NMDA-receptor antibodies alter cortical microcircuit dynamics

Richard E. Rosch[a,b,1], Sukhvir Wright[c,d], Gerald Cooray[a,e], Margarita Papadopoulou[f], Sushma Goyal[g], Ming Lim[h,i], Angela Vincent[j], A. Louise Upton[k], Torsten Baldeweg[b], and Karl J. Friston[a]

[a]Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London WC1N 3AR, United Kingdom; [b]Developmental Neurosciences Programme, Great Ormond Street Institute of Child Health, University College London, London WC1N 3EH, United Kingdom; [c]School of Life and Health Sciences, Aston University, Birmingham B4 7ET, United Kingdom; [d]Department of Paediatric Neurology, Birmingham Children's Hospital, Birmingham B4 6NH, United Kingdom; [e]Department of Clinical Neurophysiology, Karolinska Institute, 171 77 Solna, Sweden; [f]ToNIC Toulouse Neuroimaging Center, INSERM/Université Toulouse III-Paul Sabatier (UPS), 31024 Toulouse, France; [g]Department of Clinical Neurophysiology, Evelina London Children's Hospital, Guy's and St. Thomas' NHS Foundation Trust, London SE1 7EH, United Kingdom; [h]Children's Neurosciences, Evelina London Children's Hospital, Guy's and St. Thomas' NHS Foundation Trust, London SE1 7EH, United Kingdom; [i]Faculty of Life Sciences and Medicine, Kings College London, London SE1 7EH, United Kingdom; [j]Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, United Kingdom; and [k]Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford OX1 3PT, United Kingdom

Edited by Terrence J. Sejnowski, Salk Institute for Biological Studies, La Jolla, CA, and approved August 29, 2018 (received for review March 20, 2018)

NMDA-receptor antibodies (NMDAR-Abs) cause an autoimmune encephalitis with a diverse range of EEG abnormalities. NMDAR-Abs are believed to disrupt receptor function, but how blocking this excitatory synaptic receptor can lead to paroxysmal EEG abnormalities—or even seizures—is poorly understood. Here we show that NMDAR-Abs change intrinsic cortical connections and neuronal population dynamics to alter the spectral composition of spontaneous EEG activity and predispose brain dynamics to paroxysmal abnormalities. Based on local field potential recordings in a mouse model, we first validate a dynamic causal model of NMDAR-Ab effects on cortical microcircuitry. Using this model, we then identify the key synaptic parameters that best explain EEG paroxysms in pediatric patients with NMDAR-Ab encephalitis. Finally, we use the mouse model to show that NMDAR-Ab–related changes render microcircuitry critically susceptible to overt EEG paroxysms when these key parameters are changed, even though the same parameter fluctuations are tolerated in the in silico model of the control condition. These findings offer mechanistic insights into circuit-level dysfunction induced by NMDAR-Ab.

NMDA-receptor antibodies | autoimmune encephalitis | EEG | dynamic causal modeling | neural mass model

The recent incorporation of novel cellular-based molecular diagnostics into clinical practice has transformed our ability to identify molecular disruptions of synaptic functions as the cause for a range of neurological disorders (1). For example, antibodies to NMDA receptors (NMDAR-Abs) have been identified as an important cause of autoimmune encephalitis (2), with a particularly high incidence (~40% of patients) in children (3). Patients show a diverse range of symptoms including behavioral changes, movement disorders, and seizures (3, 4). Electroencephalography (EEG) abnormalities have been reported in up to 90% of patients undergoing EEG monitoring; between 20–60% of patients also have epileptiform discharges or electrographic seizures (5, 6). While some EEG features are relatively specific for NMDAR-Ab encephalitis (e.g., extreme delta brush) (6), most are nonspecific, with more global abnormalities associated with more severe disease (7).

NMDAR-Abs mainly affect glutamate transmission through reversible loss of NMDARs, resulting in a reduction of miniature excitatory postsynaptic currents (mEPSCs) in brain slices (8, 9). NMDAR hypofunction is also a hallmark of psychiatric conditions such as schizophrenia and acute psychosis (10, 11) whose clinical features resemble the neuropsychiatric symptoms also seen in early NMDAR-Ab encephalitis. At the whole-organism level, NMDAR-Abs caused an increased seizure susceptibility: Passive transfer of patient Ig containing NMDAR-Abs into a mouse model caused increased susceptibility to chemically induced seizures (12).

Linking NMDAR hypofunction at the cellular level and a predisposition to seizures at the systemic scale is challenging. In the simplified view of epileptic seizures as a consequence of excitation–inhibition imbalance (13), one would expect NMDAR hypofunction to be associated with a reduction of excitation and thus a decrease in seizure susceptibility. While NMDARs are ubiquitous across central synapses, there is differential expression of NMDARs across neuronal populations (14, 15). Therefore, when considering integrated neuronal ensembles, changes in NMDAR function at the level of a single synapse may have a multitude of different emergent effects depending on the combined influence on both excitatory and inhibitory components of the neuronal circuit. Observations in a range of experimental models motivate several mechanistic hypotheses explaining the emergent effects of NMDAR hypofunction. These include (*i*) altered excitatory dynamics with a reduction in late excitatory postsynaptic potential components (9); (*ii*) secondary neurotoxicity reducing the number of functional excitatory connections (16); and (*iii*) a reduction of cortical inhibitory interneuron activity (17). Furthermore, paradoxical changes in excitatory and inhibitory transmission resulting from maladaptive

Significance

Recently, autoantibodies against NMDA receptors (NMDARs) were identified as a major cause of autoimmune encephalitis. They cause abnormalities in brain function often associated with significant changes in patients' brain dynamics. Here we use computational modeling to identify how NMDAR dysfunction causes abnormalities in brain dynamics using patient EEGs and local field potential recordings in a mouse model of NMDAR-Ab encephalitis. NMDAR autoantibodies cause a specific shift in excitatory coupling within cortical circuits that places the circuits closer to pathological transitions between dynamic brain states. Because of the proximity to these phase transitions, otherwise benign fluctuations in neuronal coupling cause abnormal EEG responses in the presence of the antibodies. Our modeling results thus explain fluctuating abnormalities in brain dynamics observed in patients.

Author contributions: S.W. and A.L.U. designed research; R.E.R., S.W., S.G., and M.L. performed research; A.V., A.L.U., and K.J.F. contributed new reagents/analytic tools; R.E.R., S.W., G.C., M.P., S.G., M.L., and A.V. analyzed data; and R.E.R., G.C., M.P., S.G., M.L., A.V., T.B., and K.J.F. wrote the paper.

Conflict of interest statement: A.V. and the University of Oxford hold patents and receive royalties and payments for antibody assays.

This article is a PNAS Direct Submission.

Data deposition: Data and code used in this analysis have been deposited in the Open Science Framework, https://www.doi.org/10.17605/OSF.IO/YXKWD.

[1]To whom correspondence should be addressed. Email: richard.rosch@doctors.org.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1804846115/-/DCSupplemental.

Published online September 27, 2018.

homeostatic changes have been proposed as underlying NMDAR-Ab–associated abnormalities at different temporal scales (8).

In a highly nonlinear dynamic system, such as the brain, the link between synaptic abnormalities and whole-brain responses is rarely intuitive or predictable. Neuronal systems are hierarchically structured, and each observational scale is constrained by larger-scale processes as well as interacting with emergent properties arising from smaller scales (18). Some of these multiscale dynamics can be successfully captured in computational models of neuronal populations, and have been integrated into validated analytic frameworks, such as dynamic causal modeling (DCM) (19).

DCM rests on mesoscale neural mass models that capture the average behaviors of neural populations at the scale of a cortical column. The model used here is representative of generic features of layered cortex referred to as the "canonical microcircuit" (CMC) (20). Its parameters describe synaptic connection strengths and population response dynamics and can be fitted to macroscale neurophysiological recordings such as EEG or LFP recordings. Competing models can then be ranked according to their Bayesian model evidence.

We have chosen this model for two reasons. (*i*) It directly builds on models that have a long history in linking neurobiology to the dynamics of EEG during epileptic seizures (21, 22). These neural mass models constitute neuronal oscillators (i.e., reciprocally coupled inhibitory and excitatory populations) as well as incorporating some key interlaminar connectivity patterns observed across a range of different cortical areas (23, 24). (*ii*) The addition of a second pyramidal cell population in this particular model affords a greater diversity of neuronal dynamics by allowing a separation in time scales between superficial and deep neuronal oscillators, as also observed in empirical laminar recordings (25). In addition to the microcircuit structure, the prior values for the parameters are based on empirical measurements accessible in the literature, where possible (26).

The themes of cortical laminar organization recapitulated in the CMC are conserved across many mammalian species (27). This underpins the use of these models in a range of different experimental systems, ranging from LFP recordings in rodents to invasive recordings in nonhuman primates and EEG/magnetoencephalographic recordings in human subjects (28, 29). Here we exploit this conservation by combining measurements from invasive recordings in a rodent model of NMDAR-Ab encephalitis and human patient EEG recordings at the level of CMC parameters. We report the results of a DCM analysis of (*i*) changes in spontaneous activity in a mouse model of NMDAR-Ab encephalitis and (*ii*) abnormal EEG paroxysms observed in a series of pediatric patients. We first model the NMDAR-Ab effect in the mouse model using DCM to identify a minimal set of synaptic parameters required to produce the NMDAR-Ab effects on ongoing neuronal oscillations. Based on patient EEG data, we then estimate fluctuations in the parameters that explain the patient recordings. We are especially interested in how changes in the different neuronal coupling parameters in combination yield intermittent abnormalities typically recorded in patients. We leverage experimental control (afforded by the animal model) to characterize spontaneous paroxysmal abnormalities (observed in patient recordings): Operationally, we reproduce the parameter changes that explain human EEG paroxysms in the in silico models of experimental effects in mice. This allows us to identify the specific conditions that are necessary for EEG paroxysms to emerge. Understanding this neuronal context for EEG abnormalities may help improve targeted therapeutic approaches in the future.

Results

NMDAR-Abs Alter the Dynamic Response to Acute Chemoconvulsants in Mice. Cortical dysfunction associated with NMDAR-Abs was tested in C57BL/6 mice using a two-by-two design. This design tested for the effects of NMDAR-Abs (delivered via intracerebroventricular injection), the acute chemoconvulsant pentylenetetrazole (PTZ, delivered via a later i.p. injection), and their interaction. LFPs were recorded wirelessly in freely behaving animals; 45 min of recordings pre- and post PTZ injection of eight NMDAR-Ab–positive and five control animals were included for the analysis reported here.

Antibodies alone caused a moderate suppression of the LFP signal across low-frequency bands (delta and theta range) in the NMDAR-Ab–positive mice. However, additional exposure to PTZ revealed a marked difference between NMDAR-Ab–positive and control mice, with a large increase of low-frequency (delta-band, 1–4 Hz) power only in the antibody-positive PTZ-treated mice (Fig. 1). ANOVA revealed a significant main effect of NMDAR-Abs on log-delta-band power [$F(1,4601) = 9.67$; $P = 0.002$] and a significant interaction between NMDAR-Abs and PTZ exposure [$F(1,4061) = 85.05$; $P < 0.001$]. A PTZ-induced increase in paroxysmal fast activity consistent with epileptic seizures was observed in the NMDAR-Ab–positive IgG-treated mice compared with control animals, as previously reported elsewhere (12), but this did not produce a spectral difference in the frequency range analyzed here (Fig. 1). An example of induced, nonepileptiform slow activity is seen in Fig. 1*B*, *Lower*. These slow-wave cortical dynamic abnormalities were further analyzed in the modeling below.

NMDAR-Abs Potentiate PTZ-Induced Effects in Cortical Microcircuitry in Mice. To explain the observed differences in spontaneous activity, hierarchical DCM was used to infer parameter changes associated with the experimental variables over time (i.e., NMDAR-Ab exposure, PTZ infusion, and an Ab–PTZ interaction). In brief, a sliding window (length, 30 s; step size, 15 s) was used to estimate the mean power spectra over successive time points. Each time window was then modeled as the steady-state output of a CMC model (20) with fixed synaptic parameters for the duration of a single time window. By repeating this analysis over windows, we identified fluctuations in synaptic parameters that corresponded to the experimental interventions. Across windows, the evolution of spectral patterns was captured well for all experimental conditions (Fig. 2 *A* and *B*). To infer experimental effects on DCM parameters, the sequence of parameter estimates was then modeled using a parametric empirical Bayesian (PEB) approach (30). Here, slow fluctuations of cortical coupling were modeled as between-window changes in the synaptic parameters estimated within-window [see Papadopoulou

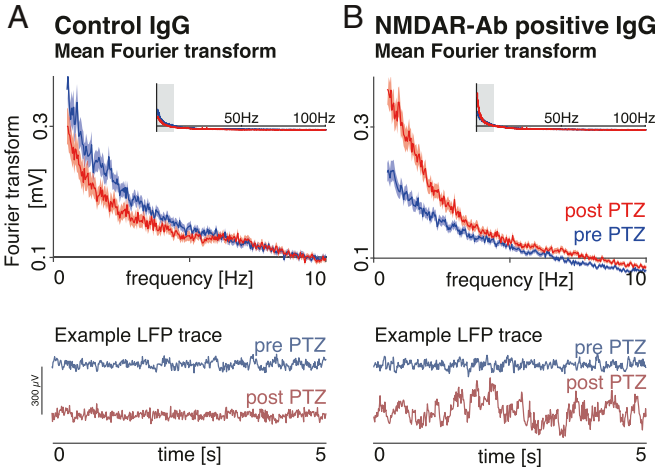

Fig. 1. NMDAR-Abs alter the spectral composition of resting-state activity following PTZ administration. Average Fourier spectra of LFP recordings of endogenous activity in mice are shown. (*A*) In control animals, PTZ injections cause a small decrease in low-frequency power. (*B*) In NMDAR-Ab–positive IgG-treated animals, PTZ causes a profound increase in low-frequency power, which is also visible as high-power slow waves in segments largely without overt epileptiform activity (example shown). Average Fourier spectra across animals are shown for 45-min recordings pre- and post-PTZ injections. Shading indicates the 95% CI. *Insets* show Fourier spectra for a broadband frequency range. Examples of 5-s LFP segments are also shown for individual animals pre- and post-PTZ injections.

et al. (28) for a worked example]. We included three main experimental effects of interest: (*i*) NMDAR-Ab, (*ii*) PTZ, and (*iii*) an NMDAR-Ab × PTZ interaction term (Fig. 2*C*).

The neuronal parameters that affect the spectral composition of spontaneous neuronal activity correspond roughly to the mechanistic hypotheses outlined above: (*i*) time constants of the neuronal populations (τ) describe the dynamics of neuronal population responses; (*ii*) excitatory coupling parameters (g_e) describe the strength of excitatory between-population connections; (*iii*) inhibitory coupling parameters (g_i) represent the strength of inhibitory between-population connections; and modulatory coupling parameters (g_m) represent the strength of inhibitory self-connections (20).

Spectral changes associated with NMDAR-Ab, PTZ exposure, and their interaction were each explained by several corresponding parameter changes. The biggest effects were associated with PTZ exposure, with a decrease in the superficial pyramidal cell population time constant (i.e., a faster return to baseline after perturbation), an increase in the spiny stellate population time constant (i.e., a slower return to baseline after perturbation), and an increase in the excitatory connectivity from spiny stellate to superficial pyramidal cells. Notably those changes were further potentiated by NMDAR-Abs and the NMDAR-Ab × PTZ interaction (Fig. 2*D*).

Shifts in Synaptic Dynamics Underlie the Emergence of Low-Frequency Power in Mice. We further investigated the effect of changes in synaptic parameters on the main spectral data feature of interest, delta-band power. For this, we first performed a principal component analysis over the slow (between-time window) parameter fluctuations separately for time constants and connection strengths, retaining the first principal component of each (Fig. 3 *A* and *B*). This analysis showed that most of the variance over time can be explained by fluctuations in a small subset of parameters, specifically, the time constants of superficial pyramidal and spiny stellate cells and the excitatory coupling between them (as is also apparent in the analysis in Fig. 2).

We use these two components to project synaptic parameter estimates at each time window onto the two dimensions explaining most of the variance (i.e., one time-constant component and one connection-strength component). To characterize different locations in this parameter space in terms of the neuronal dynamics generated by the parameters, we used the mean delta-band power of the predicted power spectral density. This functional characterization of parameter space is shown (in log-scale) with a color code and as isoclines indicating mean delta-band power centiles (Fig. 3*C*). While there is variation in delta-band power associated with both the time-constant (*x* axis) and the connection-strength (*y* axis) parameters, the time constants have the greatest effect on delta power. The difference between controls and NMDAR-Ab–positive animals in the delta-band power post-PTZ is largely conferred by shifting the time-constant component, causing it to cross the 75th delta-band power centile much more frequently than in controls (Fig. 3*D*). This differential effect of PTZ can be seen by comparing the orange and purple dots in Fig. 3*C*.

EEG Paroxysms in Patients Are Caused by Fluctuations in Synaptic Dynamics. To identify which synaptic parameters cause the paroxysmal EEG abnormalities commonly observed in NMDAR-Ab encephalitis, we used the above CMC model to perform a DCM analysis of eight pediatric cases for which EEG recordings

Free parameters fitted by the DCM

Neuronal population time constant

τ_1 superficial pyramidal cell time constant

τ_2 spiny stellate time constant

τ_3 inhibitory interneuron time constant

τ_4 deep pyramidal cell time constant

Coupling parameter

g_1 superficial pyramidal cell to spiny stellate cell inhibition

g_2 inhibitory interneuron to spiny stellate cells inhibition

g_3 inhibitory interneuron to deep pyramidal cell inhibition

g_4 spiny stellate cell to superficial pyramidal cell excitation

g_5 spiny stellate cell to inhibitory interneuron excitation

g_6 deep pyramidal cell to inhibitory interneuron excitation

g_7 superficial pyramidal cells self-modulation

g_8 spiny stellate cell self-modulation

g_9 inhibitory interneuron self-modulation

g_{10} deep pyramidal cell self-modulation

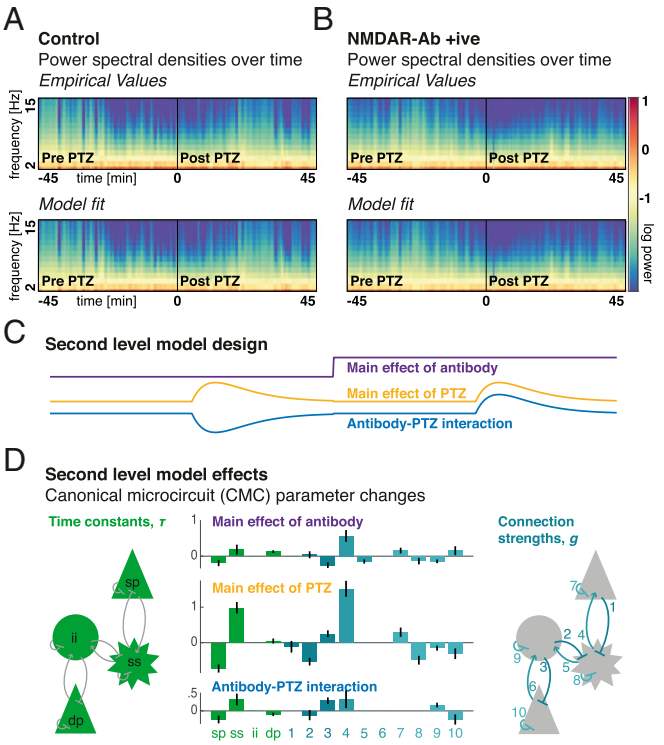

Fig. 2. Synergistic changes in synaptic coupling explain the effects of PTZ and NMDAR-Ab. (*A* and *B*) DCMs were fitted to sliding-window power spectral density summaries of LFP recordings separately for control (*A*) and NMDAR-Ab–positive (*B*) animals. (*Upper*) Observed power spectra over time. (*Lower*) Model fits. (*C*) A second-level general linear model was used to estimate parameter changes associated with NMDAR-Ab exposure, PTZ, and their interaction. The regressors for the three main effects are shown. (*D*) These experimental effects are associated with parameter changes across all populations of the CMC neural mass model. (*Left*) The population-specific synaptic time constants that parameterize the temporal dynamics of postsynaptic responses within that population. (*Right*) Excitatory connections between populations, inhibitory connections between populations, or self-inhibitory connections. (*Center*) Each of the parameters is modulated by each of the experimental effects. The strongest effects are caused by PTZ, with the biggest associated changes in superficial pyramidal cell and spiny stellate cell time constants and excitatory connection strength 4. These changes are further potentiated by NMDAR-Ab exposure. Error bars indicate Bayesian 95% CIs. *dp*, deep pyramidal cells; *ii*, inhibitory interneurons; *sp*, superficial pyramidal cells; *ss*, spiny stellate cells.

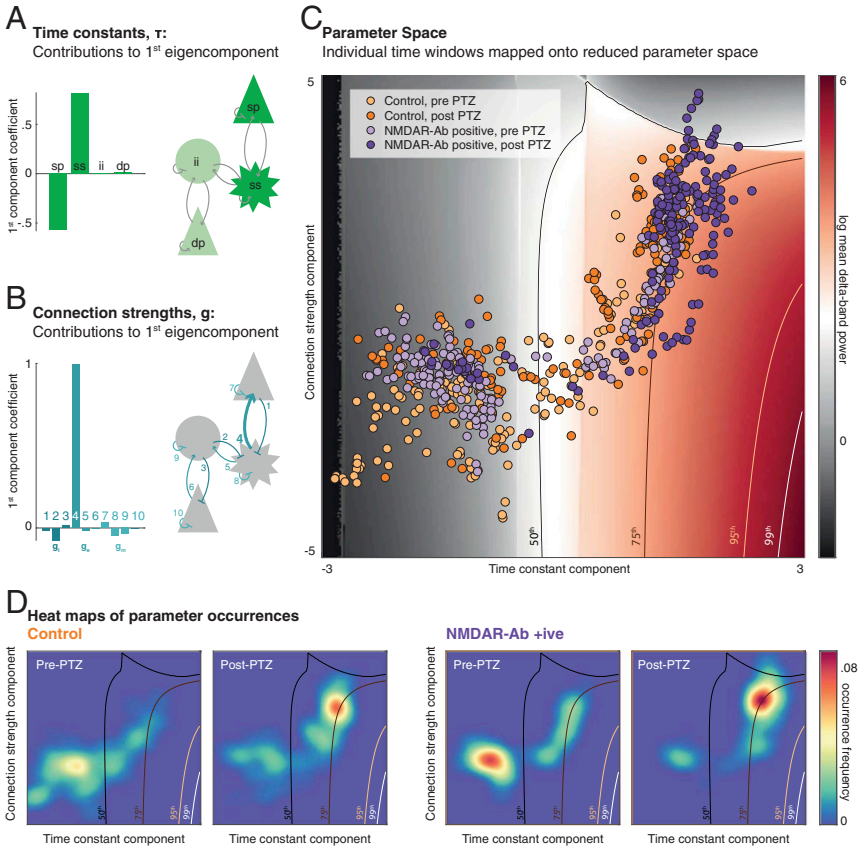

Fig. 3. NMDAR-Abs push the neuronal ensemble into high delta-band power regions of reduced parameter space. (*A* and *B*) Parameter variations between time windows are projected onto the first principal component of time-constant changes consisting predominantly of superficial pyramidal cell and spiny stellate cell changes (*A*) and onto the first principal component of connectivity-strength changes consisting predominantly of changes in coupling of spiny stellate cells to superficial pyramidal cells (*B*). (*C*) Across this parameter space, simulations can predict spectral densities, of which the log mean delta power is shown here (with selected centile isoclines shown). Individual time windows across the four conditions are then projected into the same reduced parameter space, showing an accumulation of NMDAR-Ab–positive, post-PTZ time-window estimates in high-delta ranges. (*D*) The distribution of time windows in parameter space is further illustrated with smoothed heat maps of parameter combination occurrence frequencies over the same section of parameter space for control animals (*Left*) and NMDAR-Ab–positive animals (*Right*). Estimates in NMDAR-Ab–positive animals cross the 75th centile more frequently than in controls. *dp*, deep pyramidal cells; *ii*, inhibitory interneurons; *sp*, superficial pyramidal cells; *ss*, spiny stellate cells.

were available and contained visually apparent EEG paroxysms. Briefly, routine visual EEG analysis to identify paroxysmal abnormalities was performed by two EEG-trained clinicians (R.E.R. and G.C.) (*SI Appendix*, Table S1). For each patient, 2-s time windows containing spontaneous activity, short isolated paroxysms, or rhythmic/ongoing epileptiform activity were extracted and used for further analysis (Fig. 4*A*).

Cortical source estimation for the paroxysmal EEG activity was performed, and "virtual electrode" responses were extracted from the most active sources (31). For each patient, DCMs were independently fitted to power spectral density averages of each available condition (e.g., background, short paroxysms, and ongoing rhythmic activity) (Fig. 4*B*). Individually fitted DCMs (with near-perfect model fits) (Fig. 4*C*) were subsequently combined in within-patient, between-condition hierarchical (PEB) models that explained the condition-specific differences with changes in synaptic time constants (τ), between-population inhibitory connections (g_i), between-population excitatory connections (g_e), or within-population modulatory connections (g_m). Across participants, models explaining spectral differences as arising from differences in time constants offer the best explanation of the virtual electrode data (with an exceedance probability of >95%) (Fig. 4*D*).

NMDAR-Abs Alter the Response to Intrinsic Fluctuations in Synaptic Dynamics. The DCM of human data provides us with an estimate of brain-state–specific changes in synaptic parameters. From the Bayesian model comparison of a set of reduced models, it emerged that the differences in the EEG states in human patients are best explained through variations in neuronal population time constants. We extracted DCM parameter estimates of these time-constant changes for each patient, yielding a set of time-constant parameters that explain the transition from background to paroxysmal EEG states for each participant individually. From this matrix of time constants (four time constants in eight participants with two or three EEG states), we extracted the first principal components and applied them to the control and the NMDAR-Ab–positive mouse-derived CMC model. Conceptually, we are thus enforcing the same degree of time-constant fluctuations estimated from the patient models to the in silico microcircuits derived from the mouse experiments.

The differences between the parameter estimates from the control and NMDAR-Ab–positive model result in different spectral outputs even when the same time-constant changes are applied. Overall, the NMDAR-Ab–positive context results in higher delta-band power and less high-frequency power (Fig. 5 *A* and *B*). Crucially, delta power was higher in the NMDAR-Ab–positive model across a wide range of time-constant fluctuations (Fig. 5*C*). Furthermore, small changes in the synaptic parameters identified with the patient data cause large changes in delta power in, and only in, the NMDAR-Ab–positive model. This is manifest as low-frequency paroxysmal activity when the synaptic parameters change slightly in the NMDAR-Ab–positive model but not in

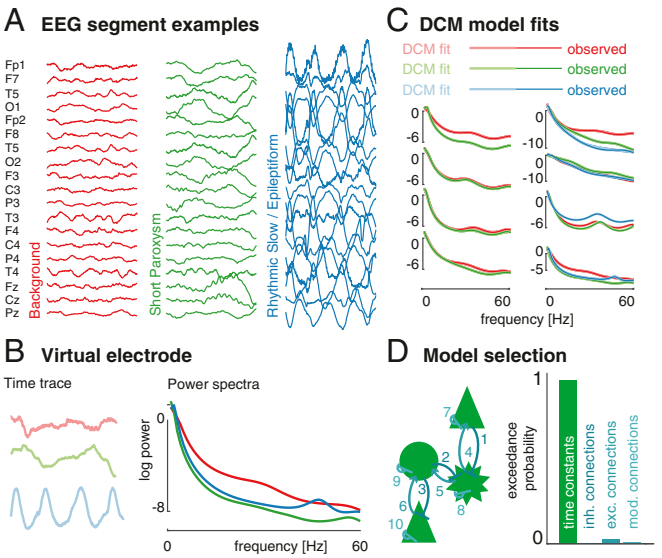

Fig. 4. EEG paroxysms in NMDAR-Ab encephalitis patients are best explained as time-constant fluctuations. (*A*) For each individual patient, 2-s time windows containing spontaneous activity, short EEG paroxysms, and, where available, longer rhythmic EEG activity were extracted. (*B*) These fluctuations were source localized, and virtual electrode time traces were extracted at the estimated cortical source. Normalized power spectral density averages across all time windows were then fitted using separate DCMs for each condition. (*C*) The normalized spectral outputs of fitted DCMs show near-perfect overlap with the observed spectral densities, illustrating that the fits provide good explanations of the observed (spectral) data features. (*D*) We then used Bayesian model reduction to test which subset of parameters best explains the differences between the different EEG states across the whole group. For each individual, between-condition effects were estimated in a number of reduced (PEB) models that differed only in which parameters were free to explain the between-window spectral variations. Of these PEB summaries of individual participants, models explaining the spectral changes with fluctuations in time constants have an exceedance probability of >95%.

the control (Fig. 5*D*). Technically, this abrupt change in dynamics with a small change in parameters is known as a "phase transition," suggesting that antibody-positive effects on synaptic coupling move the network toward a critical regime in which small fluctuations in synaptic time constants produce qualitatively different dynamics (i.e., paroxysmal EEG abnormalities).

Discussion

This study reveals common synaptic mechanisms underlying a range of electrophysiological disturbances associated with NMDAR-Abs in a mouse model and in pediatric patients: NMDAR-Abs cause a shift in cortical synaptic parameters that is associated with an increase in low-frequency oscillations and which predisposes microcircuits to the slow-wave paroxysms seen in the clinical EEG recordings.

NMDAR-Abs Are Associated with High-Amplitude Low-Frequency Discharges. NMDAR-Abs cause changes in the spectral composition of the resting-state LFP of the mouse strain tested. These differences are further revealed on additional exposure to PTZ, with a large PTZ-induced increase in mean delta power in the presence of NMDAR-Ab. This increase is largely due to intermittent rhythmic slowing without concurrent epileptic spikes. Previous analysis of seizure events shows that NMDAR-Abs also lower the seizure threshold (12), but seizure events fall largely outside the frequency spectrum analyzed here. These observations are in keeping with clinically reported EEG features, i.e., background slowing with or without additional slow-wave paroxysms.

In mouse models of NMDAR hypofunction, normal NMDAR function in parvalbumin (inhibitory) interneurons is required for gamma rhythm induction (32). Furthermore, persistent NMDAR hypofunction confers an increase in resting gamma power with a concurrent reduction in stimulus-induced gamma oscillations (33). In the mouse model presented here, we did not see such a change in gamma frequency power, which may reflect the modeled disease stage: Patients with NMDAR antibody encephalitis progress through distinct stages, initially presenting with neuropsychiatric symptoms before developing a more severe encephalopathic syndrome associated with slowing of the EEG (34). The findings of our study relate most closely to this second stage and thus may relate to a pathophysiology distinct from neuropsychiatric NMDAR hypofunction. In the early stages of disease, the antibody effects may be limited to the inhibitory interneuronal system (10). However, during the encephalopathic stages (including the emergence of paroxysmal EEG abnormalities and epileptic seizures), there may be more wide-ranging effects across cell types that are partly recapitulated in our mouse model.

An increase in the power of slow-frequency components in an EEG or LFP recording is thought to be associated with increased synchronization of local cortical firing, itself regulated by interacting cortical and subcortical systems, e.g., thalamocortical loops (35), brainstem monoamine arousal systems (36), and intrinsic cortical effects such as astrocytic regulation of synaptic function (37). Firing synchrony can occur physiologically (e.g., during sleep), can be associated with nonspecific cortical dysfunction (e.g., in the context of an encephalopathy), or can be a component of epileptic discharges (apparent in slow-wave components in spike-wave discharges) (38).

Synchrony, by definition, is an emergent feature of population dynamics rather than a property of any single neuron, but an increase in cortical synchrony may arise from a whole range of different coupling changes at the synaptic level. Many of these can be captured in mesoscale models of neuronal ensembles (39). The DCM approach uses this mesoscale modeling to identify the changes underlying the emergence of hypersynchronous slow-wave activity in the context of NMDAR-Ab.

NMDAR-Abs Cause Laminar-Specific Changes in Cortical Dynamics. DCM rests on neural mass modeling of coupled neuronal oscillators that are described using specific synaptic parameters (e.g., connection strengths and time constants) (see table in text) and that broadly resemble the laminar structure of the cortex. The neural mass model of a single electromagnetic source contains two pairs of coupled neuronal oscillators that support slower (deep oscillator: deep pyramidal cells, inhibitory interneurons) and faster (superficial oscillator: superficial pyramidal cells, spiny stellate cells) activity (25). These populations model the dynamics of an integrated cortical column. Individual parameters exert highly nonlinear effects on the system's output. The parameterization of these models is rooted in biophysical properties of individual neurons but describes average characteristics of populations of functionally related neurons, i.e., composite properties emerging from the features of individual cells.

At this mesoscale, PTZ and NMDAR-Abs produce synergistic effects that result in excessive synchrony not seen in other experimental conditions. Our results suggest that increases in low-frequency power can be explained by a combination of (*i*) an increase in superficial cortical excitatory coupling, largely associated with PTZ exposure, and (*ii*) opposing changes in the dynamics of the superficial oscillator pair (spiny stellate and superficial pyramidal cells) (Fig. 3).

The changes in synaptic dynamics align time constants in a gradient along the CMC coupling chain, with the slowest time constants in the deep pyramidal cells, the fastest time constants in the superficial pyramidal cells, and gradual steps between. This reduces the stepwise difference in time constants along the CMC chain compared with the standard CMC configuration. This parameterization allows a dominant frequency to resonate across and recruit the whole column, thus producing the high-amplitude slow-frequency patterns observed. Thus, interestingly, slow-wave activity appears to be under the control of the faster, superficial oscillator pair in the CMC model, with both NMDAR-Abs and PTZ having profound and relatively specific effects on their dynamics. This is in keeping with observations from invasive recordings of slow-wave

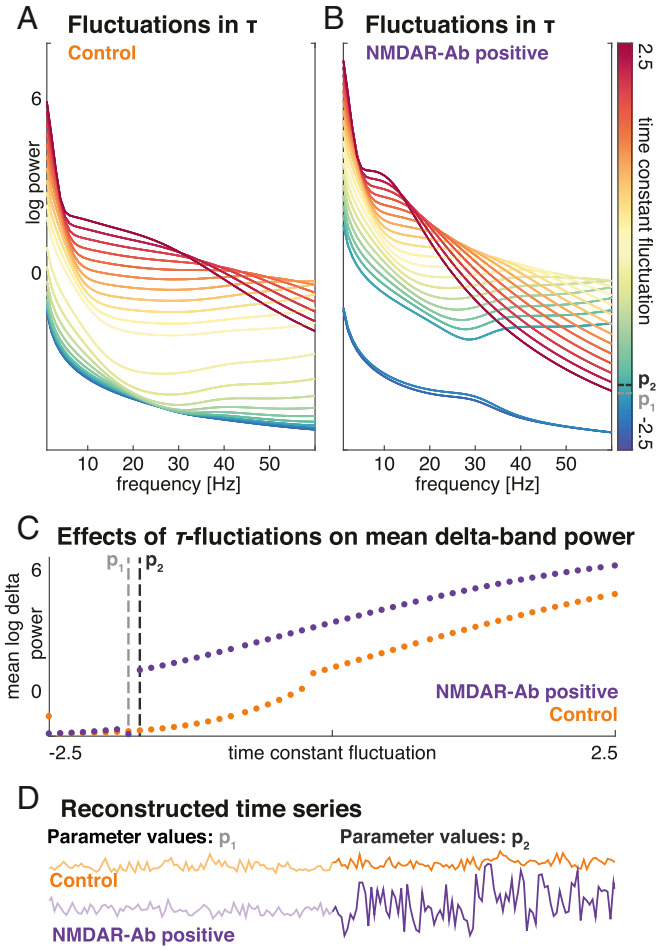

Fig. 5. NMDAR-Abs sensitize the microcircuit to intrinsic fluctuations in time constants. (*A* and *B*) Here, we apply a summary component of the time-constant fluctuations estimated from human patients to a cortical microcircuit model derived from the control mice (*A*) and the NMDAR-Ab–positive mice (*B*). The same fluctuations cause spectral outputs containing much higher relative delta power in the model estimated from NMDAR-Ab–positive mice. (*C*) This figure shows the log of mean delta power for a range of smoothly increasing time-constant fluctuations. In the low parameter range (–2.5 to –1.5), there is a large jump in delta power, suggesting that there are two distinct dynamic states separated by small differences in parameter values. (*D*) Examples of reconstructions of time series for parameter values at two very close parameter values (p_1 and p_2) are shown for control and NMDAR-Ab–positive models. The sudden increase in delta power is visible as a paroxysmal change in the time series in the NMDAR-Ab–positive context, while the control time series appears continuous. This sudden change in dynamics with a small change in parameter space is known as a "phase transition."

activity in human patients with epilepsy, which implicate superficial cortical coupling in the regulation of slow-wave sleep activity (40).

Different Molecular Changes Show Converging Effects at the Neuronal Population Level. The synaptic parameters of the CMC model employed in DCM are population summaries of a variety of cellular effects, encompassing emergent properties and multiple nonlinearities (41). Time constants at the population level are essentially descriptions of the dynamics of postsynaptic integration affected by multiple factors, such as background firing frequency, membrane conductance, intra- and extracellular ion composition, and the dynamics of receptor types present in the membrane, to name only a few (42). Connection strengths at the population level summarize the effect one population has over another. This may include effects mediated through assumed subpopulations contained within the modeled populations (e.g., self-connections are modeled as direct connectivity but represent local intralaminar inhibitory interneuronal inhibition). Because a number of different effects may converge on the same population parameters, and individual molecular effects may be expressed only in certain conditions, the link between molecular change and population parameters is nontrivial.

The approach presented here deliberately collapses much of the diversity of cortical physiology into mathematical descriptions that have specifically been developed to capture the sort of abnormal responses observable in EEGs. This offers the opportunity to integrate findings from diverse sets of recordings into a common mathematical framework describing ongoing cortical dynamics.

Exposure to NMDAR-Abs has been reported to cause a number of changes in the postsynaptic glutamate response, including a reduction in overall postsynaptic potential, a reduction in late postsynaptic currents, and a faster return to baseline (8, 9). In intact neuronal circuits, NMDARs exert differential control over excitatory and inhibitory populations, leaving the populations differentially affected by NMDAR blockade (43).

PTZ is believed to act as an antagonist to $GABA_A$ receptors by directly blocking ionophores (44). $GABA_A$ receptors are fast inhibitory receptors with a widespread region and cell-type specific set of postsynaptic effects (45). These include inhibitory postsynaptic potentials but also inhibition of dendritic excitatory postsynaptic potentials via extrasynaptic $GABA_A$ receptors, which is particularly pronounced at the cortical pyramidal cells (46). In some neuronal cell types and at certain developmental stages $GABA_A$ can cause excitatory postsynaptic potentials (47), and GABA transmission can exert direct or indirect control over excitatory NMDAR-dependent synaptic transmission (48).

With this range of different cellular effects, it is unlikely one can capture the breadth of NMDAR-Ab– and PTZ-related effects in a small subset of population model parameters. However, the effects on delta-band power can be reproduced well with a few principal components comprising largely just two main effects: (*i*) decreasing the time constants of superficial pyramidal cells relative to excitatory spiny stellate cells and (*ii*) increasing the excitatory coupling between spiny stellate and superficial pyramidal cells.

A number of possible and convergent changes at the molecular level associated with NMDAR-Ab and PTZ exposure could explain these population-level effects. The time-constant changes in superficial pyramidal cells may result from being switched toward (faster) AMPA-mediated excitatory inputs (due to the NMDAR-Ab–mediated internalization of NMDAR) and a change in membrane conductivity (due to PTZ-mediated blocking of extrasynaptic $GABA_A$ receptors). The change in excitatory connection, on the other hand, is consistent with a disinhibition of excitatory EPSPs under $GABA_A$ blockade with PTZ (i.e., a block of so-called "shunting inhibition") (46). Furthermore, different inhibitory interneuron populations characterized by different molecular markers, morphology, and functional integration show distinct and at times opposing overall effects on cortical dynamics (49). Thus, features of neuronal dynamics that in our models appear as changes in excitatory coupling at the population level may result from subpopulation-specific changes in cortical (dis) inhibition.

Experimental methods to link detailed cell type-specific cortical physiology and population dynamics exist in experimental animals (50) but cannot be accessed directly in patient EEGs. By focusing on the mesoscale descriptions here while enforcing a simplified representation of cortical dynamics, we can describe the systemic effects of NMDA Abs in recordings from the mouse model as well as in patient EEGs. The increasingly detailed characterization of mouse cortex circuitry provides an important focus for progressively refining these sorts of models in future studies.

Each of these changes at the molecular level (approximated through the population-model parameters) has nonlinear effects on neuronal responses. Thus, as in other complex systems, even small fluctuations in the parameters induced by only a relative shift in the balance of, e.g., AMPA receptor- and NMDAR-mediated transmission can have profound effects on the activity of the dynamics of the whole circuit. These effects can be quantified using a sensitivity analysis, i.e., quantifying how changes in model parameters produce distinct spectral responses (*SI Appendix*, Fig. S1).

NMDAR-Abs Sensitize the Cortical Column to Spontaneous Paroxysmal EEG Abnormalities. In the patients with NMDAR-Ab encephalitis, there is no experimental control over NMDAR-Ab exposure. Furthermore, our sample of patients is heterogeneous, representative of clinical practice (e.g., varying in age, gender, timing of EEG, timing of initial diagnosis, and other characteristics). However, these patients show a diverse range of paroxysmal, short-term changes in EEG dynamic patterns that are visually apparent, allowing us to probe spontaneous fluctuations of DCM parameters that may underlie discrete pathological brain states.

Patient-specific modeling, as facilitated by DCM, allows inference about patient-specific parameters in a generic model of the cortical column. Thus, applying DCM analysis to this diverse sample, one can access two types of results: (*i*) qualitative, i.e., identifying the parameters whose changes underlie the dynamic abnormalities seen in EEG, and (*ii*) quantitative, i.e., establishing the numerical range of parameter fluctuations that can be applied to other specified DCMs.

Consistently across patients, models with changes in time constants best explained the observed transitions between background activity and paroxysms. Furthermore, we summarized these parameter changes along a single (principal component) axis. We used this component to enforce similar fluctuations in the fully specified DCMs derived from the mouse model analysis, asking whether the baseline context (i.e., the parameterization derived from NMDAR-Ab–positive or control animals) alters the impact of parameter changes of the magnitude observed in human patients.

Indeed the dynamic responses of the two types of models are very different. In the context of NMDAR-Ab, overall greater delta-band power is observed, and there are regimes of parameter space that contain boundaries between very different dynamic states (51). This structural instability underwrites phase transitions of the sort seen in seizure activity or other EEG state transitions. In the control parameterization the same changes have a much less pronounced effect and do not induce overt slow-wave paroxysms. In short, it appears that paroxysmal EEG activity in patients may be best explained by normal fluctuations in synaptic time constants that occur in an abnormal regime of synaptic parameter space.

In the human patients we relied on EEG recordings that were obtained in the absence of external experimental controls, effectively using the modeling to describe the kinds of neuronal coupling changes that cause spectral shifts in the EEG as observed during short-term paroxysms. Furthermore, dynamic features observable in EEG are averages of larger-scale network activity than the LFP recordings in mice. However, the insight that the spectral shift induced by changes in these parameters depends on NMDAR-Abs was afforded only through the experimentally constrained DCMs estimated from the animal model. The translation between species and modalities is afforded by explicit generative neuronal models that relate data features to underlying neuronal population activity. In the case of human EEGs, we first extract local cortical time series using a virtual electrode, i.e., a beamformer source reconstruction algorithm, before fitting mesoscopic neuronal populations models using DCM. In the case of LFP recordings in mice, we can use DCM directly to explain the recorded LFPs (with the observation model consisting of a single scalar gain parameter). Crucially these generative models then allow neurobiologically relevant features (i.e., changes in connection strength and synaptic dynamics) to be translated at the same mesoscopic scales.

Overall, these findings provide integrative evidence from human patients and a mouse model of NMDAR-Ab encephalitis suggesting that (*i*) NMDAR-Abs cause electrophysiological abnormalities via a small number of synaptic changes, which may lend themselves to targeted therapeutic interventions, e.g., by exploiting laminar and/or cell type-specific effects of transcranial current stimulation (52), and (*ii*) paroxysmal abnormalities can be explained by persistent baseline changes that render cortical microcircuitry particularly sensitive to (potentially normal) fluctuations in synaptic coupling. Future research may reveal whether similar approaches have diagnostic value when performed on patient EEGs alone (53).

Limitations. The modeling approach presented here allows unique insights into possible mechanisms underlying empirically observed phenomena. Although DCM has been applied to a wide variety of neurophysiological studies, and its validity has been assessed repeatedly (54, 55), there are certain limitations to the approach adopted here.

First, that modeling can be applied only to existing data places restrictions on study design (e.g., pre–NMDAR-Ab exposure EEGs are not usually available from patients) and limits the approach to a subset of testable hypotheses. Second, like all inferences, DCM is based on specific assumptions regarding the underlying neuronal architecture. All activity presented here is presumed to emerge from microcircuitry consistent with the CMC model, and only given this assumption can we estimate the parameters and evidence for or against specific model configurations.

Most importantly, we have reduced a complex brain-wide pathology of interacting systems to changes in a cortical microcircuit. Thus, we are ignoring interactions between different cortical regions as well as the influence of subcortical structures, such as thalamus and brainstem, which (especially in the context of encephalopathy and slow-wave abnormalities) will exert a powerful influence over cortical states. Although these effects can be accommodated in the model as random effects, they are not modeled explicitly.

The approach here, by design, does not focus on single-cell dynamics but treats cortical patches as integrated units, which correspond to the mesoscale dynamics observable in EEGs (18). Our computational modeling of cortical microcircuits aims to link some of these observed dynamics with the themes that have emerged from detailed microanatomy and neurophysiology at the microscale and are implicated in NMDAR-related pathology. However, we are not attempting to make inferences about single neurons. The models describe the net effect of NMDAR-Abs on integrated circuits of neuronal populations; future studies should allow us to model how those effects emerge from single-neuron interactions. Furthermore, we note that many of the canonical models of cortical circuitry (including ours) have focused on excitatory population coupling and in the future may benefit from incorporating some of the more recent themes that have been identified in the connectivity patterns of inhibitory interneuronal populations (56). Our study aimed to answer specific questions driven by observations in a particular pathology related to NMDAR-Ab. While the models and results as presented are appropriate for this focus, there are many observations related to abnormal NMDAR function that are not currently captured in the model here. We hope that future research will integrate such experiments and observations and expect that generative models such as the one presented here will help in this work.

Methods

Our analysis uses DCM to infer the neurobiological parameters that underlie electrophysiological changes in patients with NMDAR-Ab encephalitis and a corresponding mouse model. Once these changes have been identified, we use the fully parameterized in silico models in simulation mode to integrate the findings and explore hypotheses about how NMDAR-Ab–induced changes in the neurobiology cause the EEG abnormalities observed in patients.

For this, the analysis is broadly divided into three stages, which are explained in detail below: (*i*) using DCM of LFP recordings in the mouse model (exploiting the factorial experimental design), we estimate neuronal population coupling in the cortical microcolumn induced by NMDAR-Abs (Fig. 6 *A–D*); (*ii*) in a corresponding DCM analysis of spontaneous EEG paroxysms recorded in human patients, we estimate fluctuations in microcircuit coupling (Fig. 7); and (*iii*) we implement these patient EEG-derived parameter fluctuations in the in silico representations of microcircuits derived from control and NMDAR-Ab conditions in the mouse experiment. This allows us to investigate the dynamics of the microcircuits with and without NMDAR-Ab effects, testing whether the association between certain model parameter shifts and paroxysmal spectral abnormalities (as observed in human patients) depend on specific dynamic contexts (i.e., conditions in our mouse experiments) (Fig. 6 *E–H*).

Collection and Classic Analysis of Mouse LFP. The mouse model and associated procedures have been previously described (12). Briefly, plasma with NMDAR-Ab (IgG) was obtained with informed consent from three female NMDAR-Ab–positive patients with neuropsychiatric features, movement disorder, and reduced level of consciousness; samples were deidentified before research use.

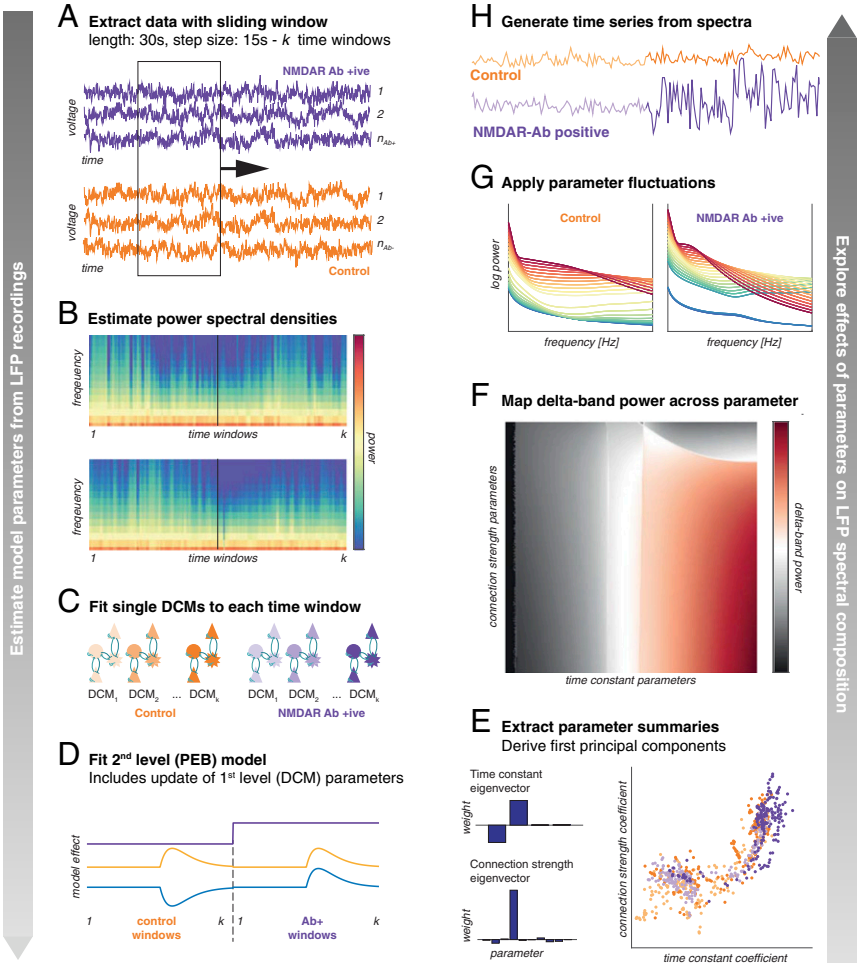

Fig. 6. Modeling approach to mouse LFP recordings. Modeling was designed to extract relevant parameters (*A–D*) and then explore the effects of those on delta power (*E–H*). (*A*) For both pre- and post-PTZ injection, 45 min of LFP recordings were extracted for each mouse. A sliding window was used to extract a sequence of time windows for further analysis. (*B*) Power spectral densities were estimated for each time window, which are the basis for the DCM model fit. (*C*) Single-source DCMs comprising a single CMC model were fitted to each time window separately. (*D*) Using a PEB approach to fit a second-level between-DCM general linear model, we extracted parameter variations explained by specific experimental effects and updated first-level DCM parameters. (*E*) From the updated first-level DCMs, we extracted all parameters and summarized them in two principal components over time constants and connection strengths, retaining the first component summaries of the fitted DCMs. (*F*) Starting from the baseline model specification, we applied the reduced (i.e., first principal component) summaries of the parameter changes to simulating cross-spectral outputs of the neural populations, yielding a map of delta power across the ensuing 2D parameter space. (*G*) We then applied quantitative parameter changes observed in patient EEGs (summarized as their first principal component) to the control and NMDAR-Ab baseline model specifications to explore the effects of parametric fluctuations on spectral output. (*H*) To further illustrate the effects of parametric fluctuations, we applied and inverse Fourier transform to generate substitute time series, illustrating the nature of the changes in a time trace.

Control IgG was purified from serum from two healthy individuals. C57BL/6 female mice aged 8–10 wk were housed and examined according to Animal Research: Reporting of in Vivo Experiments (ARRIVE) guidelines, and all analyses were performed with the observer blinded to injected antibody. Animal experiments were approved by a local ethical review committee at the University of Oxford and performed under license from the UK Home Office in accordance with the Animal (Scientific Procedures) Act 1986.

Wireless telemetry transmitters (s.c. transmitter A3028B-CC from Open Source Instruments, Inc.) were implanted in an s.c. pocket over the right flank. Two craniotomies were performed at 1 mm lateral and 1 mm caudal from bregma. Electrode screws were fixed into the drilled holes with dental cement. After a 5-d monitored recovery period, 8 μL of purified IgG (patient or control) was injected slowly into the left lateral ventricle through a single additional craniotomy made 1 mm left lateral and 0.45 mm caudal from bregma.

Mice were housed in a Faraday cage during wireless LFP data collection. To test seizure susceptibility, 40 mg/kg of PTZ was given i.p., and the mice were observed for 45 min following injection. The 45-min time period immediately preceding PTZ injection was used as the control segment.

Raw LFP data were analyzed in Matlab. Sliding-window (30-s windows, 15-s steps) Fourier estimates of power over frequency were used to statistically compare the different conditions. ANOVA over mean delta-band power (1–4 Hz) was used to estimate the effects of the two main interventions (NMDAR-Ab and PTZ) and their interaction on LFP signal composition.

Modeling Cortical Activity with the CMC Model. For the purposes of this analysis, cortical activity is assumed to arise from a cortical microcolumn that consists of four coupled neuronal populations: two main output populations (superficial and deep pyramidal cells) and local inhibitory and excitatory populations (inhibitory interneurons and spiny stellate cells, respectively). These populations are based on both established models of cortical function (22, 57) and empirically observed connectivity patterns (26, 58, 59). These populations are organized into two oscillator pairs: one superficial (consisting of superficial pyramidal cells and spiny stellate cells) and one deep (consisting of deep pyramidal cells and inhibitory interneurons). This architecture recapitulates generic themes in cortical organization while allowing a diverse range of dynamics enabled by the two coupled oscillator pairs (see Fig. 2*D* for an illustration of the model architecture) (20).

Intralaminar connectivity is largely represented within neuronal populations. Each population is parameterized by recurrent inhibitory self-connection parameters, population time constants, and a parameterized sigmoid function that models the dispersion of population responses. Interlaminar coupling is modeled explicitly through population-level connectivity between populations. Each oscillator pair has reciprocal excitatory and inhibitory connectivity. Note that indirect inhibition from superficial pyramidal cells to excitatory interneurons, mediated via assumed intralaminar inhibitory interneurons, is absorbed into a direct inhibitory connection.

DCM Analysis of Mouse LFPs. Dynamic causal modeling was performed using SPM12, an academic software package (https://www.fil.ion.ucl.ac.uk/spm/). All analysis code and raw data are available online at https://www.doi.org/10.17605/OSF.IO/YXKWD, which requires Matlab 2014b or later and SPM12. Modeling of the mouse LFP recordings can be divided into the following steps (summarized in Fig. 6):

i) Inversion of separate single-source DCM for each time window (performed on group-average data)
ii) Second-level (PEB) modeling to explain parameter changes over time, based on experimental interventions
iii) Forward modeling to explore the effects of parameter changes on specific output measures (e.g., delta power)

Individual time windows were assumed to be relatively stationary within the 30-s sliding time window, in line with previous DCM analyses of EEG seizure activity (29, 31). Each time window was modeled as originating from a single cortical source comprising four coupled neuronal populations (i.e., a single cortical column modeled as a single CMC). DCM employs a standard variational

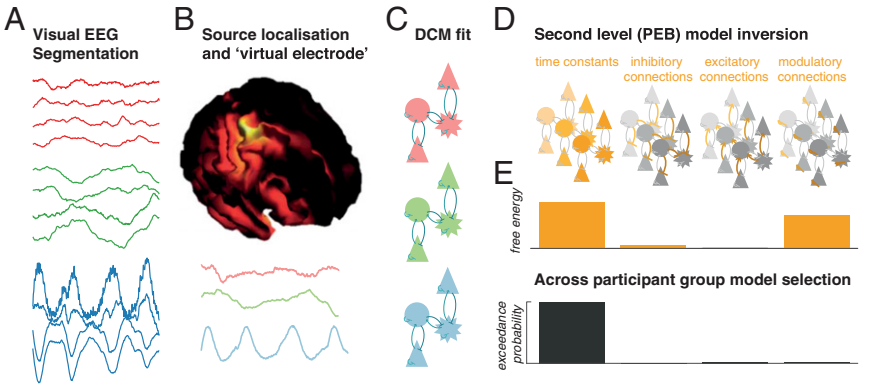

Fig. 7. DCM analysis approach for patient EEG recordings. (*A*) Visual analysis was performed to identify segments of artifact-free background EEG as well as visually apparent paroxysms of abnormal activity (which were further separated into isolated and rhythmic abnormal activity). (*B*, *Upper*) This activity was source localized using an IID approach. (*Lower*) Subsequent modeling was performed using a virtual electrode estimate of LFP activity at the identified source. (*C*) Single-source DCMs comprising a single CMC were fitted separately to power spectral density averages of background and paroxysmal activities. (*D*) PEB was employed to reduce within-subject differences between individual DCMs to specific subsets of parameters. The model space was designed to distinguish between sets of models where time constant, inhibitory connections, excitatory connections, or modulatory connections explained variations among conditions. (*E*) A random effects Bayesian model comparison between these alternative PEB models helped identify which parameters best explain the fluctuations across the whole group of subjects.

Laplace scheme to fit the parameters of a specified neural mass model to empirical data (19) while also providing a free energy measure of the Bayesian model evidence. The combination of posterior parameter estimates and free energy subsequently allows computationally efficient modeling of group effects across individual DCMs, further exploited with the PEB analysis (30).

A second-level model, PEB, was used to estimate parameter changes associated with the experimental modulations. Specifically, each time window was associated with a numerical value representing the absence or presence of NMDAR-Ab (0 or 1, respectively), the estimated PTZ concentration (range 0–1, modeled as first-order kinetics after i.p. injection), and an interaction term (range −1 to 1). PEB employs Bayesian model reduction based on the specified model parameters, effectively modeling between-window changes in parameter as a mixture of random effects and systematic modulation of each parameter by the main effects provided in the PEB model specification. Thus, inversion at the second (between-window) level provides posterior parameter estimates for first-level model parameters (i.e., neuronal physiology) that are associated with second-level parameters (i.e., experimental modulation) across the whole series of individual DCMs.

Comparison between models is based on the free energy approximation of the Bayesian model evidence. We use a Bayesian model reduction approach that is computationally efficient and provides model evidence estimates for a range of different models that differ in terms of the parameters that are free to vary to explain between-window variation in the PEB analysis (30). This approach provides a ranking of how well different combinations of free parameters explain a given dataset (here consisting of between-window changes in power spectral densities) and allows us to identify the most parsimonious model for the observed EEG or LFP effects. Note that models with the highest evidence are those that generalize, in virtue of the fact that model evidence is the difference between accuracy and complexity.

The DCMs are fully specified models of spontaneous neuronal activity and therefore can be used to explore individual parametric effects on overall spectral output. Here, we utilize the parameter estimates derived as the group mean in the PEB analysis as baseline. We then extract the first principal components of time-constant and connection-strength variations across all individual time window DCMs (Fig. 6*E*), providing a summary of covarying changes in parameters that explain most of the variance across samples. We then systematically vary the contribution of each of these two components in 300 discrete steps each around the baseline estimates. This yields $300 \times 300 = 90{,}000$ parameterizations for a single source DCM, and for each of these the spectral output can be estimated. We can use this to visualize scalar output measures (e.g., log mean delta-band power) across a section of a 2D parameter space (Fig. 6*F*). This combines the benefits of fitting generative (i.e., forward) models to empirical data and exploring the effects of specific parameters on model output through forward modeling (60).

In a last step, we implement the microcircuit parameter fluctuations estimated from paroxysmal EEGs in patients in different conditions of the in silico mouse model. Specifically, we

i) Estimate parameter changes that underlie paroxysmal EEG responses in patients (discussed below);
ii) Take the first principal component of the variations of time constants across all participants and EEG states to capture most of the variance of time-constant changes; and
iii) Implement corresponding parameter changes across the range estimated from human EEGs in mouse-derived in silico microcircuit models.

This allows us to simulate the kinds of spectral changes that would be induced if the mouse-derived in silico microcircuits experienced the same (spontaneous) fluctuations in model parameters as observed in human EEGs. We then use an inverse Fourier analysis to illustrate the sort of paroxysmal responses that would be expected based on the spectral predictions under specific parameter combinations (Fig. 6*H*).

Patient Selection and EEG Recording. Patients were selected from routine clinical service at a tertiary pediatric specialist hospital that is a regional referral center for patients with presumed autoimmune encephalitis. Patients were selected based on (*i*) symptoms consistent with autoimmune encephalitis, (*ii*) positive laboratory testing for NMDAR-Abs at some point during their clinical course, (*iii*) the availability of routine clinical EEG recordings during the acute phase of their illness, and (*iv*) the presence of visually apparent EEG abnormalities. Anonymized clinical information was provided by the patients' care team with written, informed consent provided by the patients' legal guardians. All patients met the Graus criteria for a clinical diagnosis of NMDAR-Ab encephalitis (61). Use of anonymized patient data was approved by the United Kingdom Health Regulatory Authority (Application No. 229772).

All EEGs used in this analysis were standard clinical recordings (21 electrodes, International 10–20 electrode layout, 30-min recording time, 256-Hz sampling frequency, 1- to 70-Hz digital Butterworth bandpass filter). EEGs were visually analyzed by two clinicians with expertise in EEG interpretation (R.E.R. and G.C.), identifying paroxysmal abnormalities as well as segments of artifact-free awake background EEGs that were used for further analysis.

DCM Analysis of Patient EEG Paroxysms. EEG analysis was designed to identify mechanisms underlying the frequently observed paroxysmal abnormalities in patients with NMDAR-Ab encephalitis. The purpose of this modeling approach is to identify a small set of parameters that can explain the transition between background activity and EEG paroxysms for each individual patient. The analysis can broadly be summarized as follows (also shown in Fig. 7):

i) Visual identification of paroxysmal and background EEG activity source localization and virtual electrode source wave form extraction;
ii) Fitting single-source DCM to each virtual electrode summary of paroxysmal and background data; and
iii) Inversion of hierarchical (PEB) models explaining within-subject EEG patterns through sets of reduced parameters, which then allows Bayesian model comparison at the group level (random effects analysis).

Patients were selected based on clinical EEGs with reported dynamic abnormalities (ranging from evidence of mild encephalopathy to overt epileptiform activity). EEGs were reviewed by two clinicians with EEG experience (R.E.R. and G.C.), and segments containing normal awake background as well as paroxysmal abnormalities (isolated slow waves, intermittent rhythmic slow activity, and overt epileptiform activity) were identified. Paroxysmal activity was averaged across visually identified 2-s windows and was source localized using an independent and identically distributed (IID) approach in SPM12 (62). At the cortical location with maximal activity, a single virtual electrode trace was extracted for each of the paroxysmal and background activity windows and was used for further DCM analysis (29). The virtual electrode approach reconstructs the time course of signals at points on the cortical mesh of a three-layer head model using an empirical Bayes beamformer

source reconstruction algorithm (63), thus providing LFP-like data that best explain the more distributed activity observed on the scalp (62).

This virtual LFP activity was modeled using a single CMC source. An average of all paroxysm time windows and all background time windows was inverted separately, producing two to three fully specified DCMs per subject. These were subsequently combined into a single hierarchical (PEB) model for each patient in which only a subset of specific parameters was allowed to vary. A model space was created at the level of these second-level models, where either time constants, inhibitory between-population connections, excitatory between-population connections, or inhibitory self modulatory connections were allowed to vary to explain the difference between paroxysms and background activity (see in-text table). Random-effects Bayesian model comparison across these second-level models uses the approximation to model evidence from the variational Laplace model inversion (i.e., the free energy) to compare the evidence for any given model parameterization, given the empirical data (54).

ACKNOWLEDGMENTS. We thank the patients and their families for their contribution to this study and two anonymous reviewers for their constructive feedback on an earlier version of this manuscript. This study was funded by Wellcome Trust Grants 106556/Z/14/Z (to R.E.R.), HMRVOW0 (to S.W.), and 088130/Z/09/Z (to K.J.F.).

1. Crisp SJ, Kullmann DM, Vincent A (2016) Autoimmune synaptopathies. *Nat Rev Neurosci* 17:103–117.
2. Dalmau J, et al. (2007) Paraneoplastic anti-N-methyl-D-aspartate receptor encephalitis associated with ovarian teratoma. *Ann Neurol* 61:25–36.
3. Florance NR, et al. (2009) Anti-N-methyl-D-aspartate receptor (NMDAR) encephalitis in children and adolescents. *Ann Neurol* 66:11–18.
4. Wright S, et al. (2015) N-methyl-D-aspartate receptor antibody-mediated neurological disease: Results of a UK-based surveillance study in children. *Arch Dis Child* 100: 521–526.
5. Dalmau J, et al. (2008) Anti-NMDA-receptor encephalitis: Case series and analysis of the effects of antibodies. *Lancet Neurol* 7:1091–1098.
6. Schmitt SE, et al. (2012) Extreme delta brush: A unique EEG pattern in adults with anti-NMDA receptor encephalitis. *Neurology* 79:1094–1100.
7. Nosadini M, et al. (2015) Longitudinal electroencephalographic (EEG) findings in pediatric anti-N-methyl-D-aspartate (anti-NMDA) receptor encephalitis: The Padua experience. *J Child Neurol* 30:238–245.
8. Moscato EH, et al. (2014) Acute mechanisms underlying antibody effects in anti-N-methyl-D-aspartate receptor encephalitis. *Ann Neurol* 76:108–119.
9. Hughes EG, et al. (2010) Cellular and synaptic mechanisms of anti-NMDA receptor encephalitis. *J Neurosci* 30:5866–5875.
10. Nakazawa K, Jeevakumar V, Nakao K (2017) Spatial and temporal boundaries of NMDA receptor hypofunction leading to schizophrenia. *NPJ Schizophr* 3:7.
11. Cohen SM, Tsien RW, Goff DC, Halassa MM (2015) The impact of NMDA receptor hypofunction on GABAergic neurons in the pathophysiology of schizophrenia. *Schizophr Res* 167:98–107.
12. Wright S, et al. (2015) Epileptogenic effects of NMDAR antibodies in a passive transfer mouse model. *Brain* 138:3159–3167.
13. Pinto DJ, Patrick SL, Huang WC, Connors BW (2005) Initiation, propagation, and termination of epileptiform activity in rodent neocortex in vitro involve distinct mechanisms. *J Neurosci* 25:8131–8140.
14. Baude A, et al. (1993) The metabotropic glutamate receptor (mGluR1 α) is concentrated at perisynaptic membrane of neuronal subpopulations as detected by immunogold reaction. *Neuron* 11:771–787.
15. Landwehrmeyer GB, Standaert DG, Testa CM, Penney JB, Jr, Young AB (1995) NMDA receptor subunit mRNA expression by projection neurons and interneurons in rat striatum. *J Neurosci* 15:5297–5307.
16. Farber NB, Kim SH, Dikranian K, Jiang XP, Heinkel C (2002) Receptor mechanisms and circuitry underlying NMDA antagonist neurotoxicity. *Mol Psychiatry* 7:32–43.
17. Homayoun H, Moghaddam B (2007) NMDA receptor hypofunction produces opposite effects on prefrontal cortex interneurons and pyramidal neurons. *J Neurosci* 27: 11496–11500.
18. Freeman WJ (2000) *Neurodynamics: An Exploration in Mesoscopic Brain Dynamics* (Springer, London).
19. Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *Neuroimage* 19: 1273–1302.
20. Moran R, Pinotsis DA, Friston K (2013) Neural masses and fields in dynamic causal modeling. *Front Comput Neurosci* 7:57.
21. Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys J* 12:1–24.
22. Jansen BH, Rit VG (1995) Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol Cybern* 73:357–366.
23. Gilbert CD, Wiesel TN (1983) Functional organization of the visual cortex. *Prog Brain Res* 58:209–218.
24. Jiang X, et al. (2015) Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* 350:aac9462.
25. Bastos AM, et al. (2015) A DCM study of spectral asymmetries in feedforward and feedback connections between visual areas V1 and V4 in the monkey. *Neuroimage* 108:460–475.
26. Bastos AM, et al. (2012) Canonical microcircuits for predictive coding. *Neuron* 76: 695–711.
27. Harris KD, Shepherd GMG (2015) The neocortical circuit: Themes and variations. *Nat Neurosci* 18:170–181.
28. Papadopoulou M, et al. (2016) Dynamic causal modelling of seizure activity in a rat model. *Neuroimage* 146:518–532.
29. Cooray GK, Sengupta B, Douglas PK, Friston K (2016) Dynamic causal modelling of electrographic seizure activity using Bayesian belief updating. *Neuroimage* 125:1142–1154.
30. Friston KJ, et al. (2016) Bayesian model reduction and empirical Bayes for group (DCM) studies. *Neuroimage* 128:413–431.
31. Boly M, et al. (2012) Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness. *J Neurosci* 32:7082–7090.
32. Carlén M, et al. (2012) A critical role for NMDA receptors in parvalbumin interneurons for gamma rhythm induction and behavior. *Mol Psychiatry* 17:537–548.
33. Gandal MJ, et al. (2012) GABAB-mediated rescue of altered excitatory-inhibitory balance, gamma synchrony and behavioral deficits following constitutive NMDAR-hypofunction. *Transl Psychiatry* 2:e142–e149.
34. Irani SR, et al. (2010) N-methyl-D-aspartate antibody encephalitis: Temporal progression of clinical and paraclinical observations in a predominantly non-paraneoplastic disorder of both sexes. *Brain* 133:1655–1667.
35. Steriade M, McCormick D, Sejnowski T (1993) Thalamocortical oscillations in the sleeping and aroused brain. *Science* 262:679–685.
36. Krishnan GP, et al. (2016) Cellular and neurochemical basis of sleep stages in the thalamocortical network. *eLife* 5:1–29.
37. Fellin T, Ellenbogen JM, De Pittà M, Ben-Jacob E, Halassa MM (2012) Astrocyte regulation of sleep circuits: Experimental and modeling perspectives. *Front Comput Neurosci* 6:65.
38. Beenhakker MP, Huguenard JR (2009) Neurons that fire together also conspire together: Is normal sleep circuitry hijacked to generate epilepsy? *Neuron* 62:612–632.
39. Lopes da Silva FH, Hoeks A, Smits H, Zetterberg LH (1974) Model of brain rhythmic activity. The alpha-rhythm of the thalamus. *Kybernetik* 15:27–37.
40. Csercsa R, et al. (2010) Laminar analysis of slow wave activity in humans. *Brain* 133: 2814–2829.
41. Schmidt H, Petkov G, Richardson MP, Terry JR (2014) Dynamics on networks: The role of local dynamics and global networks on the emergence of hypersynchronous neural activity. *PLoS Comput Biol* 10:e1003947.
42. Koch C, Rapp M, Segev I (1996) A brief history of time (constants). *Cereb Cortex* 6: 93–101.
43. Moreau AW, Kullmann DM (2013) NMDA receptor-dependent function and plasticity in inhibitory circuits. *Neuropharmacology* 74:23–31.
44. Kalueff AV (2007) Mapping convulsants' binding to the GABA-A receptor chloride ionophore: A proposed model for channel binding sites. *Neurochem Int* 50:61–68.
45. Lee V, Maguire J (2014) The impact of tonic GABAA receptor-mediated inhibition on neuronal excitability varies across brain region and cell type. *Front Neural Circuits* 8:3.
46. Paulus W, Rothwell JC (2016) Membrane resistance and shunting inhibition: Where biophysics meets state-dependent human neurophysiology. *J Physiol* 594: 2719–2728.
47. Song I, Savtchenko L, Semyanov A (2011) Tonic excitation or inhibition is set by GABA(A) conductance in hippocampal interneurons. *Nat Commun* 2:376.
48. Kapur A, Lytton WW, Ketchum KL, Haberly LB (1997) Regulation of the NMDA component of EPSPs by different components of postsynaptic GABAergic inhibition: Computer simulation analysis in piriform cortex. *J Neurophysiol* 78:2546–2559.
49. Muñoz W, Tremblay R, Levenstein D, Rudy B (2017) Layer-specific modulation of neocortical dendritic inhibition during active wakefulness. *Science* 355:954–959.
50. Krook-Magnuson E, Armstrong C, Oijala M, Soltesz I (2013) On-demand optogenetic control of spontaneous seizures in temporal lobe epilepsy. *Nat Commun* 4:1376.
51. Jirsa VK, Stacey WC, Quilichini PP, Ivanov AI, Bernard C (2014) On the nature of seizure dynamics. *Brain* 137:2210–2230.
52. Rahman A, et al. (2013) Cellular effects of acute direct current stimulation: Somatic and synaptic terminal effects. *J Physiol* 591:2563–2578.
53. Symmonds M, et al. (2018) Ion channels in EEG : Isolating channel dysfunction in NMDA receptor antibody encephalitis. *Brain* 141:1691–1702.
54. Kiebel SJ, Garrido MI, Moran R, Chen C-C, Friston KJ (2009) Dynamic causal modeling for EEG and MEG. *Hum Brain Mapp* 30:1866–1876.
55. Moran RJ, Stephan KE, Dolan RJ, Friston KJ (2011) Consistent spectral predictors for dynamic causal models of steady-state responses. *Neuroimage* 55:1694–1708.
56. Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M (2013) Inhibition of inhibition in visual cortex: The logic of connections between molecularly distinct interneurons. *Nat Neurosci* 16:1068–1076.
57. Lopes da Silva F (1991) Neural mechanisms underlying brain waves: From neural membranes to networks. *Electroencephalogr Clin Neurophysiol* 79:81–93.
58. Thomson AM, Bannister AP (2003) Interlaminar connections in the neocortex. *Cereb Cortex* 13:5–14.
59. Schubert D, Kötter R, Zilles K, Luhmann HJ, Staiger JF (2003) Cell type-specific circuits of cortical layer IV spiny neurons. *J Neurosci* 23:2961–2970.
60. Lytton WW (2008) Computer modelling of epilepsy. *Nat Rev Neurosci* 9:626–637.
61. Graus F, et al. (2016) A clinical approach to diagnosis of autoimmune encephalitis. *Lancet Neurol* 15:391–404.
62. Litvak V, et al. (2011) EEG and MEG data analysis in SPM8. *Comput Intell Neurosci* 2011:852961.
63. Belardinelli P, Ortiz E, Barnes G, Noppeney U, Preissl H (2012) Source reconstruction accuracy of MEG and EEG Bayesian inversion approaches. *PLoS One* 7:e51985.