **f**) when changing condition-specific E cell excitability B^{gee} (left) or I cell excitability B^{gii} (right) over the grid: [-0.5 -0.375 -0.25 -0.125 0 0.125 0.25 0.375 0.5] around the posterior expectation of the parameters (black line). Colder colours indicate increased and warmer colours reduced parameter values. Note that increasing the parameter corresponds to a loss of excitability of either superficial pyramidal (E) or inhibitory interneuron (I) cells since we parameterised excitability through an inhibitory self-connection.

Figure 4. **Interactions between E and I cell excitability.** We simulated data for changing condition-specific E (B^{gee}) and I (B^{gii}) parameters over the grid: [-0.5 -0.375 -0.25 -0.125 0 0.125 0.25 0.375 0.5]. The remaining values are linearly interpolated for visualization purposes only. To illustrate the effects of changing E/I parameters on the EEG, we selected a characteristic EEG component for each paradigm, specifically the P2 (maximum Pz amplitude between 120-250ms) for the paired-click paradigm (**a**), the mismatch negativity (**MMN**; minimum Fz amplitude between 150-250ms) in the passive oddball paradigm (**b**) and P3 (maximum Pz amplitude between 250-600ms) for the active oddball paradigm (**c**). The black colour indicates EEG component amplitudes that were equivalent to the amplitudes generated with the posterior parameter values. Warmer and colder colours indicate more positive and negative EEG component amplitudes, respectively. All component amplitudes were normalised by subtracting the ERP amplitude generated under the posterior parameter estimates. Note that increasing the self-inhibitory connection corresponds to a loss of excitability of either superficial pyramidal (E) or inhibitory interneuron (I) cells since we parameterised excitability through an inhibitory self-connection.

Tables

| Parameter | Description | Prior Mean | Prior Precision |
|--|---|---|-----------------|
| Neuronal Source Model $f(x)$ | | | |
| A{1} | Average extrinsic connectivity (forward, SP to SS) | 200 | 1/16 |
| A{2} | Average extrinsic connectivity (forward, SP to DP) | 25 | 1/16 |
| A{3} | Average extrinsic connectivity (backward, DP to SP) | 50 | 1/16 |
| A{4} | Average extrinsic connectivity (backward, DP to II) | 100 | 1/16 |
| B | Condition-specific extrinsic connectivity | 0 | 1/8 |
| C | Input Strength | 32 | 1/32 |
| G | Average intrinsic connectivity [g_{ee} g_{ii}] | [800 800] | 1/32 |
| B^g | Condition-specific intrinsic connectivity [$B^{g_{ee}}$ $B^{g_{ii}}$] | 0 | 1/8 |
| T | Synaptic time constant [SS SP II DP] | Paired Click: [16 16 2 16] ms Passive oddball: [16 32 2 2] ms Active oddball: [16 128 2 2] ms | 1/32 |
| D | Neuronal delays (all sources) [intrinsic extrinsic] | [1 8] ms | 1/64 |
| S | Slope of output firing rate function | Paired Click: exp(-2) Passive oddball: exp(-1) Active oddball: exp(-1) | 1/64 |
| R | Stimulus [onset dispersion] | Paired Click: [50 30] ms Passive oddball: [68 16] ms Active oddball: [100 16] ms | 1/1024 |
| Forward Model $g(x)$ | | | |
| Lpos | Source location [MNI coordinates] | Paired Click: [-53 -20 13; 49 -10 10; -36 23 3; 36 23 3; -1 -20 67; 1 -20 67] Passive oddball: [-42 -22 7; 46 -14 8; -61 -32 8; 59 -25 8; -46 20 8; 46 20 8] Active oddball: [-61 -32 8; 59 -25 8; -56 7 29; 50 8 30; -33 -42 64; 33 -42 64] | 0 |
| L | Dipole orientation | 0 | 64 |
| J | Neuronal populations contributing to the EEG: [SS SP II DP] | [0 1 0 0] | 0 |

Table 1. Parameter Priors. Note, that the grid search identified different priors for the slope of the sigmoid function S and the synaptic time constants T for different paradigms (cf. *Priors* section). The optimal values are determined by the overall duration of the time window that is modelled and the dominant frequencies; for example, more alpha in paired-click and passive oddball paradigms compared to more theta in the active oddball paradigm. The input parameters were selected to ensure that the input peak coincides with the first dominant peak of the event-related potential of each paradigm. For the motivation of the three different networks (**Lpos**), please see the *Networks* section. **DP** deep pyramidal cells. **II** inhibitory interneurons. **SP** superficial pyramidal cells. **SS** spiny stellate cells.