Mechanistic Inference of Brain Network Dynamics with Approximate Bayesian Computation

Timothy O. West[*1,2,3], Luc Berthouze[4,5], Simon F. Farmer[6,7], Hayriye Cagnan[1,2,3], Vladimir Litvak[3]

[1.] *Nuffield Department of Clinical Neurosciences, Medical Sciences Division, University of Oxford, Oxford OX3 9DU*

[2.] *Medical Research Council Brain Network Dynamics Unit, University of Oxford, Oxford, OX1 3TH, United Kingdom.*

[3.] *Wellcome Trust Centre for Human Neuroimaging, UCL Institute of Neurology, Queen Square, London, WC1N 3BG, UK.*

[4.] *Centre for Computational Neuroscience and Robotics, University of Sussex, Falmer, UK.*

[5.] *UCL Great Ormond Street Institute of Child Health, Guildford St., London, WC1N 1EH, UK.*

[6.] *Department of Neurology, National Hospital for Neurology & Neurosurgery, Queen Square, London WC1N 3BG, UK.*

[7.] *Department of Clinical and Movement Neurosciences, Institute of Neurology, Queen Square, UCL, London, WC1N 3BG, UK.*

**Corresponding Author*

Abstract

Brain networks and the dynamics that unfold upon them are of vital importance at many scales of systems neuroscience. The parameterization of generative models from empirical data (inverse modelling) has become of great utility to theoretical understanding in this domain. However, it has become difficult to infer mechanisms of spontaneous neural activity with existing tools due to the adoption of simplifying model assumptions that appeal to steady-state approximations of neural dynamics. In this work we describe a framework employing Approximate Bayesian Computation (ABC) for the inversion of neural models that can be used to flexibly model any statistical feature of empirically recorded data without assuming any specific form of dynamics. Further, we demonstrate how Bayesian model comparison can be applied to ABC fitted models to facilitate the selection of competing dynamical hypotheses. This work provides a validation of this approach in the context of a

Parkinsonian electrophysiology, using a set of multisite recordings from Parkinsonian rats to demonstrate convergence, consistency, and uniqueness of parameter estimation; face validity of the model comparison procedures; and its utility in answering biologically relevant problems. From the validation and example applications presented here we conclude that the proposed framework is likely to be of value to a wide range of researchers aiming to explain how brain dynamics emerge from neural circuits.

Keywords

Inverse modelling, networks, brain dynamics, oscillations, circuits

1 Introduction

Networks of connected brain regions form the scaffold upon which a diversity of neural dynamics unfold, even in the absence of structured sensory input or task demand [1–4]. A growing body of evidence shows that spontaneous brain activity can explain variation in task performance [5–7]; bias sensory perception [8–10]; modulate movement [11,12]; and encode working memory [13]. In addition, this same resting state activity is disturbed in several pathologies such as Parkinson's disease, Alzheimer's disease, and schizophrenia [14–19]. These experimental findings strongly motivate theoretical work to understand how networks shape spontaneous brain activity [20–22].

From our knowledge of neural transmission, the spatiotemporal patterning of brain activity is largely dictated by: a) the biophysical properties of the neurons acting as the sources and modulators of activity; and b) the organization of axonal tracts that form the network on which this activity propagates. Extensive mapping of white matter pathways has led to a characterization of the brain's structural connectome [23] that provides the substrate for spontaneously organized activity [24]. Functional connectivity encapsulates a range of data driven approaches to network inference directly from the correlations that emerge between connected brain regions [2,25]. Despite significant work unveiling both structural and functional connectomes, computational and experimental evidence has demonstrated that there is no straightforward one-to-one mapping between the two [26–28].

Forward modelling tools such as The Virtual Brain (Ritter et al. 2013) provide an alternative to the data-lead approaches discussed above. These methods aim to describe experimentally observed activity in terms of computational models and their simulations. This typically involves models that include a wide range of biophysical nonlinearities that yield simulated activity replete with the spatiotemporal richness of observed brain signals. These include: far-from equilibrium/critical dynamics [27,30]; multi- or meta- stable states [31–33]; and stochastic switching [34] to name but a few.

However, both data-driven and forward modelling approaches exhibit shortcomings. Data driven methods (i.e. functional connectivity) are agnostic to the underlying generative process and so no inference is made as to mechanisms, only hierarchies of relationships between putative neural sources. On the other hand, forward modelling approaches incorporate large amounts of biological detail and often involve high dimensional models with large numbers of unknown parameters. This complexity often makes it necessary to take an exploratory approach to system identification but without any formal quantification as to how well a model can explain a given set of experimental data.

In order to formalize model parameterization and identification, *inverse modelling* approaches such as Dynamical Causal Modelling (DCM; [35]) use model optimization to estimate a set of latent parameters that best reproduce the summary statistics of some experimentally observed data. Furthermore, DCM adopts a form of Bayesian optimization that allows for the deployment of Bayesian model comparison in order to make an empirically informed decision as to the best candidate hypothesis given some data [36,37]. Whilst powerful, inverse modelling invokes a parameter search that requires numerous evaluations of a given model. To render inference tractable, methods such as DCM take steps to reduce the complexity of the models to be inverted. For instance, the most commonly used DCM for electrophysiological data fits directly to the cross- and auto-spectra of a time series by invoking simplifying reductions to the model dynamics i.e. that the measured activity is generated from a quasi-stationary state that can be approximated to have a unitary and stable transfer function [38]. It is important to note that this simplification of the candidate models *removes* a large set of dynamical regimes in which the neural system *could* potentially operate.

In this paper we introduce a novel framework for the inference of brain networks from neural data. This approach draws inspiration from DCM but uses an optimization scheme based upon Approximate Bayesian Computation (ABC) that makes it possible to retain complex, high dimensional models without invoking the quasi-stationary approximation. ABC is a method of stochastic inversion that avoids the requirement of having an analytically tractable or estimable likelihood function, a major determining factor in the applicability of many schema used for parameter estimation in biological dynamical systems (such as Variational Bayes or Markov Chain Monte-Carlo). It is thus well suited for systems which are expensive to evaluate due to a large state and/or parameter space, or require numerically expensive integration schemes to solve [39]. Furthermore, we investigate schemes in which models can be flexibly fit to data features other than the auto- and cross-spectra; in the examples shown here we use Non-Parametric Directionality Analysis (NPD)- an estimator of directed functional connectivity between time series.

NPD has recently been used to estimate the changes in the functional connectivity of subcortical networks following the dopamine depletion associated with Parkinsonism [40] in order to assess a number of hypotheses regarding the emergence of pathological beta (14-30 Hz) rhythms associated with

the disease [41–43]. Parkinsonian dynamics have been the subject of numerous past modelling studies [44–48]. To date, the majority of models have investigated the emergence of pathological beta band power (however, see 49,50), whereas recent findings have implicated burst like intermittencies in these rhythms as a hallmark of the disease state [12,51–54].

The combination of: a) a well-defined anatomical network [55–57]; and b) a well characterised electrophysiological feature of the disease (i.e. beta rhythms) makes the cortico-basal ganglia system well suited to examining structure and activity relationships. In the present work we will retain steady state data features (spectra and directed functional connectivity) as the basis on which to invert models of connectivity but use a framework that does not restrict model dynamics. Thus, complex features such as intermittency in oscillations (e.g. Parkinsonian beta bursts) contribute to the model output and are open to be explored in post-hoc simulations of the fitted system.

In this paper we formulate a generic framework based on ABC for the inversion of forward models of neural dynamics and model selection. We will use Parkinsonian network dynamics as a context in which to set up example implementations and validation. This work provides to our knowledge the first application and validation of the ABC framework to large scale neuroscience by demonstrating convergence of models upon the data, consistency of the optimization, uniqueness of solutions, robust system identification, application of model comparison to biological networks, and demonstrate the scalability of the platform to investigate larger and increasingly complex model spaces. Additionally, we provide a post-hoc analysis of simulations from fitted models and compare their temporal dynamics (specifically beta bursts) with that measured in empirical data.

2 Results

2.1 Approximate Bayesian Computation for Model Inversion and Comparison

2.1.1 Formulation of the Framework

The main results of this paper assess the validity of a framework utilizing Approximate Bayesian Computation (ABC; 58,59) for the inference of mechanistic models of neural connectivity. In the examples presented here we use coupled neural mass models [60,61] to describe the population activity of a network with a structure similar to that used in previous studies investigating the cortico-basal-ganglia-thalamic network [44–46]. We however adapt the state equations such that they explicitly incorporate delays and stochastic inputs that can be integrated to provide time domain simulations that are equivalent in form to the original field potential recordings (see methods and supplementary information for details). This renders their solutions ill-suited to inversion with deterministic optimization schemes such as Variational Bayes. Instead we use a method based on sequential ABC

[59,62–64] that eschews the estimation of likelihood functions required by other stochastic methods such as Markov Chain Monte-Carlo or Gibbs sampling.

ABC provides a method of likelihood free inference [65]. In the case where a forward model of some empirical data D is known with sample parameters drawn from some prior distribution $\theta^* \sim P(\theta)$ to simulate sample *pseudo-data* $D^* \sim P(D|\theta^*)$, it is possible to make approximate inference by rather than sampling from the exact posterior $P(\theta|D)$, instead sampling from something that lies approximately close i.e. $P(\theta|\|D^* - D\| < \epsilon)$. This tolerance ϵ can then be interpreted as a model of measurement error. ABC in its most basic form uses a rejection sampler to approximate the posterior, however we adopt a sequential form in which the approximate posterior is reached by a sequence of intermediate proposal distributions with shrinking ϵ [64]. Further, rather than fit directly by the distance of data from pseudodata, we compactly represent the data by some summary statistic $s(x)$ such that the algorithm presented here samples from $P(\theta|\|s(D^*) - s(D)\| < \epsilon_q)$ over Q iterations. For full technical details of the algorithm please see the methods section of this paper.

An overview of the proposed framework is presented in figure 1. This schema allows for the inference of latent parameters of dynamical models directly from empirical time series data. The algorithm takes a form in which several nested processes are repeated multiple times within their parent process (figure 1; inset). The schema is contingent on simulation of pseudo-data by the generative model – a description of the neural dynamics - (figure 1A; green box) given a set of proposal parameters sampled from a prior distribution (figure 1C; turquoise box). This pseudo-data can then be compared against the empirical

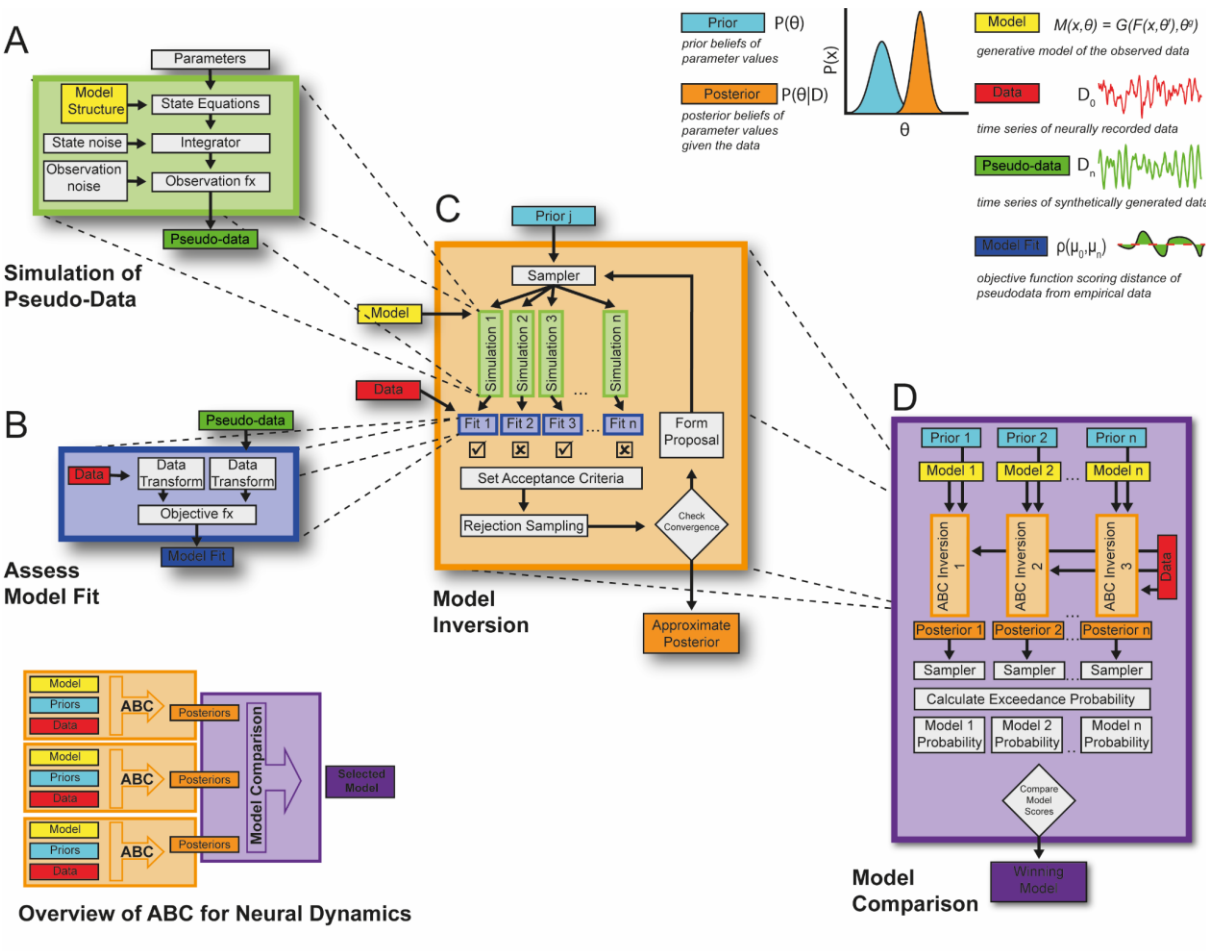

Figure 1 – **Framework for application of Approximate Bayesian Computation for mechanistic inference of brain network dynamics. (Inset)** This schematic gives an overview of the structure of the framework proposed and validated in this paper. Individual generative models (yellow boxes) are specified as a set of state equations and prior distribution of parameters (light blue boxes) that will be used to approximate the parameters of system generating the observed data (red boxes) with varying degrees of fit (blue boxes) using ABC. This inversion scheme estimates a posterior distribution of parameters (orange boxes) that can then be used to compare models and decide on a winning model or family of models (purple box). **(A) Generation of pseudo-data** by integrating state equations parameterized by a sample of parameters drawn from a prior or proposal distribution. Models can incorporate stochastic innovations as well as a separate observation model to produce samples of pseudo-data (green box). **(B) Pseudo-data is compared against the real data** using a data transform common to both that provide a summary statistic of the time series data (i.e. spectra and functional connectivity). The simulated and empirical data are then compared with some objective function (e.g. R2 or RMSE) that can be used to score the model fit (blue box). **(C) ABC sequentially repeats the processes in boxes A and B** by iteratively updating a proposal distribution formed from accepted samples. Samples are rejected depending on an adaptive thresholding of the objective scores with the aim to reduce the distance between summary statistics of the data and pseudo-data. This process iterates until the convergence criterion is met and the proposal distribution is taken as an approximation of the posterior distribution. **(D) By repeating the ABC process in box (C) over multiple models**, the approximate posteriors can be used to evaluate the model probabilities. This process samples from the posterior many times to compute the probability of each model exceeding the median accuracy of all models tested. This "exceedance" probability can then be used to compare the model's ability to accurately fit the data and select the best candidate model given the data.

data using a common data transform (i.e. a summary statistic of the data) that allows the fit of the model to the data to be assessed (figure 1B; blue box). This model fit provides a measure by which parameter samples are either rejected or carried forward to generate a proposal distribution depending on some threshold model score. When the process in figure 1C is iterated with a shrinking tolerance schedule, ABC can be used to approximate the posterior parameter distribution at convergence (figure 1C; orange box). Finally, if the process described above is repeated over several competing models then the approximate posterior distribution may be used to assess each model's fitness via model comparison (figure 1D; purple box). The exact details of each process outlined in the figure are given in the relevant methods sections of this paper.

2.2 Convergence Properties of ABC Applied to a Simple Model of the Pallido-Subthalamic Subcircuit

We provide an example model inversion in figure 2 to demonstrate that the ABC algorithm converges to yield close approximations to the empirical summary statistics. This example uses a simple forward model of a basal-ganglia subcircuit comprising the reciprocally connected subthalamic nucleus (STN) and external segment of the globus-pallidus (GPe) shown in figure 2A. Specifically, this model was fit to the autospectra and NPD (an estimate of directed functional connectivity with results originally reported in [40]; see methods) estimated from the empirical data (local field potentials recorded from a rodent model of Parkinsonism; see methods for experimental details). By tracking the value of the evaluated objective function (estimated with R^2) over iterations (figure 2B) we demonstrate a fast-rising initial trajectory in the first 15 iterations that eventually plateaus towards convergence. This iteration history is well fit by a logistic function (shown by purple dashed line). In figure 2C and D the simulated features (autospectra and NPD respectively) gradually move closer to the empirically estimated features with each iteration of the algorithm. By the 25th iteration, the optimization reaches convergence criteria (change in $R^2 < 0.002$) and posterior densities over parameters are stable and unimodal. Overall, fits reach 40% explained variance for the autospectra (figure 2C; $R^2 = 0.41$) and 25% for the functional connectivity estimates (figure 2D; $R^2 = 0.25$).

The evolution of the proposed marginal densities (figure 2E) demonstrates that over the optimization, parameter means, and variances deviate significantly from the prior. Estimation of some parameters is better informed by the data than for others, as indicated by the different precision of the proposal densities (for example the parameter density outlined in blue exhibits reduced posterior precision). Additionally, learnt multivariate structure in the joint parameter densities is apparent from the correlation matrices (figure 2F). As all priors were initiated with diagonal covariance, any significant dependencies between parameters were informed by the data. The evolution of these matrices shows the emergence of distinct correlated modes. For instance, the parameters numbered 6, 7, and 8 (indicating STN endogenous noise; STN → GPe strength; and GPe → STN strength respectively) are

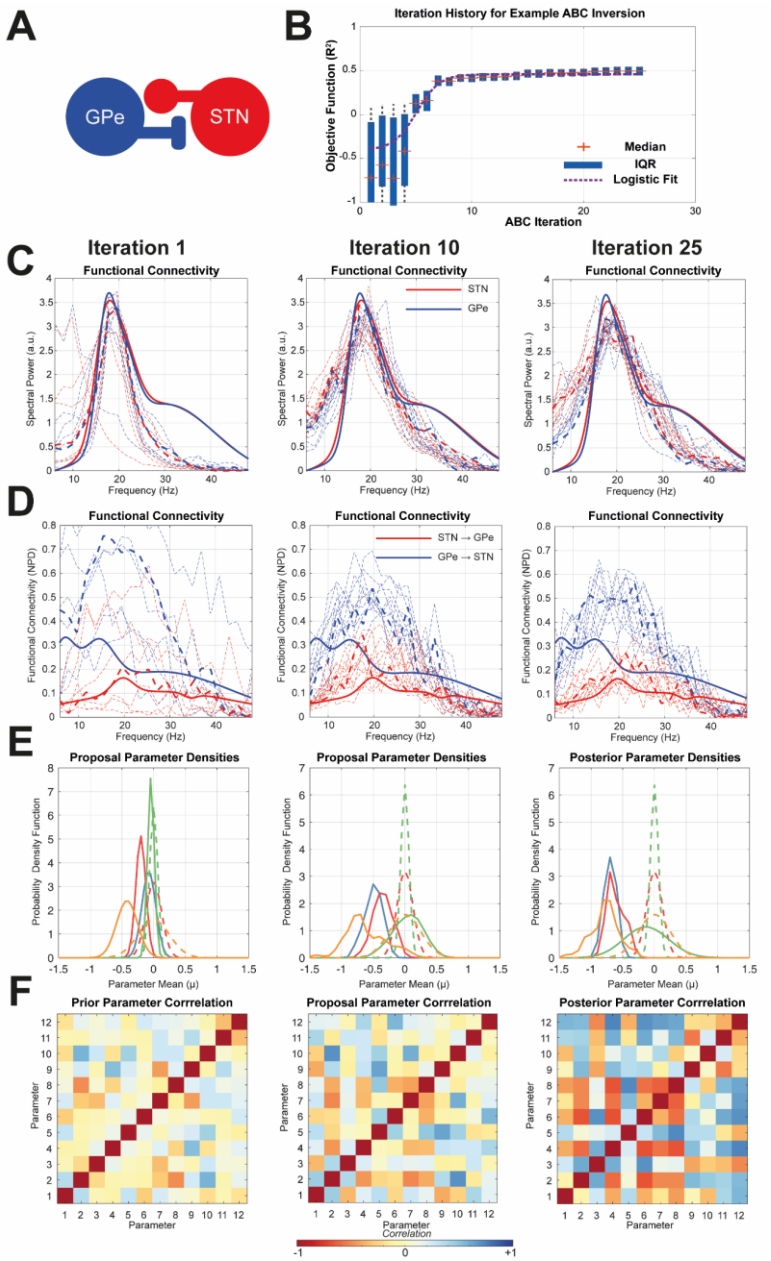

Figure 2 – **Examining the convergence of ABC optimization upon summary statistics from recordings of the STN and GPe in Parkinsonian rats –** Parameters of neuronal state space models were optimized using the ABC method detailed in the text. Snapshots of the optimization are taken at the 1st, 10th, and 25th iteration at which the optimization converges. **(A)** Schematic of the STN/GPe neural mass model. **(B)** The iteration history of the ABC algorithm is presented as a sequence of box plots indicating the distribution of fits (R^2) at each sampling step, with mean and interquartile range indicated by individual crosses and boxes. Optimization shows a logistic convergence with a fit indicated by purple dashed line. **(C)** Autospectra of the empirical data (bold) and simulated data (dashed) are shown. The best fitting parameter sample for each iteration is given by the bold dashed line. **(D)** Similarly, the functional connectivity (non-parametric directionality; NPD) is shown in red and blue with the same line coding. **(E)** Examples of the prior (dashed) and proposal (bold) marginal distributions for a selection of five parameters are shown (note some priors have identical specifications and so overlap). It is seen that over iterations the proposal and posterior deviate from the prior as the latent parameter densities are estimated. **(F)** Correlation matrices from copula estimation of joint densities over parameters. Colour bar at bottom indicates the correlation coefficient. Correlated modes appear between parameters as optimization progresses.

bound into a correlated mode. The appearance of correlated modes acts to effectively reduce the dimensionality of the optimization problem: when estimating effective dimensionality of the parameter space by estimating the number of significant principle components (see methods) we find that optimized models show a reduction of 50-70% from that of the prior.

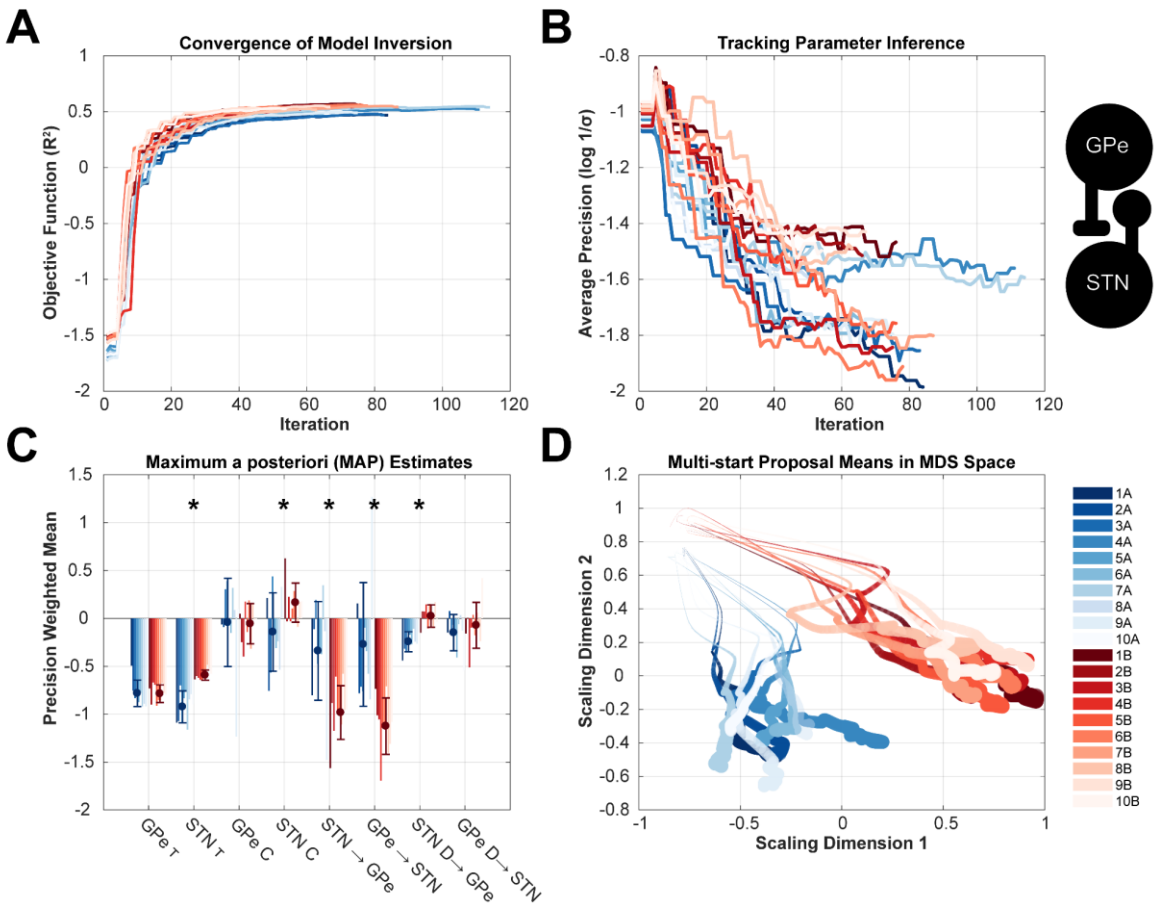

Figure 3 – **Multi-start analysis to test the consistency of the ABC optimization of model parameters and separability by empirical data.** A two-node model of the STN/GPe circuit (inset) was fitted to two different data sets: A (blue) and B (red). Each estimation was performed 10 times with identical specification of prior distributions for all initializations. **(A)** Tracking of the objective function (R^2) over the iterations demonstrated consistent convergence to 45-52% explained variance for both datasets. **(B)** Optimization shows a consistent reduction in the average precision (logarithm of the inverse standard deviation of the data) of the posteriors indicating that data was informative in constraining parameter estimates. **(C)** Examination of the MAP estimates (weighted by their inverse precision) demonstrates a consistent inference of parameter values. Some parameters are drawn to common values with both data A and B (e.g. GPe time constant), whilst others show differences informed by the data (e.g. STN time constant). MAP values are scaled as log scaling parameters of the prior mean. The prior values are set to equal zero. Error bars are the standard deviations of the weighted estimates across initializations. Asterisks indicate significant t-test for difference in means between parameters estimated from data A and B. **(D)** To visualize trajectories of the multistarts, the high dimensional parameter space was reduced to 2D using multi-dimensional scaling. Evolutions of the means of the proposal parameters exhibit a clear divergence between data sets A and B.

2.3 Demonstrating Estimator Consistency and Uniqueness with Multi-start

This section of the results addresses two concerns regarding the validity of the proposed framework: (i) the consistency of the ABC scheme in its estimation of model parameters across multiple instances of the optimization, and (ii) the uniqueness of posterior parameter estimates upon data arising from different generative models. This is achieved using the multi-start approach [66] described in the methods (section 4.6.2) in which we perform 10 instances of the algorithm on upon two separate datasets. The results of the multistarts are shown in Figure 3.

The evolution of the objective function (R^2) over the progress of the ABC optimization is presented in figure 3A. This demonstrates that all the multi-starts converge to ~50% explained variance for both sets of data. The value of R^2 reached by convergence is 0.55±0.02 for dataset A and 0.51±0.02 for dataset B showing consistency in the accuracy at convergence across the starts. In figure 3B, the average log-precision of the marginal densities is tracked over the progress of the optimization. These data show that across all initializations, the average precision of the posterior densities ($1/\sigma = 0.025$) was four times smaller than the initial priors at $1/\sigma = 0.1$ demonstrating that parameters estimates were well constrained by the data.

In figure 3C, we present the maximum a posterior (MAP) estimates for each parameter across the multi-starts. The posterior means were weighted by their precision (as described in the methods) such that parameters that were poorly informed by the data tended towards zero. In figure 3C, there are clear differences between parameters inverted upon the two separate sets of data (red versus blue bars; asterisks indicate significant t-tests). For instance, the mean STN time constant (2nd group of bars from the left) is smaller for data A compared to B. Other parameters were well informed by the data, but not different between either data sets (e.g. GPe time constant; 1st set of bars from the left). A third category of parameters were poorly inferred from the data (e.g. input gain to GPe; 3rd set of bars from the left) where both estimates are close to zero deviation from the prior mean.

Finally, we apply statistical testing to the MAP estimates to determine (i) estimator consistency and (ii) differentiability with respect to empirical data. To determine (i): estimator consistency, we applied a one-sample Hotelling test within a 10-fold, leave-one-out cross-validation to each set of MAP estimates from the multi-start. For both samples of parameters estimated from data A and B we find there to be a 0% rejection of the null hypothesis that the mean of the fold is significantly different from that of the left-out sample. This suggests that samples are consistent between initializations. To test (ii): that the posteriors were differentiated by the data on which they were estimated, we apply the Szekely and Rizzo energy test. We find there to be a significant difference in the means of the two samples ($\Phi = 6.68$; P =

0.001). This finding is supported by a MANOVA test that demonstrates that the two data sets are significantly segregated by their posterior parameter means (D = 1, P < 0.001).

Visualization of the parameter space using multidimensional scaling (MDS) confirms the segregation of the posterior samples into two clusters determined by the datasets from which they are estimated. These results confirm that the ABC optimization is consistent across initializations and that its output is determined by the empirical data.

2.4 Face Validation of the Model Comparison Approach

To verify that the model comparison approach provides a valid identification of the original generative model of the data (c.f. face validity) we constructed a confusion matrix (as detailed in the methods section 4.6.3) using variations on the STN/GPe model presented in the previous sections and shown in figure 4A. In the case of validity, the analysis should correctly identify the system that originally generated the data (i.e. the best model scores should lay along the diagonal of the confusion matrix).

In figure 4B we present the estimated posterior model probabilities *P(M|D)* derived from exceedance probabilities computed on the joint model space (see methods). This analysis demonstrates that, in terms of accuracy, the most probable models lie on the diagonal of the confusion matrix showing that the comparison of models by their posterior accuracies is enough to correctly identify the generating model. In figure 4C, analyses show that the divergence of each model's posteriors from priors (so called *complexity* or D_{KL}) remain minimized along the diagonal such that best fitting models in panel A are not resulting from overfitting. In the case of model 1 (which is the most flexible in terms of numbers of

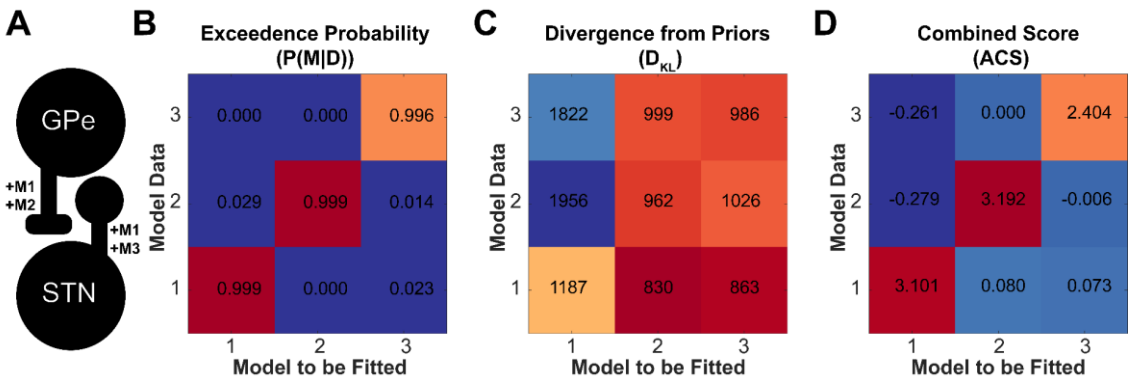

Figure 4 – **Testing face validity of the ABC model comparison approach to model identification.** Confusion matrices were constructed by fitting the three models of the STN/GPe circuit. Synthetic data was generated using the fitted models and then the three original models were fitted back to the synthetic data to test whether model comparison could identify the generating model. **(A)** Schematic of neural mass model to be fitted. Annotations of connections indicate the presence of each for models 1-3. **(B)** Matrix of posterior model probabilities *P(M|D)* computed as exceedance probabilities across the joint model space. **(C)** Matrix of D_{KL} divergences of posteriors from priors. When $P(\theta|D_0) = P(\theta)$ then the divergence is zero thus larger values indicate a bigger shift from the prior. **(D)** Combined scoring to simultaneously account for model accuracies and divergence (ACS). Large values indicate better fits with parsimonious posteriors (small D_{KL}).

free parameters) there are inflated divergences in the first column. These are the result of a large deviation of posteriors when attempting to fit the data generated from the alternative models. This shows that a post-hoc analysis of model parameter deviation (using KL divergence) can be used to discriminate models which have been overfitted. When combining these two measures into the ACS metric (accuracy minus complexity) in figure 4D, it is seen that the best fitting models are still correctly identified even when accounting for the increased complexity of posterior parameter densities. These results demonstrate that the model comparison approach can properly identify models from which the data originated, thus providing a face validation of the model comparison procedures.

2.5 Utility of Model Comparison for Biological Problems: Investigating Hypotheses of the STN/GPe Circuit

Having established the face validity of the model comparison approach, we next demonstrate its application to estimating biological parameters within the STN/GPe circuit (figure 5). We again use the three models of the circuit introduced in section 2.3 and test their ability to explain the observed autospectra and directed functional connectivity. Namely, we test a set of three models (figure 5A) with: 1) reciprocal coupling of the STN/GPe; 2) predominant STN excitation; and 3) predominant GPe inhibition. Examples of the simulated time series generated from the inverted models are shown in the bottom row of figure 5A. The simulated data features resulting from the posterior distribution over parameters are shown in figure 5B. We demonstrate that given suitable parameterization, the reciprocal model (model 1) can generate beta oscillations that exhibit intermittent burst-like behaviour (figure 5A; blue trace) similar to that seen *in vivo* [12,67]. Please see results section 2.7 for a fuller analysis of simulated bursts in comparison to empirical and surrogate data.

The summary statistics of the inverted models are shown alongside that of the data in figure 5B. It is apparent that model 1 (blue) can generate good fits to the autospectra (R^2 = 0.54) and functional connectivity (R^2 = 0.17) estimated from the empirical recordings (black). Neither of the other two models give good fits to either autospectra (R^2 = -2.5 for model 2 in red**;** R^2 = -1.6 for model 3 in green; please see methods for interpretation of negative values of R^2) or functional connectivity (R^2 = -6.1 for model 2**;** R^2 = -0.5 for model 3).

In this simple example it is clear from the posterior distributions of the R^2 values in figure 5C that model 1 is the best fitting model and corresponds well to the known biology of the circuit [68]. We however demonstrate that the model comparison framework can yield a quantitative assessment of the relative model performances. In figure 5C the distributions of the model fits are shown as violin plots of the model accuracies (in terms of the R^2 of the empirical data features). Model 1 has a distribution that far exceeds the other two models in terms of its accuracy.

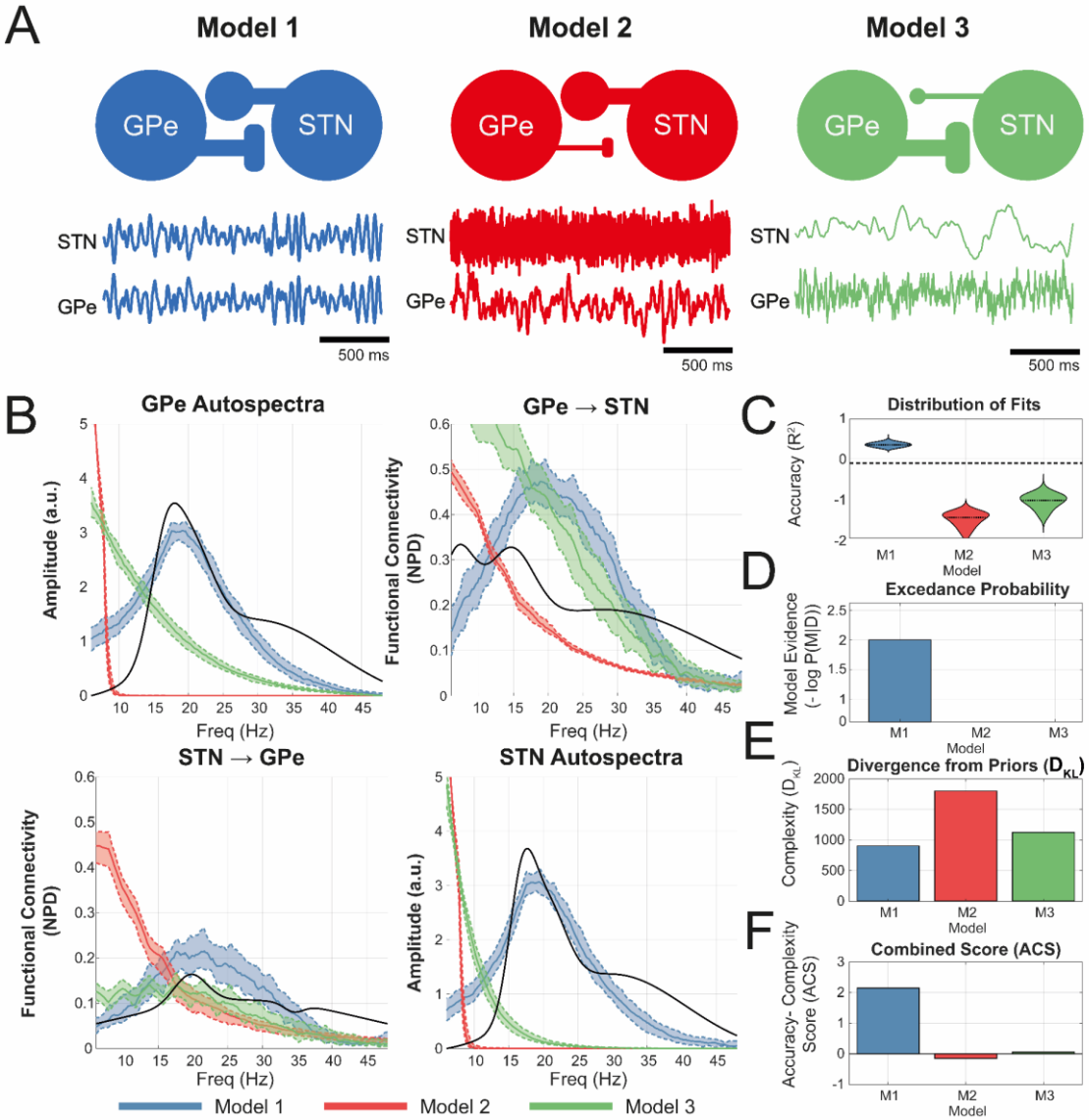

Figure 5 – **Application of ABC inversion and model comparison to competing models of the STN/GPe circuit. (A)** Three competing models of the STN/GPe circuit were fitted using ABC. Model 1 (blue) has reciprocal coupling whilst models 2 (red) and 3 (green) show coupling biased towards STN excitation or GPe inhibition respectively. Two-second samples of simulated time series generated from the inverted models are given below the model diagrams. **(B)** The summary statistics of the data are presented with autospectra along the diagonal and directed functional connectivity on the off-diagonal. The empirical data are shown in black whilst the posterior model fits are given by the colour coding described in the legend. **(C)** Violin plots of the distribution of R^2 values computed from 1000 draws from each of the posterior distributions over parameters. The median accuracy is given by the dashed line. **(D)** Bar plots of the estimates of the negative posterior model probabilities as calculated by the probability of model accuracies exceeding that of the median R^2 of the whole model space. **(E)** Bar plot of the complexity associated with each fitted model estimated using the KL divergence between the posterior and prior. **(F)** Bar plot of the accuracy-complexity score (ACS) to combine model probabilities (computed from goodness-of-fits) with the model divergence (computed as the divergence of priors from posteriors).

Using the median R^2 of all models as a common baseline, the exceedance probability is computed as an approximation to the model evidence (figure 5D). Model 1 shows a P(M|D) = 0.98. A more interpretable comparison can be made by computing the Bayes factor between Model 1 and 2 (the

next best fitting model). This yields K = 17.8, equivalent to *substantial* evidence (in terms of the Jeffrey's interpretation of K) of model 1 over that of the next best fitting model. To determine the degree of overfitting, we compute D_{KL} of posteriors from priors which we term model complexity (figure 5E). This analysis shows that the best fitting model also has a smaller complexity associated with its posterior distribution over parameters indicating it as the more parsimonious of hypotheses. Both remaining models (2 and 3) exhibit a larger D_{KL} indicating that posterior parameter estimates have shifted significantly from the prior estimates. The ACS score combining the accuracy and complexity (figure 5F) also confirm model 1 as the best candidate hypothesis of the data.

2.6 Scaling up ABC based Model Comparison: Application to Models of the Cortico-Basal Ganglia-Thalamic Circuit

Finally, we apply the ABC framework to a larger and more complex model space to test the scalability of the methodology. Specifically, we devise a set of 12 models (illustrated in figure 6) incorporating combinations of pathways in the cortico-basal ganglia-thalamic circuit amongst a set of six neural populations motor cortex (M2); striatum (STR), GPe, STN, and thalamus (Thal.). Models are split into sets including/excluding the indirect (M2 → STR → GPe → STN); hyperdirect (M2 → STN); and thalamocortical relay (M2↔Thal.). Models are further subdivided to include or exclude the subthalmo-pallidal feedback connection (STN → GPe; models prefixed M *x*.2 to denote inclusion of the connection). For a full description and defence of model choices please see methods. These models were fit individually to the empirical data and then model comparison used to determine the best candidate model.

In figure 7 we show the resulting model fits and then the subsequent model comparison done in order to determine the best model or set of models. From visual inspection of the fits to the data features in figure 7A as well as the distribution of posterior model accuracies in figure 7B there is a wide range of model performances with regards to accurate fitting of the models to the data. Inspection of the posterior model fits to the data features in 7A shows that even in the case of the best fitting model (M 5.2) fitting fails to account for the multiple peaks in the autospectra, with models showing a preference for fitting the largest peak around 20 Hz. There is also a systematic overestimation of the directed functional connectivity (NPD) from cortex to subcortex. When the R^2 of the fitted models are segregated between their autospectra and functional connectivity, the latter is more accurately fitted (M 5.2; $R^2 = 0.45$) than for the power (M 5.2; $R^2 = -0.3$). It is also found that the simplest models in families 1 and 2 can fit the data to a reasonable degree ($R^2 \sim 0.2$) although inspection of the distributions of model accuracies in 7B indicates that poor posterior precision over model parameters means that model fitness is variable.

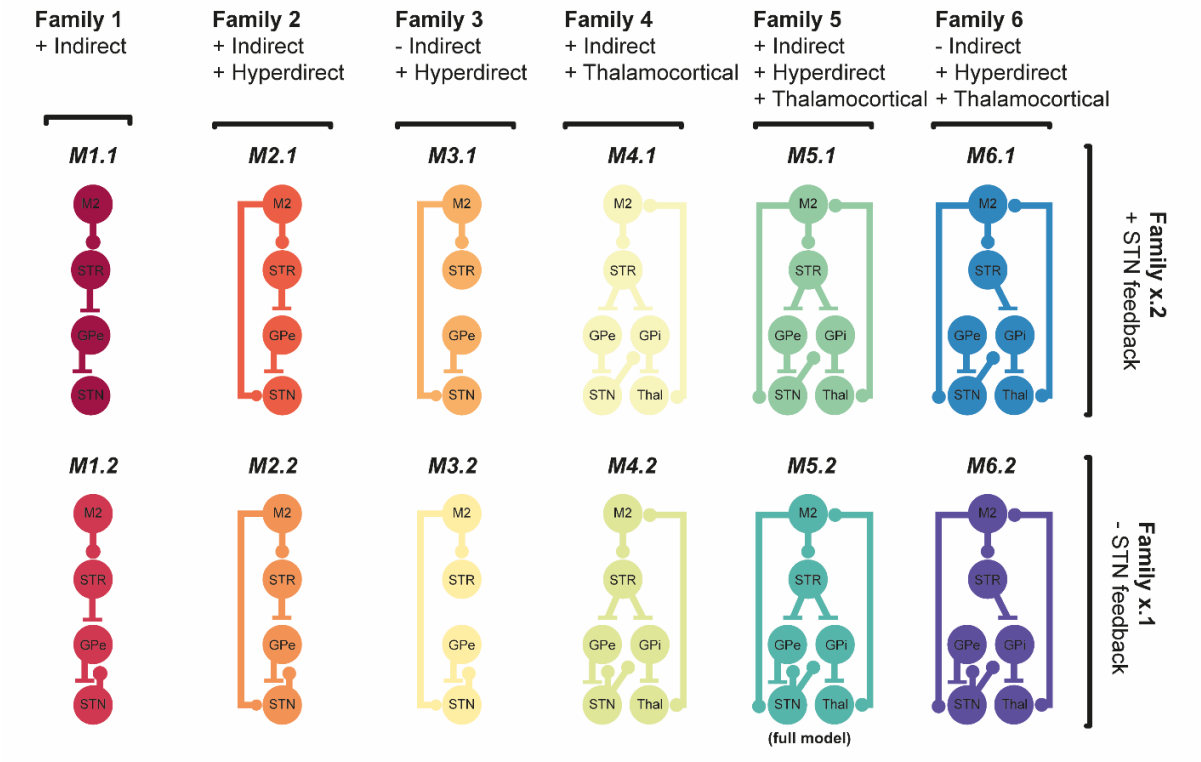

Figure 6 – **Illustration of the model space of the cortico-basal-ganglia network fitted with ABC and compared with Bayesian model selection.** The model space comprises six families which can be further subdivided into two subfamilies yielding 12 models in total. Family (1) models the indirect pathway; family (2) contains models with both the indirect and hyperdirect pathways; family (3) contains models with the hyperdirect pathway but not indirect pathway; family (4) contains models with the indirect and thalamocortical relay; family (5) contains models with indirect, hyperdirect, and thalamocortical pathways; family (6) contains models with hyperdirect and thalamocortical pathway but indirect pathway. Finally, each family comprises two sub-families that either exclude (Mx.1) or include (Mx.2) subthalamopallidal feedback excitation. Excitatory projections are indicated by ball-ended connections, whilst inhibitory connections are flat-ended.

In all cases the models containing the subthalamopallidal excitatory connection (M *x*.2) perform better than those without (M *x*.1) and in good agreement with known Parkinsonian electrophysiology [68]. Notably we find that the model families 3 and 6 that both lack indirect pathway connections perform poorly with many of the posterior distributions of model fits falling far below the median of the whole model space. In the case of M 6.1 where there is a non-negligible exceedance probability, we see this accompanied by a high KL divergence and so both are penalized when measured with the combined ACS scoring. M 4.2 and 5.2 are the strongest models, with distributions of fits tightly clustered around high values yielding high model evidences. This suggests the importance of including thalamocortical feedback connections in the model. Scoring with ACS suggests model 4.2 is the best of the models due to the smaller model complexity (parameter divergence from prior).

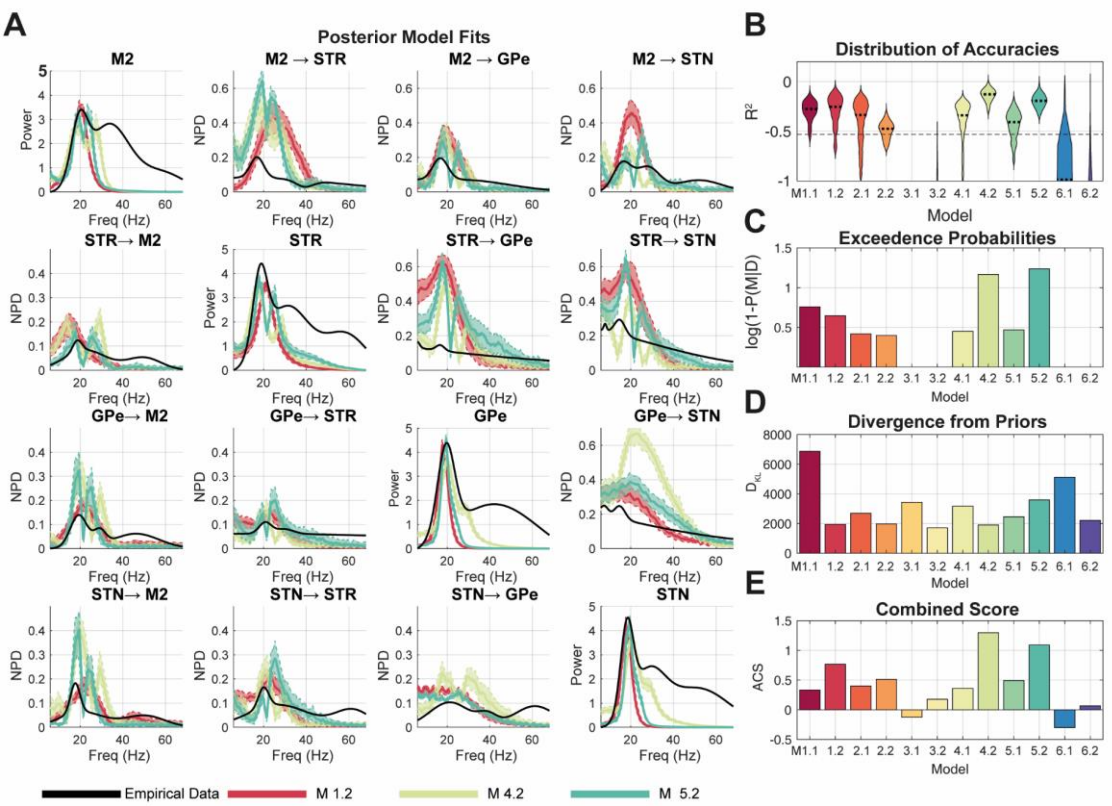

Figure 7 – **Scaling up the ABC model comparison framework – investigating models of the cortico-basal ganglia-thalamic network.** 12 competing models (six families subdivided each into two sub-families) were fitted to empirical data from Parkinsonian rats. Models were fitted to summary statistics of recordings from the motor cortex (M2), striatum (STR), subthalamic nucleus (STN), and external segment of the globus pallidus (GPe). Models were first fitted using ABC to estimate approximate posterior distributions over parameters. To assess relative model performances, 1000 draws were made from each model posterior and corresponding data was simulated. **(A)** The maximum *a posteriori* data features for the top three performing models are shown, with autospectra on the diagonal and NPD off the diagonal (M 5.2 in light green; M 4.2 in turquois; and 1.2 in red). Bounds indicate the interquartile range of the simulated features. **(B)** Violin plots of the distributions of model accuracies (R^2) of the simulated pseudo-data from the empirical data. **(C)** The exceedance probability approximation to the model evidence *P(M|D)* is determined by computing the number of samples from the posterior that exceed the median model accuracy (R^2). **(D)** The Kullback-Leibler divergence of the posterior from prior is shown for each model. Large values indicate high divergence and overfitting. **(E)** Combined scores for accuracy and divergence from priors using ACS.

2.7 Analysis of Simulated Network Dynamics: Beta Bursts

In the final set of results, the simulated output of the fitted model was analysed for its temporal dynamics- specifically beta bursts. Beta bursts were computed using a simple narrowband and threshold method ([12]; see methods section for details). Bursts from the beta band (14-30 Hz) were analysed from data (figure 8) derived from a white noise surrogate (red), an example animal recording (blue), and the simulated output from the best fitting model (M 4.2; green) reported in the previous section. From the broadband time series (figure 8A, top row) it is clear that both the simulated and surrogate data lack the low frequency fluctuations observed in the

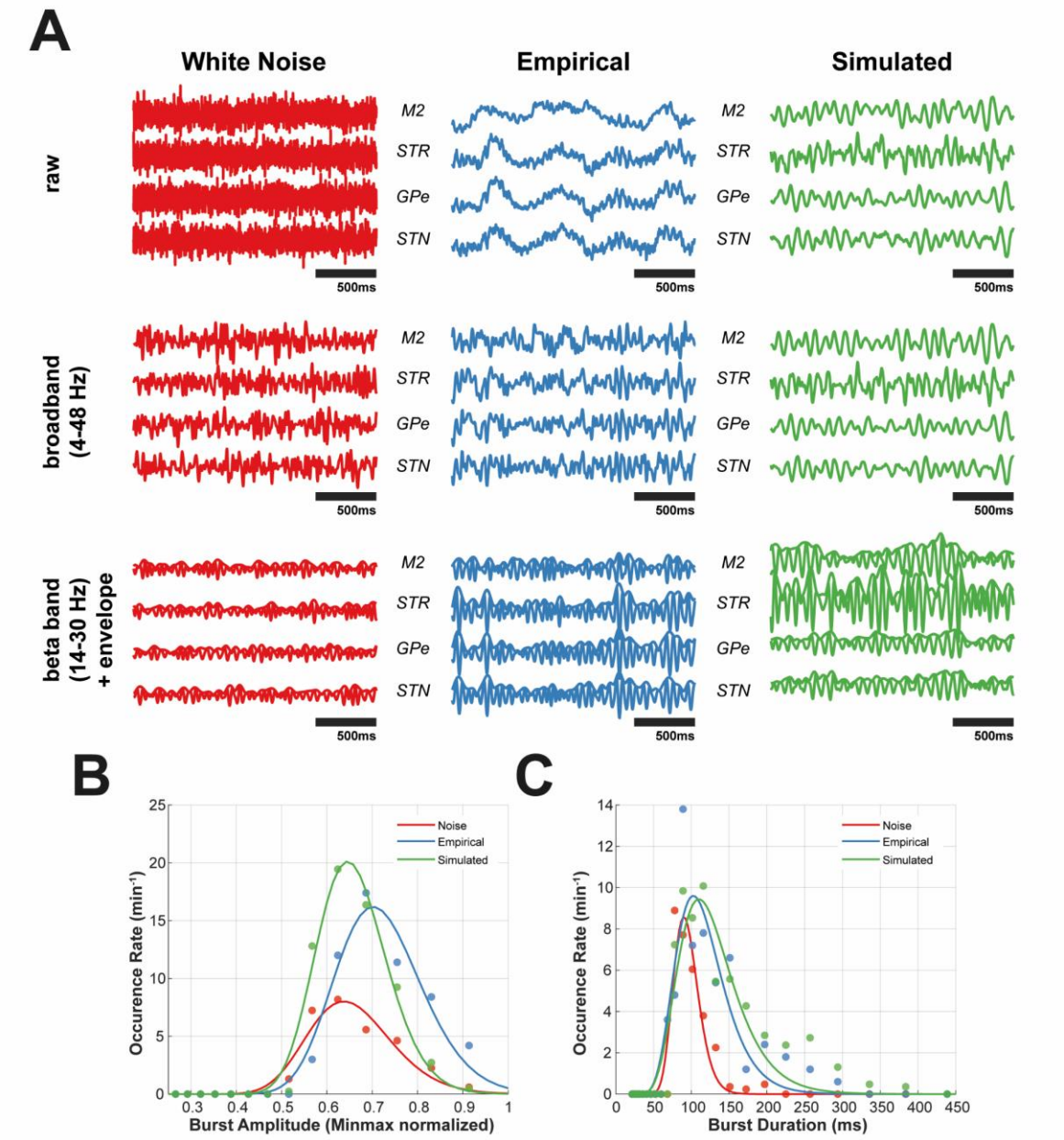

Figure 8 – **Analysis of simulated dynamics from fitted model and comparison with surrogate and empirical data. (A)** Data is compared between a white noise surrogate (red), empirically recorded data (from one example animal; blue), and simulated output from the fitted model (M5.2 in figure 7; green). All raw data was Z-normalized. **(B)** Analysis of burst amplitude distributions. Burst rates were computed within 25 logarithmically spaced bins (filled points). Distributions were approximated by a scaled log-normal function (bold lines). Analyses show that burst rates are roughly matched between empirical (peak at 16 bursts min^{-1}) and simulated data (peak at 20 bursts min^{-1}). **(C)** Same as in (B) but for burst durations. Again, both empirical and simulated bursts exhibit longer tails indicating burst elongation compared to surrogate.

empirical data. When broadband filtered (4-48 Hz; figure 8B, middle row) the simulated data series shows alternating periods of quiescence and bursting that is shared with the empirical data but absent in the surrogate series. When filtered at 14-30 Hz, it is apparent that some of these fluctuations occur in the beta range. These bursts are again absent in the filtered white noise surrogate. These comparisons

were formalized by segregating the data into bursts and comparing distributions of burst amplitude and duration (figure 8B and C).

In figure 8B we show that the burst amplitudes occur with similar spread between the three sets of data. However, the peak rate of occurrence is over twice as large for the empirical (16 bursts min^{-1}) and simulated (20 bursts min^{-1}) data when compared to the surrogate (8 bursts min^{-1}). Analysis of the distribution of burst durations (figure 8C) shows that the median burst length is around 100ms for all data analysed. However, the distribution of burst durations in the empirical data exhibit an elongated tail up to 350ms (up to 5 cycles of 14 Hz) that is well matched by the simulated data.

3 Discussion

3.1 Validity of Approximate Bayesian Computation for Mechanistic Inference of Brain Network Dynamics

In this paper we have formulated a framework for the inference of latent variables from implicit generative models of neural dynamics based upon a sequential Monte Carlo approach to ABC. Further, we provide a first-pass evaluation of these procedures with application to models and data typically used in large scale neuroscience. This work has shown that given a suitable model, the framework will reliably yield valid and useful results. Firstly, we have examined the convergence properties of our modified sequential ABC algorithm and shown that parameter estimates converge to yield accurate reproductions of summary statistics of empirical data. Importantly, using multistarts we demonstrate that these estimates are consistent across multiple instances of the same inversion, and that model parameters are uniquely determined by the data features from which they have been estimated. This validation is important as reproducibility studies have shown that optimization algorithms (such as the deterministic schemes used in DCM) are prone to local minima [69]. Furthermore, we have tested the capacity for a model comparison framework based upon ABC to properly identify the correct generative models of data, in a way similar to that done for steady-state DCM [35].

We apply these techniques to two biological problems, i.e. the STN/GPe subcircuit and the full model space of the cortico-basal-ganglia-thalamic circuit. Results from these analyses have demonstrated that inference using ABC can yield biologically relevant results useful for the post-hoc in-silico analysis of the brain networks of interest. Finally, we provide an analysis of the fitted model and demonstrate that its temporal dynamics (i.e. beta bursts) have features that are like that observed in empirical recordings.

3.2 Comparison with Existing Techniques

The limited resolution of biological recordings means that researchers typically have incomplete observations of the system that they are trying to understand. To make inference, a biologist has to

connect the dots between points of observation, as for example when a neuroscientist must approximate the unseen paths taken by information propagating between two observed parts of the brain. In the introduction we reviewed three approaches: data driven functional connectivity, forward modelling, and finally, data-constrained inverse modelling. We argue that inverse modelling strategies provide the optimal way to infer neural connectivity, even if the need for model optimization and the computational complexity of the inverse problem may make that a difficult proposition.

The enthusiasm for inverse modelling approaches in systems neuroscience has led to an expansive effort to develop generic frameworks such as DCM that can be applied broadly across different modalities of electrophysiological data spanning invasive recordings (e.g. Moran et al. 2011a); electroencephalography (e.g. Boly et al. 2012), and recently calcium imaging [71]. Current approaches in DCM directly model the steady state, frequency-domain statistics of a neural time series (auto and cross spectral densities) rather than the time series itself [35,72,73]. Fitting of models directly from their approximate frequency response at equilibrium enables a large reduction in the computational complexity of the problem by removing the need for numerically expensive integration of the underlying equations of motion [39]. However, recent work has suggested that approximations made to simplify model dynamics may hinder the accuracy of parameter estimation [74].

In this paper, we have introduced a framework which allows for the formal parameter estimation and model comparison of stochastic, non-equilibrium models of neural activity such that have typically restricted to more exploratory modelling approaches such as that employed in The Virtual Brain software [29]. Recent work developing optimization of large scale neural models has shown that by improving parameter estimation algorithms, it is possible to fit high-dimensional, complex models of neural dynamics without appealing to steady-state approximations [75].

By nature of the expansion of model behaviours that can be explored, the ABC framework comes with an increased computational demand over that of estimation from DCM. In the future, work should be done to characterize the specific demands and scalability of the problems that can be investigated within this framework. To provide a sense of time complexity, the evaluation of the model fitting and comparison presented in section 2.6 of the results took approximately 4 hours per model and approximately 30 hours for the complete analysis. We note however, the parallel nature of sampling steps would make the algorithm amenable to deployment on cluster computing such as that done in [76].

In future work, a comparison between the performance of different inversion schemes in application to questions regarding neural network dynamics would be useful. Existing methods such as variational Laplace used in DCM [77]; deterministic sampling approaches [75]; Generative Adversarial Networks [78]; and variants of ABC such as that presented here all provide potential routes to estimation of network mechanisms of spontaneous dynamics yet each are likely suited to different pairings of data

and model types. A quantitative assessment of their applicability in different scenarios would provide concrete answers to the inevitable question of which scheme is best is for a particular type of problem.

3.3 Moving Beyond the Steady State Approximation

Considerations of the steady-state follow from conventions in functional connectivity which have typically involved trial-averaged statistics that appeal to assumptions of stationarity. This convenience, has however obscured the direct investigation of time-dependent changes in connectivity [79,80]; non-equilibrium dynamics [24]; intermittent and spontaneous brain activity such as bursting in brain oscillations [12,53,67]; and the transient recruitment of brain networks [81,82]. Whilst development of novel statistical techniques to determine these phenomena is well underway, the tools to provide a mechanistic explanation for these features are currently lacking. A large body of theoretical work accounts for how these dynamics may be achieved [83,84] but there is up till now a major difficulty in inferring how to link data with theory.

To date, approaches that aim to infer parameters of models associated with these types of behaviours have typically taken the approach of characterization of the pseudo-equilibria associated with each state. For instance the slow changes in underlying parameters that may give rise to epileptic seizures have been investigated by estimating drifts in model parameters by application of DCM within sliding windows [71,85]. Alternatively, in [86] cortical recordings were subdivided into either high or low power windows in order to investigate changes in coupling associated with Parkinsonian beta bursts. In both cases, discretization of the long-time recordings into shorter windows makes the problem of inversion easier by assuming a separation of time scales between the dynamics to be analysed and some 'slow' variable that shifts those dynamics across the entirety of the recording (the so-called *adiabatic* approximation; 85). This acts to displace the estimable unknown to the slow variable and dispenses with an explanation of the mechanisms that generate the underlying itinerancy.

Steady state assumptions have been used previously to make inverse modelling of neural data more computationally amenable [35,88] by restricting the potential operating behaviour of a candidate system to assume that it rests near or at equilibrium. To relax these constraints, we have adopted ABC for neural model inversion to facilitate the exploration of mechanisms of complex brain dynamics. Beyond the typical priors specifying beliefs over parameter values, it is also possible to supply information regarding the expected behaviour of a system's evolution. In this study, dynamical priors restrict solutions of the inversion to regimes in which there is neither entirely stable (convergent) or unstable (divergent) solutions. This is important for several reasons: a) to ensure that stochastic innovations play a role in shaping the system's trajectories; b) to ensure that solutions are non-periodic i.e. have irregular or intermittent dynamics; c) to ensure that solutions are not exponentially divergent; and d) to reduce the potential space of solutions.

These priors are based upon the observation that resting state neural data is both noisy and non-stationary. By making no assumptions as to the causes of these non-stationarities, it would be possible in the future to elucidate mechanisms that are responsible for the features of interest by using ABC model comparison in combination with fitting to summary statistics of the time dynamics themselves (rather than the steady state features as is done currently). Additionally, using the posterior estimates over parameters, these models can be taken forward into other analysis tools such as XPPAUT [89] that are designed for numerical analysis of dynamical systems in order to make an identification of a model's bifurcation structure, the existence of equilibria, and the role of stochastic fluctuations in shaping the dynamics of the inverted systems.

3.4 Limitations of ABC for Model Optimization and Modifications of the Framework

In recent years ABC has become a well-established tool for parameter estimation in systems biology [59,90–93]. Nonetheless, it is known that ABC will not perform well in all cases and faces issues common to all model optimization problems [94] . Specifically, it has been shown that the simplest form of ABC algorithms based upon rejection-sampling approaches are inefficient in the case where the prior densities lie far from the true posterior [95]. When this is the case, it becomes highly improbable that sampled pseudo-data will lie anywhere close to the actual data. Fortunately, for biological models there is often a good degree of a priori knowledge as to the potential range of values in which a parameter is likely to fall. For instance, in the case of neural transmission delays we know that neural activity takes anywhere from a minimum of 1 ms up to a maximum of ~ 50ms to propagate which allows the precision of prior parameter densities to be well constrained. This is a benefit of using biophysically rooted models over more abstract, phenomenological models where the range of potential parameter values are unknown.

As well as issues regarding uninformative priors, all optimization algorithms are subject to the 'curse of dimensionality' by which the exponential increase in the volume of the search space with each new parameter effectively means that any practical sampling procedure will be extremely sparse. Furthermore, nonlinear models often exhibit non-convex, non-smooth objective functions in which multiple local minima exist, making the identification of the global minima difficult. In addition to highlighting the value of proper prior constraints on model parameters, we make use of two other modifications of the ABC procedure which aim to avoid the issue of ill-informed priors as well the local minima problem. The first relates to the non-parametric estimation of proposal densities using kernel density approximation. This property aids evaluation of non-convex objective functions during intermediate steps of the optimization as a) long tailed and skewed distributions can emerge to aid sampling when the true posterior lies far to the tails of the prior; and b) it facilitates simultaneous sampling of multiple regions over parameter space and thus can reduce the risk of converging towards

a local extrema upon initialization. The second addition to the method employs copula estimation of the joint parameter densities from the non-parametric marginals. This facilitates the binding of parameters into correlated 'modes' that can reduce the effective volume of the parameter space to be searched.

It is also important to note that in the case of ABC for model selection, the usage of a Bayes factor derived from an approximate posterior model probability can be problematic in the case of adoption of an insufficient summary statistic [96]. Necessary and sufficient conditions on summary statistics to ensure consistency in Bayesian model selection have been previously described [97]. Despite this, approaches such as the validation performed here for model identification arguably provide the most flexible methods to determine the usage of a given statistic for data reduction.

Model comparison methods such as the one presented here have been criticized for covering only a finite hypothesis space and so 'winning models' are on the surface only the most plausible of a small selection of candidates [98,99]. It has been stated previously that this hypothetically infinite model space can be effectively reduced by specifying prior model probabilities [100], a decision implicit in the selection of a finite space of candidate models to be tested. This comes from a desire to test only models that may be expected to be reasonable *a priori*, and to ask hypotheses that are experimentally useful. For instance, in the model space presented in figure 6, we include models lacking STN → GPe feedback for purposes of demonstrating that the model comparison framework can identify models with poor anatomical grounding. In a real case of model comparison, we advocate a utilitarian selection of candidate models [101] where including hypotheses where there is no strong *a priori* reasoning is not considered useful.

3.5 Conclusions

Overall, we have presented a novel framework for parameter estimation and model comparison for models of large-scale neural activity. This framework provides a robust method by which non-stationary brain dynamics may be explored in terms of their underlying neural dynamics. We have demonstrated that this framework exhibits convergence upon empirical data; provides consistent results; can identify true generative models; can determine biologically plausible models; and shows initial evidence that this platform can be scaled to larger models and datasets. Whilst this work constitutes a first validation and description of the method, more work will be required to establish its validity in the context of more complex models as well as statistics of non-stationary properties of neural dynamics. Nevertheless, this provides a unique opportunity by which many intermittent or stochastic features of large-scale brain activity may be explored in terms of their underlying neurophysiological parameters.

4 Methods

4.1 Generative Modelling of Neural Pseudo-Data

The ABC inversion scheme is based upon sampling from a proposal distribution of parameters to generate realizations of the generative process that we refer to as pseudo-data (figure 1A; green box). A model M is specified by the state equations of the dynamical system F and observation model G :

$$M(x, \theta) = G(F(x, \theta^F), \theta^G).$$

EQUATION 1

The state equations of the model F describe the evolution of states x with parameters θ^F. This model describes the underlying dynamics that give rise to the evolution of the neural states. Model G describes the effects upon the signals that are introduced by the process(es) that occur during experimental acquisition and recording of the neural signal(s) x and is parameterized by θ^G. This observation model accounts for confounds introduced during the recording and acquisition of the data such as changes in signal-to-noise ratio, crosstalk, or filtering of the data.

In general, model M could describe any time evolving dynamic of neural data such as spiking networks, conductance models, or phenomenological models (e.g. phase oscillators, Markov chains). In this paper we use a model based upon the anatomy of the cortico-basal-ganglia-thalamic network that is based on coupled neural mass equations [61,102]. The biological and theoretical basis of this model has been described previously [44–46]. The state equations of this model are given in Supplementary Information I. Furthermore, we adapt the original equations to explicitly incorporate stochastic inputs and finite transmission delays. This system of stochastic delay differential equations is solved using the Euler-Maruyama method. For details of the integration of the state equations see Supplementary Information, part 2. For sections of the results 2.1 through to 2.4 we use a subnetwork of the full model comprising the reciprocally coupled STN/GPe subcircuit. This model can be again divided into separate models (for the purposes of performing face validation and example model comparison) by constraining priors on the connectivity between the STN and GPe. Section 2.5 uses a wider model space comprising variations upon the full model (see figure 6).

The generative and observation models both require parameters and initial states to be evaluated. We take the initial states as zero and let the system evolve for an initialization of three seconds to allow for an initial transient that is then removed. For optimization parameters are either be fixed or variable. In the case of variable parameters (parameters to be fit), the experimenter sets a prior density encoding their *a priori* beliefs about the values that the parameters take. This is encoded through a mean and variance for each parameter, with the variance encoding the inverse-precision of a prior belief. In this way fixed parameters can be thought as being known with complete confidence. We take the majority of prior values for parameters of the cortico-basal-ganglia-thalamic network from [44] but set some delays and connection strengths given updated knowledge in the literature. A table of parameters can be found in Supplementary Information III.

4.2 Model Inversion with ABC

4.2.1 Algorithm Overview

In order to estimate the parameters of the model M given the information provided by the empirical recordings we use a stochastic optimization scheme based upon Approximate Bayesian Computation (ABC; [58,59,65]; and see [94] for a tutorial). Specifically, we use an approach based upon a sequential Monte Carlo variant of ABC [59,103]. Most generally, the algorithm forms a sample of draws (particles) taken from a prior distribution and then goes on to estimate a sequence of proposal distributions via iterative rejection of the particles. Given a shrinking tolerance schedule, the simulated pseudo-data (generated from the sample-parameterized forward model) and the empirical data should converge as the proposal distribution approaches the true posterior.

The ABC algorithm is illustrated in figure 1C (orange box) and follows the procedure below. Probability densities are given by $P(\cdot)$ parameters are indicated by θ; models by M; data by D; and distances by ρ. Samples are indicated by hat notation (i.e. $\hat{\cdot}$); subscripts indicate the sample number; proposal distributions are indicated by an asterisk (i.e. $P(\cdot)^*$); and subscripts equal to zero denote belonging to the empirical data (i.e. μ_1 and $\mu_0$are summary statistics of sample 1 and of the empirical data respectively).

1. *Specify prior distribution of parameters, $P(\theta)$ of the model M.* *(Model Prior)*
2. *Randomly sample N times from the prior to yield samples $\hat{\theta}_n$.* *(Sampler)*
3. *Simulate pseudo-data $\hat{D}_n \sim M(x_n, \hat{\theta}_n)$.* *(Simulation of joint distribution)*
4. *Compute summary statistic of μ_n of pseudo-data $\hat{D}_n$.* *(Data Transform)*
5. *Compute distance ρ of μ_n from μ_0.* *(Assess Model Fit)*
6. *Reject $\hat{\theta}_n$ if distance $\rho(\mu_n, \mu_0) \geq \epsilon_q$.* *(Rejection Sampling)*
7. *Form proposal distribution, $P^*(\theta|D_0)$, from the accepted parameter samples.* *(Form Proposal)*
8. *Iterate for $q = 1, \ldots, Q$, setting $P(\theta) = P^*(\theta|D_0)$, and shrinking the distance threshold.* *(Adaptive Tolerance Schedule)*
9. *At convergence criteria, accept proposal distribution as posterior: $P(\theta|D_0) \cong P^*(\theta|D_0)$,* *(Estimation of Approximate Posterior)*

To avoid sample wastage across iterations, we store the particles from step (2) and their resulting distance from the data (step 5) across the Q iterations such that they are propagated through a sequence of intermediate distributions. In this way, at step (7) the updated proposal distribution comprises samples from both current and past draws selected on the basis of a threshold calculated over the current draw.

The distance of the data from pseudo-data uses the coefficient of determination (R^2) as the objective function $\rho(\mu_n, \mu_0)$ closely related to the mean squared error (MSE). We note that in nonlinear models (and linear models without an intercept), R^2 takes value in [-∞ 1] with 1 indicating a perfect fit, and negative R^2 denoting a fit worse than that of the mean of the data [104]. Setting a shrinking distance threshold ϵ ensures that the posterior estimates converge upon solutions that most accurately reproduce the summary statistics of the observed data [103]. With non-negligible ϵ_Q, the algorithm samples from an approximate posterior distribution $P\left(\theta|\rho\left(\hat{D}, D\right) < \epsilon_Q\right)$ rather than the true posterior $P(\theta|D)$. Thus the upper bound on the error of parameter estimates is therefore determined by how far ϵ_Q is from zero [105].

4.2.2 Adaptive Tolerance Schedule

We set an adaptive tolerance schedule that facilitates sampling from a sequence of increasingly constrained target distributions. The tolerance schedule is specified by a gradient on the expected distance of the next set of samples:

$$\epsilon_{q+1} = \epsilon_q + \Delta\epsilon_q$$

EQUATION 2

where the expected change in the distance of the new samples $\Delta\epsilon_q$ is given by:

$$\Delta\epsilon_q = \begin{cases} \epsilon_{\mathrm{q}} - \epsilon_{q-1}, & N_{accept} \geq \gamma \\ \epsilon_q^* - \epsilon_{q-1}, & N_{accept} < \gamma \end{cases}$$

EQUATION 3

where N_{accept} is the number of accepted samples and γ is a minimum criterion on the accepted sample size to carry forward the tolerance shrinkage at its current gradient. If $N_{accept} < \gamma$ then this gradient is assumed to be too steep and the expected

gradient is recalculated using a modified tolerance ϵ_q^* that is computed using the median distance ρ between the sample pseudo-data from that real (i.e. $\epsilon_q^* = \tilde{\rho}$, where ~ indicates the median). Thus γ parameterizes the coarseness of the optimization. If γ is very large (e.g. >99% of *N*) then the algorithm will converge slowly but accurately, whereas if γ is very small (e.g. 1% of *N*) the algorithm will be inaccurate and biased. We set γ to be the two times the estimated rank of the parameter covariance matrix i.e. $rank(\Sigma)$ (for details of estimation see bottom of section 4.2.3).

4.2.3 Formation of Proposal Distributions

Following rejection sampling, the proposal density $P^*(\theta|D_0)$ must be sampled from the accepted parameters sets. In the initial draw of samples from the prior we treat the data as Gaussian and form a multivariate normal:

$$P^*(\theta|D_0) = N(\mu, \Sigma)$$

EQUATION 4

where μ is a vector of the univariate expectations and Σ their covariances. In subsequent iterations whereby a minimum sample is accumulated, we use nonparametric estimation of the marginal densities over each parameter using a non-parametric kernel density estimator [106]. This approach does not assume normality of parameter distributions and may estimate multimodal parameter distributions. This flexibility allows for sampling across multiple possible maxima at once, particularly at intermediate stages of the optimization. The bandwidths (determining the smoothness) of the kernel density estimators are optimized using a cross-validation approach [107].

We then form the multivariate proposal distribution using the t-copula [108]. Copula theory provides a mathematically convenient way of creating the joint probability distribution whilst preserving the original marginal distributions. Data are transformed to the copula scale (unit-square) using the kernel density estimator of the cumulative distribution function of each parameter and the transformed to the joint space with the t-copula.

The copula estimation of the correlation structure of the parameter distributions acts to effectively reduce the dimensionality of the problem by binding correlated parameters into modes. The effective rank of the posterior parameter space (used in the computation of the adaptive tolerance schedule and reported in the results as a post-hoc assessment of parameter learning) can be estimated by taking the eigenvalues of the covariance matrix and normalizing the coefficients by their sum. Using a cumulative sum of the ordered coefficients we can then determine the number of modes that can explain 95% of the variance of the parameter samples.

4.2.4 Priors on Model Dynamics: Finding Metastable States

As the modelling is explicitly concerned with regimes in which dynamics are non-stationary, we explicitly incorporate this into the prior specification of model dynamics. Concretely, we reject simulations in which there is either strong periodicity in the envelope of the signal, or if the envelope is divergent, respective signs of a signal that is either close to equilibrium or unstable. We define the envelope *S* as the magnitude of the analytic signal computed from the Hilbert transformed broadband signal $S(t) = |H(x(t)|$. We compute envelope periodicity using the autocorrelation (scaled to unity at zero-lag) of this signal and remove signals with statistically significant correlations at lags ±2ms. To determine divergence, we test for monotonicity of the envelope by applying a non-parametric rank correlation test. In the case of a divergent signal, ranks will be significantly increasing or decreasing and so we remove signals with statistically significant correlation. Models failing these criteria are automatically removed by setting their R^2 to -∞.

4.3 Model Comparison

In the process of model-based inference, hypotheses must be compared in their ability to explain the observed data. In this section we describe a method [62,109] in which models fit with ABC can be formally compared. At convergence ABC approximates the posterior parameter distribution $P(\theta|D_0)$ that may be used to approximate the model evidence $P(M|D_0)$.

This estimate is made by drawing N times from the posterior and then computing an exceedance probability to estimate the marginal probability of the jth model M given data D_0 (the approximate model evidence):

$$P\left(M^j \middle| D_0\right) \cong \frac{\#\, \rho(\mu_n, \mu_0) \geq \epsilon^*}{N}$$

EQUATION 5

where ϵ^* is a threshold on the distance metric ρ that is suitably small to give an acceptable fit on the data. If $\epsilon^*$is held constant and the data is identical between models, then the exceedance probabilities may be compared to yield the model that gives the most accurate fit. In practice we set ϵ^* to be the median distance of all sets of models. In order to compare models a Bayes factor K can be constructed:

$$K = \frac{P(M^1|D_0)}{P(M^2|D_0)}$$

EQUATION 6

where K gives the strength of evidence of model 1 over model 2 that can be interpreted as proposed by [110]. In addition to calculating a Bayes factor derived from the estimated model evidences, we also provide a post-hoc examination of model complexity in the form of a divergence of the posterior from the prior [36,77]. Specifically, we estimate the Kullback-Lieber divergence D_{KL} of the posterior density $P(\theta|D_0)$ from the prior density $P(\theta)$ over F discretized bins of the density:

$$D_{KL}(P(\theta|D_0)||P(\theta)) = -\sum_{i=1}^{F} P(\theta)\, log\left(\frac{P(\theta)(i)}{P(\theta|D_0)(i)}\right)$$

EQUATION 7

This is a simplification of the full multivariate divergence and ignores the dependencies between variables encoded in the posterior covariance. We use the full multivariate divergence (given in supplementary information IV) that makes a Gaussian approximation to the marginals, whilst accounting for their covariance structure. This is a reasonable approximation to make as we only expect densities to be non-Gaussian in the proposal densities. Given that the algorithm has reached convergence, proposal densities over parameters will be normally distributed. The D_{KL} can then be used as a post-hoc regularization of the model comparison in which model accuracy provides the primary objective function, whilst D_{KL} provides a secondary discriminator to determine the degree of overfitting. When two models have identical accuracy, they can be further segregated by examining D_{KL} and favouring the model that deviates the least from the prior parameters as being the most parsimonious of the two explanations of the data since it requires less deviation from the experimenter's prior expectations of parameter values (as an appeal to Occam's razor; [111]). Finally, we combine both the model accuracy and D_{KL} to yield a combined metric (accuracy-complexity score; ACS) accounting for both accuracy and complexity for model j:

$$ACS = -\log\left(1 - P\left(M^j \middle| D_0\right)\right) - \log\left(\frac{D_{KL}^j}{\widetilde{D}_{KL}}\right)$$

EQUATION 8

where D_{KL}^j is the divergence of posteriors from priors for the jth model and $\widetilde{D}_{KL}$ is the median divergence across the whole model space. With increasing accurate and/or minimally divergent models the ACS metric will tend towards large positive values.

4.4 Empirical Data: Recordings from Parkinsonian Rats

Summary statistics are computed from empirical data and then used to fit the generative forward model. In the example implementation used in this paper we use multisite basal ganglia and single site cerebral cortex recordings in rats (n = 9) that

have undergone a 6-hydroxydopamine (6-OHDA) induced dopamine depletion of the midbrain, a lesion model of the degeneration associated with Parkinsonism in humans [112,113]. The animals were implanted with two electrodes to measure local field potentials (LFP) from multiple structures in the basal ganglia: dorsal striatum (STR), external segment of the globus pallidus (GPe), and the subthalamic nucleus (STN). Additionally electrocorticography was measured over area M2 of the motor cortex, a homologue of the Supplementary Motor Area (SMA) in humans [114]. Animals were recorded under isoflurane anaesthesia and during periods of "cortical-activation" induced by a hind-paw pinch [115]. The details of the experimental procedures were previously published [112,113].

Pre-processing of time series data (LFP and ECoG) was done as follows: data were 1) truncated to remove 1 second (avoid filter artefacts); 2) mean corrected; 3) band-passed filtered 4-100 Hz with a finite impulse response, two-pass (zero-lag) with optimal filter order; 4) data were split into 1 second epochs with each epoch subjected to a Z-score threshold criterion such that epochs with high amplitude artefacts were removed.

4.5 Computation of Summary Statistics

We derive a set of summary statistics from signal analyses of the experimental and simulated time series. These statistics transform both the data and pseudo-data into the same feature space such that they can be directly compared (figure 1B; blue box). Here we give details of the specific summary statistics used in this implementation but also note the flexibility of the framework in hypothetically fitting any time-series statistic. It is important to note that the summary statistic is vital in determining the outcome of the inverse modelling. The set of statistics must effectively encode all phenomena of the original data that the experimenter wishes to be modelled.

4.5.1 Frequency Spectra

We use the autospectra to constrain the oscillatory characteristics of each neural mass population. Auto-spectral analyses were made using the averaged periodogram method across 1 second epochs and using a Hanning taper to reduce the effects of spectral leakage. Frequencies between 49-51 Hz were removed so that there was no contribution from 50 Hz line noise. 1/f background noise was removed by first performing a linear regression on the log-log spectra (at 4–48 Hz) and then subtracting the linear component from the spectra. This ensured that the inversion scheme was focused upon fitting the spectral peaks in the data and not the profile of 1/f background noise. To simplify observation modelling of differences in experimental recording gains between sites, all spectra were normalized by dividing through by their summed power at 4-48 Hz.

4.5.2 Directed Functional Connectivity

To quantify interactions between populations, we use non-parametric directionality (NPD; Halliday 2015), a directed functional connectivity metric which describes frequency resolved, time lagged correlations between time series. The NPD was chosen as it directly removes the zero-lag component of coherence and so statistics are not corrupted by signal mixing. This has the effect of simplifying observation modelling by reducing the confounding effects signal mixing.

Estimates of NPD were obtained using the Neurospec toolbox (http://www.neurospec.org/). This analysis combines Minimum Mean Square Error (MMSE) pre-whitening with forward and reverse Fourier transforms to decompose coherence estimates at each frequency into three components: forward, reverse and zero lag. These components are defined according to the corresponding time lags in the cross-correlation function derived from the MMSE pre-whitened cross-spectrum. This approach allows the decomposition of the signal into distinct forward and reverse components of coherence separate from the zero-lag (or instantaneous) component of coherence which can reflect volume conduction. The method uses temporal precedence to determine directionality. For a detailed formulation of the method see [116]; and for its validation see [117]. We ignored the instantaneous component of the NPD and use only the forward and reverse components for all further analyses.

4.5.3 Data Pooling and Smoothing

In all optimizations using empirical data in this study we used the group-averaged statistics computed from recordings from a group of unilaterally 6-OHDA lesioned animals. As a final processing step, both the autospectra and NPD were smoothed to

remove noise as well as to reduce them to their main components (peaks). This was achieved via fitting of a sum of one, two, or three Gaussians to the spectra using the adjusted R^2 as a guide for selecting the most appropriate sum of Gaussians. This allows the spectra to be reduced to a smooth sum of normal modes corresponding to the main peaks in the spectra (i.e. alpha (8-12), low beta (12-21) and high beta/gamma frequencies (21-30 Hz)). Empirical and simulated data were transformed similarly to produce equivalent autospectra and NPD. The simulated data did not undergo regression of 1/f background in the autospectra; nor Gaussian smoothing of the data features as the simulations aimed to approximate the smoothed data features as directly as possible. In both cases these additional transforms may be misappropriated by the optimization procedures to yield superficially close fits.

4.6 Validation of ABC for Optimization and Model Comparison

4.6.1 Testing for Convergence of ABC Parameter Estimation

To examine the convergence properties of the model inversion, we provide an example inversion using the STN/GPe s tracking the model fit to the data, parameter estimates and their dependencies across the iterations of ABC. If the optimization converges then it is expected that the summary statistics (directed functional connectivity and autospectra) of the data and the model closely match; the posterior density over parameters will increase in precision (as they are informed by the data); and the posterior covariances of the parameters (i.e. their dependencies) will be learnt.

4.6.2 Testing for ABC Estimator Consistency with Multi-start, and Uniqueness of Empirical Inference with Multi-Start

To test whether the ABC estimator will: a) yield parameter estimates that are unique to the data from which they have been optimized; and b) yield consistent estimation of parameters across multiple instances of the schema, we performed a multi-start. We performed two procedures of ten multi-starts for a single STN/GPe model (using identical priors). The two multi-starts were identical except for the data to which they were fitted. Over the evolution of each optimization, the posterior parameters are reported as the maximum a posteriori (MAP) estimates taken as the median of the marginal distributions.

When testing the point (a)- that parameter estimates are unique to the data from which they are fitted - we performed a 10-fold cross-validation procedure in which we used a one-sample Hotelling procedure to test for significant difference of each fold's mean from that of the left-out sample. We report the probability of the folds that yielded a significant test, with high probability indicating that the left-out MAP estimates are likely to deviate from the rest of the fold. In this way we can identify the probability of an ABC initialization yielding a non-consistent sample. Secondly, we test (b) – that MAP estimates are unique to the data on which they have been fitted- using the Szekely and Rizzo energy test [118] between the samples from data A and B, with the null-hypothesis that the multistart samples derived from different data arise from the same distribution. Finally, we use a Multivariate Analysis of Variance (MANOVA) procedure to test for difference in means between the two multivariate samples. In the case of a positive MANOVA test, post-hoc analyses of differences in means of each parameter are made using two sample ANOVAs.

4.6.3 Testing the Face Validity of Model Comparison Framework – Model Identification

To test the face validity of the model comparison framework, we constructed a confusion matrix, an approach commonly used in machine learning to examine classification accuracy. Three different models of the STN/GPe network were fit to the empirical data and then using the fitted parameters three synthetic data sets were simulated. We chose a model with reciprocal connectivity: (1) STN ↔ GPe; and then two models in which one connection predominated: (2) STN → GPe and (3) GPe → STN. The three models (with the original priors) were then fitted back onto the synthetic data. A model comparison procedure was then performed to see whether it could correctly identify the best model as the one that had generated the data. The model comparison outcomes (accuracy of the fit; D_{KL} of posteriors from priors; and the combined ACS measure) were then plotted in a 3 x 3 matrix of data versus models. In the case of valid model selection, the best fitting models should lay on the diagonal of the confusion matrix.

4.6.4 Application of Model Fitting and Comparison to STN-GPe Network

To test the biological utility of model comparison, we used the simple STN/GPe model and investigated the role of reciprocal coupling. We fitted the three models from the previous section to the data from the STN and GPe and performed a model comparison to identify the best explanations for the data. From a wide body of anatomical and computational evidence it is expected that feedback between the STN and GPe is a probable generation mechanism of subcortical beta rhythms. Thus, we expect that models with asymmetric connectivity will perform poorly.

4.6.5 Testing the Scalability of the Framework with Application to the Full Model Space

In order to demonstrate the scalability of the optimization and model comparison framework, we used the space of 12 models described below. We individually fitted the models and then performed model comparison to select the best fitting model. A set of null models were included which are anatomically implausible. If model selection performed correctly, then it is expected that these models would perform poorly.

To investigate the importance of known anatomical pathways in reconstructing the observed steady state statistics of the empirical local field potentials (i.e. autospectra and NPD), we considered a set of competing models. Specifically, we look at the role of four pathways and their interactions: the cortico-striatal indirect; the cortico-subthalamic hyperdirect; thalamocortical relay; and the subthalamic-pallidal feedback. In total we tested 6 families of models (presented later in the results section- figure 6):

1. + indirect.
2. + indirect/ + hyperdirect pathway.
3. + hyperdirect / - indirect.
4. + indirect / - hyperdirect/ + thalamocortical.
5. + indirect / + hyperdirect/ + thalamocortical.
6. - indirect / + hyperdirect/ + thalamocortical.

We considered these six families and divide them into two sub-families that do or do not include the subthalamopallidal (STN → GPe) feedback connection. Family (1) investigates whether the indirect pathway alone can explain the pattern of observed spectra and functional connectivity. In the case of family (2), previous work has highlighted the importance of hyper-direct connections in the functional connectivity [119,120], yet anatomical research has shown dopamine to weaken the density of synaptic projections [121]. Thus, family (2) provides an ideal set of models to examine the nonlinear mapping of anatomical to functional connectivity described in the introduction of this paper. Families (3) and (6) represent *null* models in which the indirect pathway is excluded and are used as implausible models to test the model comparison procedure. This is because it is thought that the indirect pathway is vital to explain activity within the network following dopamine depletion associated with PD [55–57]. The functional role of the thalamocortical subnetwork is relatively unknown (but see recent work: Reis et al. 2019) and so families (4) and (5) provide an examination of whether the addition of the thalamic relay can better explain the empirical data. The second level of families (i.e. x.1-2) investigates whether the reciprocal network formed by the STN and GPe is required to explain observed patterns of connectivity in the data. This network has been the subject of much study and is strongly hypothesized to play an important role in the generation and/or amplification of pathological beta rhythms [68,122,123].

4.6.6 Analysis of Beta Bursts and Comparison Between Surrogate, Empirical, and Simulated Data

In order to analyse beta bursts [12], we apply a narrowband filter in the beta range and then define burst events with respect to a threshold on the amplitude envelope. Specifically, data were filtered using a two-pass, Butterworth filter between 14-30 Hz and filter order optimized. The magnitude of the Hilbert transformed narrowband series is used to compute the amplitude envelope of the signal. We then set a threshold at the 75th percentile of the envelope. Suprathreshold events that are shorter than 1.5 periods of the highest frequency bound are excluded. Burst amplitudes are reported as the mean of the envelope during

the burst. Durations are reported as the total amount of time that the burst remains suprathreshold. When plotting histograms, we normalize counts by the total duration of the data in order to give burst frequency in terms of occurrence rates (burst min^{-1}).

We apply this analysis to data from a white noise surrogate, an example 6-OHDA experimental recording, and simulated data from ABC fitted models. Surrogate and simulated data lengths were matched to that of the experimental series (100s).

5 Acknowledgments and Funding

T.O.W. acknowledges funding from UCL CoMPLEX doctoral training program, and the UCL Bogue Fellowship. H.C. receives funding from an MRC Career Development award (MR/R020418/1). S.F.F. receives funding support from the UCL/UCLH NIHR Biomedical Research Centre. L.B. acknowledges funding support from the Leverhulme Trust (RPG- 2017-370). We thank Peter Magill and Andrew Sharott at the Brain Network Dynamics Unit, Oxford University for making available the experimental data used in this study. We thank all authors of the publicly available toolboxes used in this paper.

6 References

1. Mantini D, Perrucci MG, Del Gratta C, Romani GL, Corbetta M. Electrophysiological signatures of resting state networks in the human brain. Proc Natl Acad Sci. 2007;104: 13170–13175. doi:10.1073/pnas.0700668104

2. Deco G, Jirsa VK, McIntosh AR. Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat Rev Neurosci. 2011;12: 43–56. doi:10.1038/nrn2961

3. Brookes MJ, Woolrich M, Luckhoo H, Price D, Hale JR, Stephenson MC, et al. Investigating the electrophysiological basis of resting state networks using magnetoencephalography. Proc Natl Acad Sci. 2011;108: 16783–16788. doi:10.1073/pnas.1112685108

4. Zalesky A, Fornito A, Cocchi L, Gollo LL, Breakspear M. Time-resolved resting-state brain networks. Proc Natl Acad Sci U S A. 2014;111: 10341–6. doi:10.1073/pnas.1400181111

5. Arieli A, Sterkin A, Grinvald A, Aertsen A. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science. 1996;273: 1868–71. doi:10.1126/science.273.5283.1868

6. Fox MD, Snyder AZ, Zacks JM, Raichle ME. Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. Nat Neurosci. 2006;9: 23–25. doi:10.1038/nn1616

7. Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, et al. Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat Neurosci.

2010;13: 369–78. doi:10.1038/nn.2501

8. Iemi L, Chaumon M, Crouzet SM, Busch NA. Spontaneous Neural Oscillations Bias Perception by Modulating Baseline Excitability. J Neurosci. 2017;37: 807–819. doi:10.1523/JNEUROSCI.1432-16.2016

9. Shin H, Law R, Tsutsui S, Moore CI, Jones SR. The rate of transient beta frequency events predicts behavior across tasks and species. Elife. 2017;6. doi:10.7554/eLife.29086

10. Busch NA, Dubois J, VanRullen R. The Phase of Ongoing EEG Oscillations Predicts Visual Perception. J Neurosci. 2009;29: 7869–7876. doi:10.1523/JNEUROSCI.0113-09.2009

11. Torrecillos F, Tinkhauser G, Fischer P, Green AL, Aziz TZ, Foltynie T, et al. Modulation of Beta Bursts in the Subthalamic Nucleus Predicts Motor Performance. J Neurosci. 2018;38: 8905–8917. doi:10.1523/JNEUROSCI.1314-18.2018

12. Feingold J, Gibson DJ, DePasquale B, Graybiel AM. Bursts of beta oscillation differentiate postperformance activity in the striatum and motor cortex of monkeys performing movement tasks. Proc Natl Acad Sci. 2015;112: 13687–13692. doi:10.1073/pnas.1517629112

13. Lundqvist M, Rose J, Herman P, Brincat SL, Buschman TJ, Miller EK. Gamma and Beta Bursts Underlie Working Memory. Neuron. 2016;90: 152–164. doi:10.1016/J.NEURON.2016.02.028

14. Litvak V, Jha A, Eusebio A, Oostenveld R, Foltynie T, Limousin P, et al. Resting oscillatory cortico-subthalamic connectivity in patients with Parkinson's disease. Brain. 2011;134: 359–374. doi:10.1093/brain/awq332

15. Baggio H-C, Segura B, Sala-Llonch R, Marti M-J, Valldeoriola F, Compta Y, et al. Cognitive impairment and resting-state network connectivity in Parkinson's disease. Hum Brain Mapp. 2015;36: 199–212. doi:10.1002/hbm.22622

16. Göttlich M, Münte TF, Heldmann M, Kasten M, Hagenah J, Krämer UM. Altered Resting State Brain Networks in Parkinsonâ€™s Disease. Bodurka J, editor. PLoS One. 2013;8: e77336. doi:10.1371/journal.pone.0077336

17. Wang X-J, Introduction I, Synchronization a, Resonance B, Subthreshold C, Rhythms VS, et al. Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. Physiol Rev. 2010;90: 1195–1268. doi:10.1152/physrev.00035.2008.

18. Greicius MD, Srivastava G, Reiss AL, Menon V. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc Natl Acad Sci U S A. 2004;101: 4637–42. doi:10.1073/pnas.0308627101

19. Cagnan H, Mallet N, Moll CKE, Gulberti A, Westphal M, Gerloff C, et al. Temporal evolution of beta bursts in the parkinsonian cortico-basal ganglia network. bioRxiv. 2018; 458414. doi:10.1101/458414

20. Deco G, Jirsa VK, McIntosh AR. Resting brains never rest: computational insights into potential cognitive architectures. Trends Neurosci. 2013;36: 268–74. doi:10.1016/j.tins.2013.03.001

21. Fornito A, Zalesky A, Breakspear M. The connectomics of brain disorders. Nat Rev Neurosci. 2015;16: 159–72. doi:10.1038/nrn3901

22. Avena-Koenigsberger A, Misic B, Sporns O. Communication dynamics in complex brain networks. Nat Rev Neurosci. 2017;19: 17–33. doi:10.1038/nrn.2017.149

23. Park H-J, Friston K. Structural and functional brain networks: from connections to cognition. Science. 2013;342: 1238411. doi:10.1126/science.1238411

24. Deco G, Kringelbach ML, Jirsa VK, Ritter P. The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core. Sci Rep. 2017;7: 3095. doi:10.1038/s41598-017-03073-5

25. Friston KJ. Functional and Effective Connectivity: A Review. Brain Connect. 2011;1: 13–36. doi:10.1089/brain.2011.0008

26. Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran JP, Meuli R, et al. Predicting human resting-state functional connectivity from structural connectivity. Proc Natl Acad Sci U S A. 2009;106: 2035–40. doi:10.1073/pnas.0811168106

27. Cabral J, Kringelbach ML, Deco G. Exploring the network dynamics underlying brain activity during rest. Prog Neurobiol. 2014;114: 102–31. doi:10.1016/j.pneurobio.2013.12.005

28. Ton R, Deco G, Daffertshofer A. Structure-Function Discrepancy: Inhomogeneity and Delays in Synchronized Neural Networks. Daunizeau J, editor. PLoS Comput Biol. 2014;10: e1003736. doi:10.1371/journal.pcbi.1003736

29. Ritter P, Schirner M, McIntosh AR, Jirsa VK. The virtual brain integrates computational modeling and multimodal neuroimaging. Brain Connect. 2013;3: 121–45. doi:10.1089/brain.2012.0120

30. Deco G, Jirsa VK. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. J Neurosci. 2012;32: 3366–75. doi:10.1523/JNEUROSCI.2523-11.2012

31. Freyer F, Roberts JA, Becker R, Robinson PA, Ritter P, Breakspear M. Biophysical mechanisms of multistability in resting-state cortical rhythms. J Neurosci. 2011;31: 6353–61.

doi:10.1523/JNEUROSCI.6693-10.2011

32. Freeman WJ, Holmes MD. Metastability, instability, and state transition in neocortex. Neural Networks. 2005;18: 497–504. doi:10.1016/j.neunet.2005.06.014

33. Friston KJ. Transients, Metastability, and Neuronal Dynamics. Neuroimage. 1997;5: 164–171. doi:10.1006/nimg.1997.0259

34. Deco G, Jirsa V, McIntosh AR, Sporns O, Kötter R. Key role of coupling, delay, and noise in resting brain fluctuations. Proc Natl Acad Sci U S A. 2009;106: 10302–7. doi:10.1073/pnas.0901831106

35. Moran RJ, Stephan KE, Seidenbecher T, Pape H-C, Dolan RJ, Friston KJ. Dynamic causal models of steady-state responses. Neuroimage. 2009;44: 796–811. doi:10.1016/j.neuroimage.2008.09.048

36. Penny WD. Comparing dynamic causal models using AIC, BIC and free energy. Neuroimage. 2012;59: 319–330. doi:10.1016/j.neuroimage.2011.07.039

37. Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, et al. Comparing families of dynamic causal models. PLoS Comput Biol. 2010;6: e1000709. doi:10.1371/journal.pcbi.1000709

38. Friston KJ, Bastos A, Litvak V, Stephan KE, Fries P, Moran RJ. DCM for complex-valued data: cross-spectra, coherence and phase-delays. Neuroimage. 2012;59: 439–55. doi:10.1016/j.neuroimage.2011.07.048

39. Kloeden PE, Platen E. Numerical Solution of Stochastic Differential Equations [Internet]. Springer Berlin Heidelberg; 1992. Available: https://books.google.co.uk/books?hl=en&lr=&id=r9r6CAAAQBAJ&oi=fnd&pg=PR17&ots=Zr_eiYMmgX&sig=XkgHJh2SitUqiw3cSATzd9_8TW4&redir_esc=y#v=onepage&q&f=false

40. West TO, Berthouze L, Halliday DM, Litvak V, Sharott A, Magill PJ, et al. Propagation of Beta/Gamma Rhythms in the Cortico-Basal Ganglia Circuits of the Parkinsonian Rat. J Neurophysiol. 2018; jn.00629.2017. doi:10.1152/jn.00629.2017

41. Ray NJ, Jenkinson N, Wang S, Holland P, Brittain JS, Joint C, et al. Local field potential beta activity in the subthalamic nucleus of patients with Parkinson's disease is associated with improvements in bradykinesia after dopamine and deep brain stimulation. Exp Neurol. 2008;213: 108–113. doi:10.1016/j.expneurol.2008.05.008

42. Brown P, Oliviero A, Mazzone P, Insola A, Tonali P, Di Lazzaro V. Dopamine Dependency of Oscillations between Subthalamic Nucleus and Pallidum in Parkinson's Disease. J Neurosci.

2001;21: 1033–1038. Available: http://www.jneurosci.org/content/21/3/1033.abstract

43. Levy R. Dependence of subthalamic nucleus oscillations on movement and dopamine in Parkinson's disease. Brain. 2002;125: 1196–1209. doi:10.1093/brain/awf128

44. van Wijk BCM, Cagnan H, Litvak V, Kühn AA, Friston KJ. Generic dynamic causal modelling: An illustrative application to Parkinson's disease. Neuroimage. 2018;181: 818–830. doi:10.1016/J.NEUROIMAGE.2018.08.039

45. Marreiros AC, Cagnan H, Moran RJ, Friston KJ, Brown P. Basal ganglia–cortical interactions in Parkinsonian patients. Neuroimage. 2013;66: 301–310. doi:10.1016/j.neuroimage.2012.10.088

46. Moran RJ, Mallet N, Litvak V, Dolan RJ, Magill PJ, Friston KJ, et al. Alterations in brain connectivity underlying beta oscillations in parkinsonism. Kording KP, editor. PLoS Comput Biol. 2011;7: e1002124. doi:10.1371/journal.pcbi.1002124

47. Holgado AJN, Terry JR, Bogacz R. Conditions for the generation of beta oscillations in the subthalamic nucleus-globus pallidus network. J Neurosci. 2010;30: 12340–52. doi:10.1523/JNEUROSCI.0817-10.2010

48. Terman D, Rubin JE, Yew AC, Wilson CJ. Activity patterns in a model for the subthalamopallidal network of the basal ganglia. J Neurosci. 2002;22: 2963–2976. doi:20026266

49. Park C, Worth RM, Rubchinsky LL. Neural dynamics in Parkinsonian brain: The boundary between synchronized and nonsynchronized dynamics. Phys Rev E. 2011;8319. doi:10.1103/PhysRevE.83.042901

50. Ahn S, Rubchinsky LL, Lapish CC. Dynamical Reorganization of Synchronous Activity Patterns in Prefrontal Cortex-Hippocampus Networks During Behavioral Sensitization. Cereb Cortex. 2014;24: 2553–2561. doi:10.1093/cercor/bht110

51. Tinkhauser G, Torrecillos F, Duclos Y, Tan H, Pogosyan A, Fischer P, et al. Beta burst coupling across the motor circuit in Parkinson's disease. Neurobiol Dis. 2018;117: 217–225. doi:10.1016/j.nbd.2018.06.007

52. Cagnan H, Duff EP, Brown P. The relative phases of basal ganglia activities dynamically shape effective connectivity in Parkinson's disease. Brain. 2015;138: 1667–1678. doi:10.1093/brain/awv093

53. van Ede F, Quinn AJ, Woolrich MW, Nobre AC. Neural Oscillations: Sustained Rhythms or Transient Burst-Events? Trends Neurosci. 2018;41: 415–417. doi:10.1016/J.TINS.2018.04.004

54. Rubchinsky LL, Park C, Worth RM. Intermittent neural synchronization in Parkinson's disease. Nonlinear Dyn. 2012;68: 329–346. doi:10.1007/s11071-011-0223-z

55. Bolam JP, Hanley JJ, Booth PA, Bevan MD. Synaptic organisation of the basal ganglia. J Anat. 2000; 527–42. doi:10.1046/j.1469-7580.2000.19640527.x

56. Albin RL, Young AB, Penney JB. The functional anatomy of basal ganglia disorders. Trends Neurosci. 1989;12: 366–375. doi:10.1016/0166-2236(89)90074-X

57. Alexander GE, DeLong MR, Strick PL. Parallel Organization of Functionally Segregated Circuits Linking Basal Ganglia and Cortex. Annu Rev Neurosci. 1986;9: 357–381. doi:10.1146/annurev.ne.09.030186.002041

58. Beaumont MA, Zhang W, Balding DJ. Approximate Bayesian Computation in Population Genetics. Genetics. 2002;162: 2025 LP – 2035. Available: http://www.genetics.org/content/162/4/2025.abstract

59. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MP. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. J R Soc Interface. 2009;6: 187–202. doi:10.1098/rsif.2008.0172

60. Lopes da Silva FH, Hoeks A, Smits H, Zetterberg LH. Model of brain rhythmic activity. Kybernetik. 1974;15: 27–37. doi:10.1007/BF00270757

61. Jansen BH, Rit VG. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biol Cybern. 1995;73: 357–366. doi:10.1007/BF00199471

62. Toni T, Stumpf MPH. Simulation-based model selection for dynamical systems in systems and population biology. Bioinformatics. 2009;26: 104–110. doi:10.1093/bioinformatics/btp619

63. Del Moral P, Doucet A, Jasra A. An adaptive sequential Monte Carlo method for approximate Bayesian computation. Stat Comput. 2012;22: 1009–1020. doi:10.1007/s11222-011-9271-y

64. Sisson SA, Fan Y, Tanaka MM. Sequential Monte Carlo without likelihoods. Proc Natl Acad Sci U S A. 2007;104: 1760–5. doi:10.1073/pnas.0607208104

65. Wilkinson DJ. Stochastic modelling for systems biology: Second edition. Stochastic Modelling for Systems Biology: Second Edition. Chapman and Hall/CRC; 2011.

66. Baritompa B, Hendrix EMT. On the Investigation of Stochastic Global Optimization Algorithms. J Glob Optim. 2005;31: 567–578. doi:10.1007/s10898-004-9966-0

67. Tinkhauser G, Pogosyan A, Tan H, Herz DM, Kühn AA, Brown P. Beta burst dynamics in

Parkinson's disease OFF and ON dopaminergic medication. Brain. 2017;140: 2968–2981. doi:10.1093/brain/awx252

68. Cruz A V., Mallet N, Magill PJ, Brown P, Averbeck BB. Effects of dopamine depletion on information flow between the subthalamic nucleus and external globus pallidus. 2011;106: 2012–2023. doi:10.1152/jn.00094.2011

69. Garrido MI, Kilner JM, Kiebel SJ, Stephan KE, Friston KJ. Dynamic causal modelling of evoked potentials: A reproducibility study. Neuroimage. 2007;36: 571–580. doi:10.1016/J.NEUROIMAGE.2007.03.014

70. Boly M, Moran R, Murphy M, Boveroux P, Bruno M-A, Noirhomme Q, et al. Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness. J Neurosci. 2012;32: 7082–90. doi:10.1523/JNEUROSCI.3769-11.2012

71. 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. Jbabdi S, editor. PLOS Comput Biol. 2018;14: e1006375. doi:10.1371/journal.pcbi.1006375

72. Friston KJ, Kahan J, Biswal B, Razi A. A DCM for resting state fMRI. Neuroimage. 2014;94: 396–407. doi:10.1016/J.NEUROIMAGE.2013.12.009

73. Rowe DL, Robinson PA, Rennie CJ. Estimation of neurophysiological parameters from the waking EEG using a biophysical model of brain dynamics. J Theor Biol. 2004;231: 413–433. doi:10.1016/J.JTBI.2004.07.004

74. Lemaréchal J-DD, George N, David O. Comparison of two integration methods for dynamic causal modeling of electrophysiological data. Neuroimage. 2018;173: 623–631. doi:10.1016/j.neuroimage.2018.02.031

75. Hadida J, Sotiropoulos SN, Abeysuriya RG, Woolrich MW, Jbabdi S. Bayesian Optimisation of Large-Scale Biophysical Networks. Neuroimage. 2018;174: 219–236. doi:10.1016/J.NEUROIMAGE.2018.02.063

76. Aponte EA, Raman S, Sengupta B, Penny WD, Stephan KE, Heinzle J. mpdcm: A toolbox for massively parallel dynamic causal modeling. J Neurosci Methods. 2016;257: 7–16. doi:10.1016/J.JNEUMETH.2015.09.009

77. Friston K, Mattout J, Trujillo-Barreto N, Ashburner J, Penny W. Variational free energy and the Laplace approximation. Neuroimage. 2007;34: 220–234. doi:10.1016/J.NEUROIMAGE.2006.08.035

78. Arakaki T, Barello G, Ahmadian Y. Inferring neural circuit structure from datasets of heterogeneous tuning curves. Marinazzo D, editor. PLOS Comput Biol. 2019;15: e1006816. doi:10.1371/journal.pcbi.1006816

79. Breakspear M. “Dynamic” Connectivity in Neural Systems. Neuroinformatics. 2004;2: 205–226. doi:10.1385/NI:2:2:205

80. Cabral J, Luckhoo H, Woolrich M, Joensson M, Mohseni H, Baker A, et al. Exploring mechanisms of spontaneous functional connectivity in MEG: How delayed network interactions lead to structured amplitude envelopes of band-pass filtered oscillations. Neuroimage. 2014;90: 423–435. doi:10.1016/j.neuroimage.2013.11.047

81. Baker AP, Brookes MJ, Rezek IA, Smith SM, Behrens T, Probert Smith PJ, et al. Fast transient networks in spontaneous human brain activity. Elife. 2014;3. doi:10.7554/eLife.01867

82. O’Neill GC, Tewarie P, Vidaurre D, Liuzzi L, Woolrich MW, Brookes MJ. Dynamics of large-scale electrophysiological networks: A technical review. Neuroimage. 2018;180: 559–576. doi:10.1016/J.NEUROIMAGE.2017.10.003

83. Tognoli E, Kelso JAS. The metastable brain. Neuron. 2014;81: 35–48. doi:10.1016/j.neuron.2013.12.022

84. Rabinovich MI, Varona P. Robust transient dynamics and brain functions. Front Comput Neurosci. 2011;5: 24. doi:10.3389/fncom.2011.00024

85. Papadopoulou M, Cooray G, Rosch R, Moran R, Marinazzo D, Friston K. Dynamic causal modelling of seizure activity in a rat model. Neuroimage. 2017;146: 518–532. doi:10.1016/J.NEUROIMAGE.2016.08.062

86. Reis C, Sharott A, Magill PJ, van Wijk BCM, Parr T, Zeidman P, et al. Thalamocortical dynamics underlying spontaneous transitions in beta power in Parkinsonism. Neuroimage. 2019;193: 103–114. doi:10.1016/J.NEUROIMAGE.2019.03.009

87. Friston KJ, Bastos AM, Oswal A, van Wijk B, Richter C, Litvak V. Granger causality revisited. Neuroimage. 2014;101: 796–808. doi:10.1016/j.neuroimage.2014.06.062

88. Moran RJ, Stephan KE, Dolan RJ, Friston KJ. Consistent spectral predictors for dynamic causal models of steady-state responses. NeuroImage. 2011. doi:10.1016/j.neuroimage.2011.01.012

89. Ermentrout B. Simulating, analyzing, and animating dynamical systems a guide to XPPAUT for researchers and students [Internet]. Philadelphia, Pa. : Society for Industrial and Applied

Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 2002.; 2002. Available: https://search.library.wisc.edu/catalog/9985509453602122

90. Excoffier CLDWL. Bayesian Computation and Model Selection in Population Genetics. 2009; Available: http://arxiv.org/abs/0901.2231

91. Liepe J, Kirk P, Filippi S, Toni T, Barnes CP, Stumpf MPH. A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation. Nat Protoc. 2014;9: 439–456. doi:10.1038/nprot.2014.025

92. Ratmann O, Andrieu C, Wiuf C, Richardson S. Model criticism based on likelihood-free inference, with an application to protein network evolution. Proc Natl Acad Sci U S A. 2009;106: 10576–81. doi:10.1073/pnas.0807882106

93. Turner BM, Sederberg PB. Approximate Bayesian computation with differential evolution. J Math Psychol. 2012;56: 375–385. doi:10.1016/j.jmp.2012.06.004

94. Sunnåker M, Busetto AG, Numminen E, Corander J, Foll M, Dessimoz C. Approximate Bayesian computation. PLoS Comput Biol. 2013;9: e1002803. doi:10.1371/journal.pcbi.1002803

95. Lintusaari J, Gutmann MU, Dutta R, Kaski S, Corander J. Fundamentals and Recent Developments in Approximate Bayesian Computation. Syst Biol. 2016;66: syw077. doi:10.1093/sysbio/syw077

96. Robert CP, Cornuet J-M, Marin J-M, Pillai NS. Lack of confidence in approximate Bayesian computation model choice. Proc Natl Acad Sci U S A. 2011;108: 15112–7. doi:10.1073/pnas.1102900108

97. Marin J-M, Pillai NS, Robert CP, Rousseau J. Relevant statistics for Bayesian model choice. J R Stat Soc Ser B (Statistical Methodol. 2014;76: 833–859. doi:10.1111/rssb.12056

98. Lohmann G, Erfurth K, Müller K, Turner R. Critical comments on dynamic causal modelling. Neuroimage. 2012;59: 2322–2329. doi:10.1016/J.NEUROIMAGE.2011.09.025

99. Templeton AR. Statistical hypothesis testing in intraspecific phylogeography: nested clade phylogeographical analysis vs. approximate Bayesian computation. Mol Ecol. 2009;18: 319–31. doi:10.1111/j.1365-294X.2008.04026.x

100. Friston K, Daunizeau J, Stephan KE. Model selection and gobbledygook: Response to Lohmann et al. Neuroimage. 2013;75: 275–278. doi:10.1016/J.NEUROIMAGE.2011.11.064

101. Box GEP. Science and Statistics. J Am Stat Assoc. 1976;71: 791–799. doi:10.1080/01621459.1976.10480949

102. David O, Friston KJ. A neural mass model for MEG/EEG: Neuroimage. 2003;20: 1743–1755. doi:10.1016/j.neuroimage.2003.07.015

103. Del Moral P, Doucet A, Jasra A. An adaptive sequential Monte Carlo method for approximate Bayesian computation. Stat Comput. 2012;22: 1009–1020. doi:10.1007/s11222-011-9271-y

104. Cameron CA, Windmeijer FAG. An R-squared measure of goodness of fit for some common nonlinear regression models. J Econom. 1997;77: 329–342. doi:10.1016/S0304-4076(96)01818-0

105. Dean TA, Singh SS, Jasra A, Peters GW. Parameter Estimation for Hidden Markov Models with Intractable Likelihoods. Scand J Stat. 2014;41: 970–987. doi:10.1111/sjos.12077

106. Silverman BW. Density estimation for statistics and data analysis. Chapman and Hall/CRC; 2003.

107. Bowman AW. An Alternative Method of Cross-Validation for the Smoothing of Density Estimates. Biometrika. 1984;71: 353. doi:10.2307/2336252

108. Nelsen RB. An Introduction to Copulas [Internet]. New York, NY: Springer New York; 1999. doi:10.1007/978-1-4757-3076-0

109. Grelaud A, Robert CP, Marin J-M, Rodolphe F, Taly J-F. ABC likelihood-free methods for model choice in Gibbs random fields. Bayesian Anal. 2009;4: 317–335. doi:10.1214/09-BA412

110. Kass RE, Raftery AE. Bayes Factors. J Am Stat Assoc. 1995;90: 773–795. Available: http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1995.10476572

111. MacKay DJC. Information theory, inference, and learning algorithms [Internet]. Cambridge University Press; 2003. Available: http://www.inference.org.uk/itprnn/book.html

112. Magill PJ, Sharott A, Bolam JP, Brown P. Brain State–Dependency of Coherent Oscillatory Activity in the Cerebral Cortex and Basal Ganglia of the Rat. J Neurophysiol. 2004;92: 2122–36. doi:10.1152/jn.00333.2004

113. Magill PJ, Pogosyan A, Sharott A, Csicsvari J, Bolam JP, Brown P. Changes in functional connectivity within the rat striatopallidal axis during global brain activation in vivo. J Neurosci. 2006;26: 6318–6329. doi:10.1523/JNEUROSCI.0620-06.2006

114. Paxinos G, Watson C. The rat brain in stereotaxic coordinates. Elsevier; 2007.

115. Steriade M. Corticothalamic resonance, states of vigilance and mentation. Neuroscience. 2000;101: 243–276. doi:10.1016/S0306-4522(00)00353-5

116. Halliday DM. Nonparametric directionality measures for time series and point process data. J Integr Neurosci. 2015;14: 253–277. doi:10.1142/S0219635215300127

117. West TO, Halliday DM, Bressler S, Farmer SF, Litvak V. Measuring directed functional connectivity in neural data using Non-Parametric Directionality Analysis: validation and comparison with Non-Parametric Granger Causality. bioRxiv. 2019; 526566. doi:10.1101/526566

118. Aslan B, Zech G. New test for the multivariate two-sample problem based on the concept of minimum energy. J Stat Comput Simul. 2005;75: 109–119. doi:10.1080/00949650410001661440

119. Nambu A, Tachibana Y, Chiken S. Cause of parkinsonian symptoms: Firing rate, firing pattern or dynamic activity changes? Basal Ganglia. 2015;5: 1–6. doi:10.1016/j.baga.2014.11.001

120. Jahfari S, Waldorp L, van den Wildenberg WPM, Scholte HS, Ridderinkhof KR, Forstmann BU. Effective Connectivity Reveals Important Roles for Both the Hyperdirect (Fronto-Subthalamic) and the Indirect (Fronto-Striatal-Pallidal) Fronto-Basal Ganglia Pathways during Response Inhibition. J Neurosci. 2011;31. Available: http://www.jneurosci.org/content/31/18/6891.long

121. Chu HY, McIver EL, Kovaleski RF, Atherton JF, Bevan MD. Loss of Hyperdirect Pathway Cortico-Subthalamic Inputs Following Degeneration of Midbrain Dopamine Neurons. Neuron. 2017;95: 1306-1318.e5. doi:10.1016/j.neuron.2017.08.038

122. Bevan MD, Magill PJ, Terman D, Bolam JP, Wilson CJ. Move to the rhythm: oscillations in the subthalamic nucleus–external globus pallidus network. Trends Neurosci. 2002;25: 525–531. doi:10.1016/S0166-2236(02)02235-X

123. Plenz D, Kital ST. A basal ganglia pacemaker formed by the subthalamic nucleus and external globus pallidus. Nature. 1999;400: 677–82. doi:10.1038/23281