Supplementary Material for "A computational model of Major Depression: The role of glutamate dysfunction on cingulate-frontal network dynamics"

September 16, 2015

Contents

1 Full firing rate model

We model the interaction between the ventral anterior cingulate cortex (vACC) and the dorso-lateral prefrontal cortex (dlPFC) via four coupled differential

equations. The equations are

$$\tau_e \dot{r_e^v} = -r_e^v + \phi_e\Big(G_{ee} f_D r_e^v - G_{ei} r_i^v + I_e^v + I_e f_D\Big), \tag{S.1}$$

$$\tau_i \dot{r_i^v} = -r_i^v + \phi_i\Big(G_{ie} f_D r_e^v - G_{ii} r_i^v + G_x r_e^d + I_i^v + I_i f_D\Big), \tag{S.2}$$

$$\tau_e \dot{r_e^d} = -r_e^d + \phi_e\Big(G_{ee} r_e^d - G_{ei} r_i^d + I_e^d + I_e\Big), \tag{S.3}$$

$$\tau_i \dot{r_i^d} = -r_i^d + \phi_i\Big(G_{ie} r_e^d - G_{ii} r_i^d + G_x r_e^v + I_i^d + I_i\Big), \tag{S.4}$$

where r_e^v (r_i^v) represents the mean firing rate of excitatory principal neurons (inhibitory interneurons) in the ventral network (vACC). Similarly, r_e^d (r_i^d) represents the mean firing rate of excitatory principal neurons (inhibitory interneurons) in the dorsal network (dlPFC). The inputs to a population of excitatory (inhibitory) neurons is converted into a firing rate via the nonlinear activation function ϕ_e (ϕ_i), which we have chosen of the form

$$\phi(x) = A \begin{cases} 0 & \text{if } x < 0 \\ x^2 & \text{if } 0 \le x \le 1 \\ 2\sqrt{x - 3/4} & \text{if } x > 1, \end{cases} \tag{S.5}$$

where we take $A = 20$. This choice captures the expansively nonlinearity of the neuronal response at low firing rates, and the decreased gain at higher rates in the simplest way possible.

The strength of recurrent connections within the same cortical area are given by: G_{ee} (E-to-E), G_{ie} (E-to-I), G_{ei} (I-to-E) and G_{ii} (I-to-I), whereas the connection strength across areas is G_x. Note that the cross-area interaction is via excitatory synaptic activation of inhibitory neurons. Therefore increased excitatory activation in one of the two areas tends to decrease the excitatory activation of the other via recruitment of local inhibitory interneurons. We model both non-selective (I_e, I_i) and selective (I_e^v, I_i^v,I_e^d, I_i^d) external inputs. The increased excitation in the ventral area due to slowed re-uptake of glutamate is modelled via the term f_D which takes on a value of one in the "healthy" condition. Finally, excitation and inhibition have characteristic time-scales of $\tau_e = \tau_i = \tau = 20$ms.

We seek a regime in Eqs.S.1-S.4 in which each area is bistable between a low- and a high-activity state, and in which there is strong competition. This means that when one area is activated, e.g. ventral, it should shut down the other. In this way we can model the robust switching on and off of activity

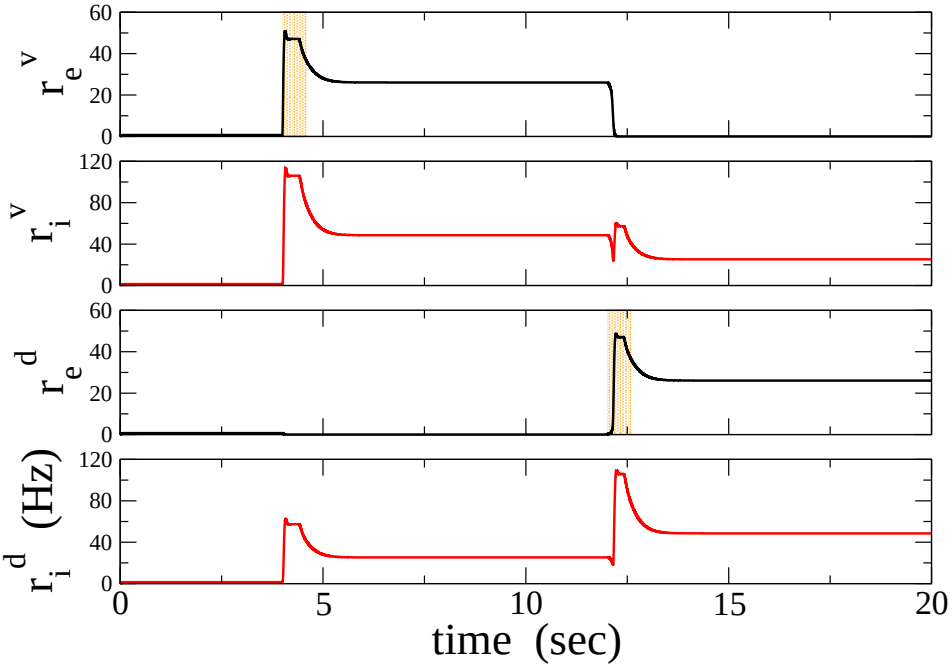

Fig. S.1: An example of the competitive dynamics between the ventral and dorsal areas. The shaded orange regions show periods of elevated selective input, e.g. from $t = 4$sec to $t = 4.6$sec there is an increase in the selective input I_e^v. From $t = 12$sec to $t = 12.6$sec the dorsal input is increased, turing on the dorsal network and turing off the ventral one. Parameters are: $G_{ee} = 0.09$, $G_{ie} = 0.04$, $G_{ei} = 0.0275$, $G_{ii} = 0.0075$, $G_x = 0.025$, $I_i = 0.1$, $I_i^d = I_i^v = 0$ and f_D. I_e^v and I_e^d are set to 0.65 during 400ms of the shaded period and then decay to zero exponentialy with a time constant of 100ms.

related to emotional (ventral) or cognitive (dorsal) content. An example of such dynamics is shown in Fig.S.1.

Fig.S.2 shows the firing rate of the excitatory population in the ventral area as a function of the selective excitatory input to that population. The upper panel shows that there is indeed a region of bistability (coexistence of solid black lines for the same value of I_e^v). There is, however an oscillatory instability nearby (dashed orange line). Therefore, the region of bistability is not strictly delineated by the two saddle-node (SN) bifurcations (see Fig.S.2), but rather the rightmost SN and a Hopf bifurcation. When the parameter f_D is increased, the increased excitation broadens the region between the two

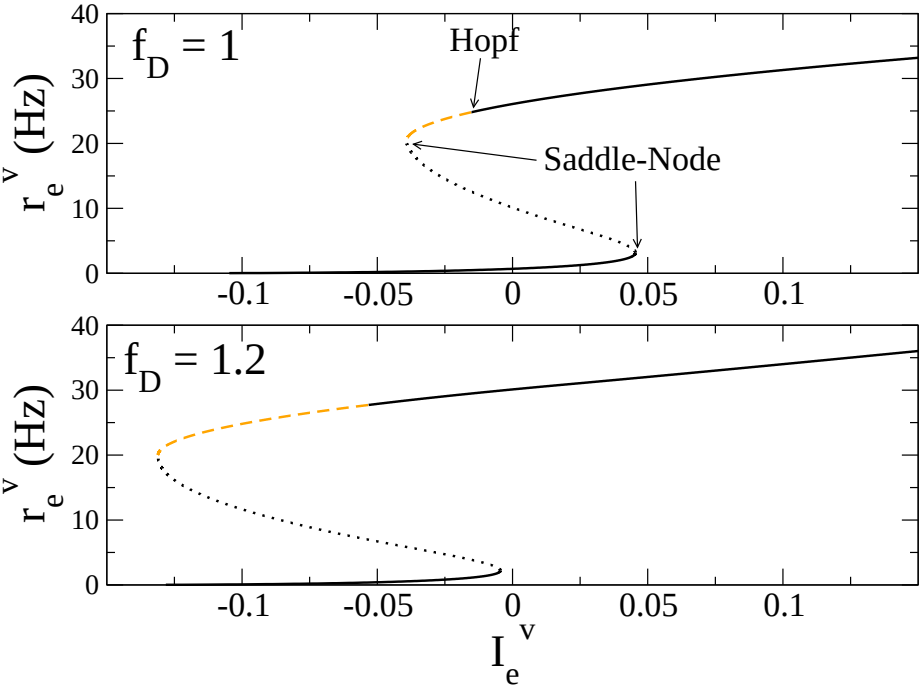

Fig. S.2: The bifurcation diagram for the firing rate of the excitatory population in the ventral area. Upper: The "healthy" case with $f_D = 1$. Lower: The "depressed" case with $f_D = 1.2$. Solid black line: stable fixed point. Dotted black line: unstable (saddle) point. Dashed orange line: unstable (spiral) point. Parameters are: $G_{ee} = 0.09$, $G_{ie} = 0.04$, $G_{ei} = 0.0275$, $G_{ii} = 0.0075$, $G_x = 0.025$, $I_i = 0.1$, $I_e = 0.163$, $I_e^d = I_i^d = I_i^v = 0$.

SNs, but not the region of bistability, because the oscillatory instability has grown, see bottom panel.

Fig.S.3 shows the bifurcation diagram for all four populations when $f_D = 1.1$. The diagram for r_e^d makes clear that when the selective input to the ventral area is increased, the activity of the excitatory population in the dorsal area is almost completely shut down. Therefore, in the regime in which we study the model, each area acts as a self-contained switch; once it is switched on it strongly inhibits the activity of the other area.

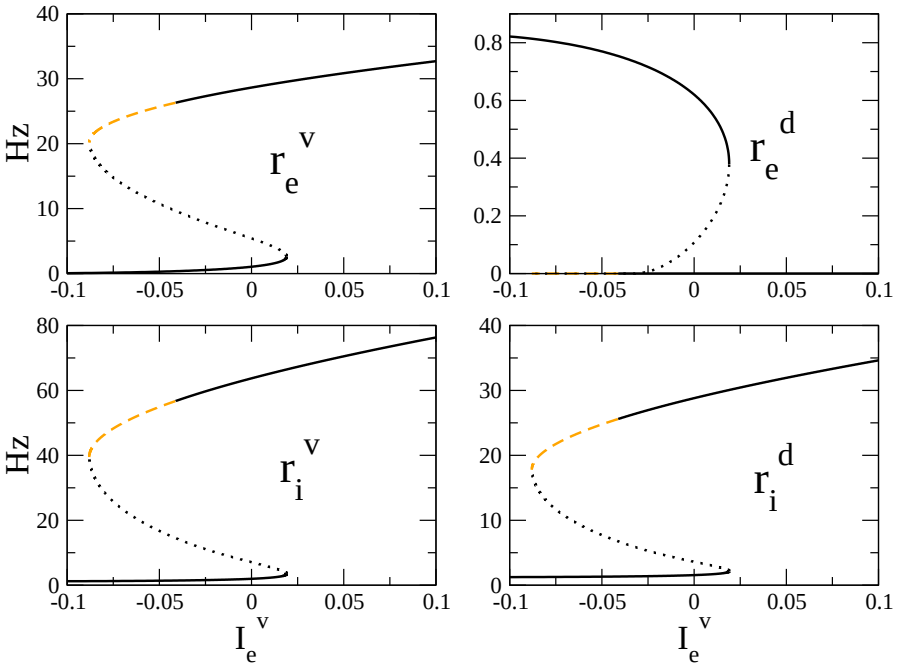

Fig. S.3: The bifurcation diagram for the activity of all four populations as the selective input to the excitatory population in the ventral area is varied. Parameters are the same as in Fig.S.2 with $f_D = 1.1$. These are the same parameters as in Fig.7 of the main manuscript.

2 Stability of the solutions

2.1 4D system and 2D reduction

The stability of the fixed-point solutions in Eqs.S.1-S.4 is determined by assuming small perturbations to the solutions and determining whether or not they grow in time. This is a standard linear stability analysis. In the case of our 4D system, the exponent of linear perturbations λ (the eigenvalue) is found by assuming the ansatz $\mathbf{r} = \mathbf{r}_0 + \delta\mathbf{r}e^{\lambda t}$, where the bold font indicates a vector (of all of the firing rates), $\mathbf{r_0}$ is the fixed point solution, and the perturbation $\delta\mathbf{r} \ll 1$. Plugging this ansatz into Eqs.S.1-S.4 yields the following characteristic polynomial for λ

$$\Big[(\tau\lambda + 1 - G_{ee}f_D\phi_1^{'})(\tau\lambda + 1 + G_{ii}\phi_2^{'}) + G_{ei}G_{ie}f_D\phi_1^{'}\phi_2^{'}\Big] \cdot \Big[(\tau\lambda + 1 - G_{ee}\phi_3^{'})$$
$$(\tau\lambda + 1 + G_{ii}\phi_4^{'}) + G_{ei}G_{ie}\phi_3^{'}\phi_4^{'}\Big] = G_x^2G_{ei}^2\phi_1^{'}\phi_2^{'}\phi_3^{'}\phi_4^{'}, \tag{S.6}$$

where $\phi_1^{'}$ is the gain (slope of the transfer function at the fixed point) of the excitatory ventral population, $\phi_2^{'}$ of the inhibitory ventral population, $\phi_3^{'}$ of the excitatory dorsal and $\phi_4^{'}$ of the inhibitory dorsal population. The left-hand side of Eq.S.6 is simply the product of the characteristic polynomials for the two areas separately. That is, when the right-hand side is zero the stability of the each area is independent of the other. This occurs precisely when $G_x = 0$, i.e. when there is no interaction. When this is the case, the first term (second term) is zero when there is a Saddle-Node (SN) or Hopf (H) bifurcation in the ventral (dorsal) area, the conditions for which are given by Eqs.5 and 6 in the main text respectively. The presence of a finite interaction does not strongly alter the nature nor the position of instabilities of each area. This is particularly true since the interaction term depends on the product of all of the gains in the neuronal transfer functions, which will always be small if one of the excitatory populations is inactivated. This suggests that the stability of the full 4D system can be studied by looking only at the reduced 2D system corresponding to one of the two areas, e.g. the ventral area. This means that the stability of the solutions shown in Fig.S.3 should be well-described by just the first term of the product on the left-hand side of Eq.S.6, i.e. precisely the conditions shown in Eqs.5 and 6 of the main text.

In order to confirm this we have calculated the real and imaginary parts of λ for the full 4D system shown in Fig.S.3, i.e. we solve for $\lambda = \lambda_r + i\omega$ in Eq.S.6, where λ_r is the growth rate of perturbations and ω is their frequency. The growth rate λ_r is shown in Fig.S.4(top) for the same parameters and color-code as in Fig.S.3. Note that in the 4D system there are four distinct eigenvalues (or three when two are complex conjugates) and here we only show the most unstable one. The values of the eigenvalue corresponding to the upper and lower branches of the bifurcation diagram in Fig.S.3 are labeled for clarity. In the same figure we also draw, in red, λ_r evaluated solely from the 2D system. Note that it is in good agreement with the full, 4D system, with only slight discrepancies in the upper branch. The same is true of the frequency of perturbations, shown in Fig.S.4(bottom).

Of particular importance to the conclusions of the paper is the fact that the amplitude of slow (theta) oscillations increases with decreasing I_e^v. Fig.7 from the main manuscript also shows that the increase in amplitude is accompanied by a decrease in oscillatory frequency. Fig.S.4 explains why this is so. The Hopf bifurcation (shown with the open circle), which indicates the onset of an oscillatory instability, occurs as I_e^v is decreased from high values. This means that near the bifurcation, the growthrate of instabilities must necessarily increase for decreasing I_e^v, i.e. the curve for λ_r has a negative slope.

How is the amplitude of fluctuations related to the largest growthrate of the linearized system λ_r? If we include an additive white noise term with amplitude σ to the equation for the excitatory ventral population, then the autocorrelation (AC) of the firing rate r_e^v is just

$$AC(\tau) = \frac{\sigma^2}{|\lambda_r|} e^{-|\lambda_r|t} \cos \omega t, \tag{S.7}$$

as long as $\lambda_r < 0$, where we define the autocorrelation of a signal $x(t)$ as

$$AC(\tau) = \int_{-\infty}^{\infty} dt (x(t) - x_0)(x(t+\tau) - x_0) \tag{S.8}$$

and x_0 is the mean of $x(t)$. Fig.S.5 shows the results of three simulations where the excitatory dorsal population is driven by a source of white noise with ampltiude σ. The AC calculated from the numerical time series (solid line) is compared to Eq.S.7 (dotted line), and the agreement is good.

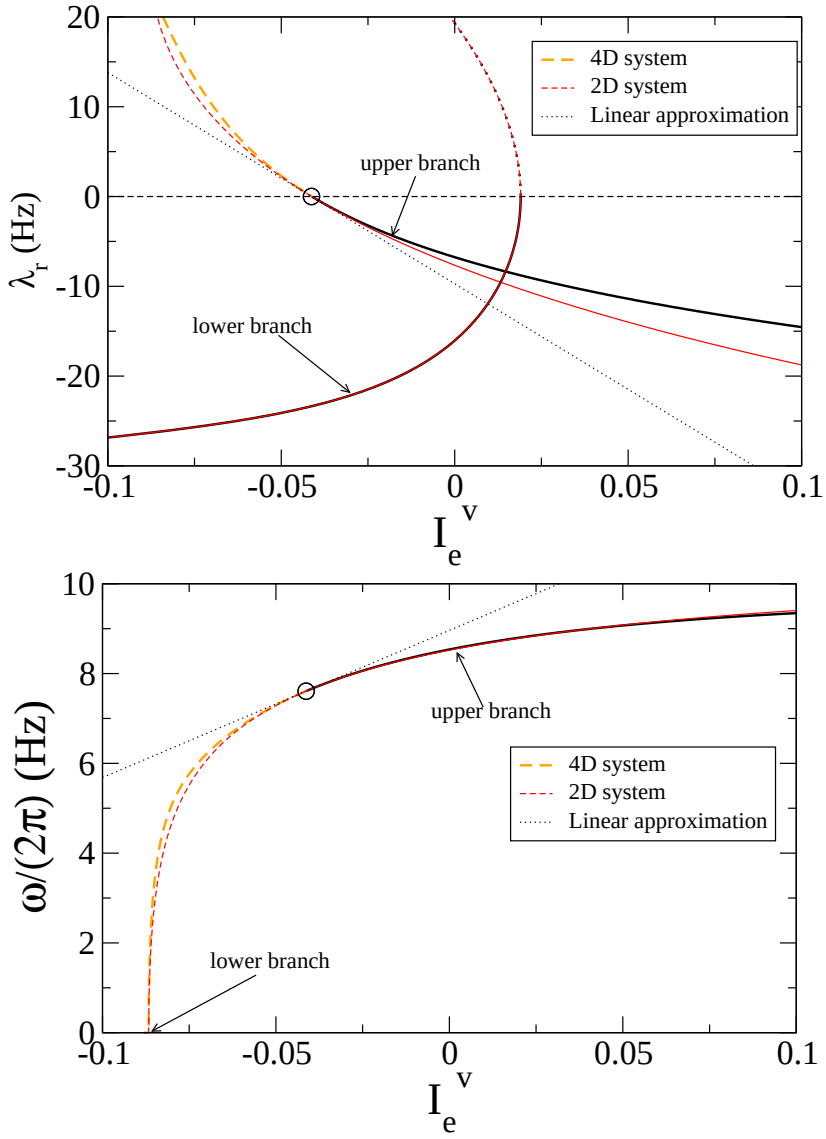

Fig. S.4: The real part (top) and imaginary part (bottom) of the eigenvalue with largest real part of the 4D system and 2D system (red) for the parameters used in Fig.S.3. The dotted black line is the linear approximation of the slope (see below).

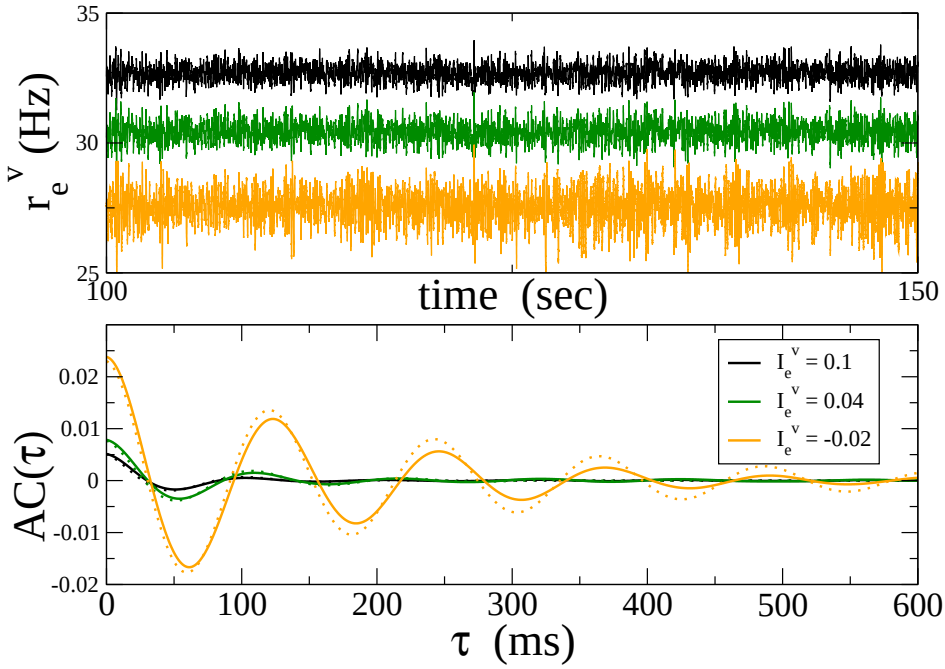

Fig. S.5: The autocorrelation (AC) of activity can be predicted from the eigenvalue of the linearized syste. Top: The firing rate r_e^v for three separate simulations with three different values of I_e^v. The population is driven by a white noise term with amplitude σ. Bottom: The AC of the three time series from simulations (solid lines), and predictions from Eq.S.7 (dotted lines). Note that the random seed for the noise source is identical for the three simulations.

2.2 Robustness of results

Although the dependence of slow oscillations on the selective input I_e^v described above is observed in numerical simulations of the model, and explained by the underlying eigenvalues of the linearized system, we can ask how general and robust a phenomenon it is. Specifically, we have chosen a specific set of parameters. Does this scenario hold for other parameter values and, if so, under what conditions? The simplest answer is that when the Hopf bifurcation occurs close to the Saddle-Node (SN) bifurcation, then the frequency must decrease for decreasing I_e^v. This is because the frequency must go to zero at the SN (there is no twist at a SN), and in the 2D system any other possibility is topologically prohibited. In the 4D system something more exotic could potentially occur, although we have seen that the stability of solutions is likely to mirror that of the 2D system. Furthermore, the growthrate must also increase for decreasing I_e^v, that is, the upper branch should stabilize as I_e^v is increased and hence the amplitude of fluctuation-driven oscillatory activity must decrease. This is because the condition for the Hopf bifurcation is a large enough excitatory gain, which in general decreases as the firing rate goes up. This is the case if the neuronal transfer function ϕ has a single inflection point, for example.

We can ask if the results still hold even if the Hopf bifurcation is not "close" to the SN, in the mathematical sense. In order to address this question we calculate the slope of the eigenvalue as a function of the selective input to the ventral area, i.e. $\frac{\partial \lambda}{\partial I_e^v}$. The upshot is that we find that

$$\begin{aligned} \frac{\partial \lambda_r}{\partial I_e^v} &= A\phi_e'' - B\phi_i'', \\ \frac{\partial \omega}{\partial I_e^v} &= -D\phi_e'' + E\phi_i'', \end{aligned} \tag{S.9}$$

where A, B, C and D are all positive coefficients for most parameter values. This means that changes in the amplitude and frequency of slow oscillations depend essentially on the curvature of the neuronal transfer function. Specifically, the trends we report here are most robust when the excitatory neuronal response function has negative curvature and the inhibitory one positive curvature. This means the excitatory population must be in a high-activity state whereas the inhibitory population is in the fluctuation-driven regime where the response exhibits an expansive nonlinearity. The interested

reader can read through section 2.2.1 for details. Otherwise one can skip to section 3.

2.2.1 Details of calculation of how eigenvalue changes with external input

Note that it is not sufficient to naively expand the first term in Eq.S.6 in λ and I_e^v. The reason is that changing I_e^v changes the value of the fixed point, which therefore must be taken into account. There are several ways to do this. We choose to use an ansatz which is similar to a standard linear stability analysis with one difference: we allow for a slowly-varying amplitude of the perturbation which reflects the marginal change in stability (growthrate and frequency) as I_e^v is varied. The method is similar in spirit to the so-called multi-scale analysis [1] (for a detailed example with nonlinear effects see [2]).

Specifically, we write out 2D system as

$$\dot{\mathbf{r}} = \mathbf{f}(\mathbf{r}), \tag{S.10}$$

and consider the ansatz $\mathbf{r} = \mathbf{r_0} + \epsilon\Big(A(T)\mathbf{r_1}e^{\lambda t} + c.c.\Big) + ...$, where $\mathbf{r_0}$ is the fixed-point solution and $\epsilon \ll 1$ is a small parameter. We also assume $I_e^v = I_{e0}^v + \epsilon^2 I_{e1}^v$ and define a slow time $T = \epsilon^2 t$. Plugging the ansatz into Eq.S.10, we collect all terms by their order in ϵ. At order ϵ we just recover the simple linear stability problem $\mathcal{L}_0\mathbf{r_1} = 0$ which leads to the quadratic polynomial in λ, see first bracketed term on l.h.s. of Eq.S.6. The vector $\mathbf{r_1}$ is therefore the right eigenvector of $\mathcal{L}_0$. If we denote by $\mathbf{r}^\dagger$ the left eigenvector, then we find after some algebra, at order ϵ^3

$$\partial_T A = \frac{\bar{\mathbf{r}}^\dagger \cdot \mathbf{N}}{\bar{\mathbf{r}}^\dagger \cdot \mathbf{r_1}} I_{e1}^v A, \tag{S.11}$$

where the overbar means complex conjugate and

$$\mathbf{r_1} = (g_{ei}\phi_e^{'}, -\lambda - 1 + g_{ee}f_D\phi_e^{'})^T, \tag{S.12}$$

$$\mathbf{r}^\dagger = (-g_{ie}f_D\phi_i^{'}, -\bar{\lambda} - 1 + g_{ee}f_D\phi_e^{'})^T, \tag{S.13}$$

$$\mathbf{N} = \left(\phi_e^{''}g_{ei}(1+\lambda)\left[1 + \frac{\phi_e^{'}}{C}\Big(g_{ee}f_D + f_D\phi_i^{'}(g_{ee}g_{ii} - g_{ei}g_{ie})\Big)\right], \phi_i^{''}\left[g_{ii}(1+\lambda) + f_D\phi_e^{'}(g_{ei}g_{ie} - g_{ee}g_{ii})\right]\frac{g_{ie}f_D\phi_e^{'}}{C}\right)^T, \tag{S.14}$$

where $C = (1 - g_{ee} f_D \phi_e^{'})(1 + g_{ii}\phi_i^{'}) + g_{ei} g_{ie} f_D \phi_e^{'} \phi_i^{'}$. The real and imaginary parts of this ratio of vector dot-products are just the slopes of the top and bottom curves in Fig.S.4 respectively. When the formula is evaluated right at the Hopf bifurcation, it produces the straight dotted black lines shown in Fig.S.4 and therefore captures the slope well. What do we hope to gain from such a complex formula? We can use it to show under which conditions the slope of the growth rate is negative and that of the frequency is positive. In fact, from the above formula these slopes can be written as in Eqs.S.9 where A, B, C and D are all positive coefficients for most parameter values. For example, if we evaluate them at the Hopf bifurcation for slow oscillations, and with the additional assumption $g_{ee} g_{ii} - g_{ei} g_{ie} = 0$ then $A = C$ and $B = D$, with

$$A = \frac{g_{ie} g_{ei} f_D \phi_i^{'} \phi_e^{'} (2 + g_{ii}\phi_i^{'})}{2\omega^2 (1 + g_{ii}\phi_i^{'})^2}, \tag{S.15}$$

$$B = \frac{g_{ii}}{2}. \tag{S.16}$$

3 How the region of bistability changes as a function of f_D

We would like to be able to draw general conclusions regarding how the region of bistability changes as we increase the parameter f_D, which is meant to model the affects of MDD. In our simulations we found that increasing f_D lead to an increase in the distance between the two SN bifurcation which would usually delineate the region of bistability. However, the presence of an oscillatory instability between the two SNs actually lead to the region of bistability shrinking as we increased f_D.

Due to the simplicity in our choice of the neuronal transfer function ϕ we are in a position to extract some analytical results. Specifically, for each neuronal population there are two distinct qualitative regimes: 1 - The quadratic regime (Q) where the nonlinearity is expansive and hence the curvature is positive, and 2 - The square-root regime (S) where the nonlinearity is compressive and hence the curvature is negative. Therefore there are four possible regimes in total: QQ, SQ, QS, SS, where the first letter corresponds to the state of the excitatory population. For example, the QQ regime is always the relevant one when firing rates are low in both populations, and conversely when firing rates are high in both then the regime is likely to be SS, etc.

We can see which regimes are relevant for the dynamics we have considered by looking at Fig.S.6, which shows how the SN and Hopf bifurcation lines vary as a function of both the selective external input I_e^v as well as the parameter f_D. The remaining parameters are the same as in Fig.S.3 as well as in Fig.7 of the main text. The regimes are color-coded: QQ (red), SQ (black) and SS (blue). Clearly the SQ regime is the relevant one for the simulations we have carried out in this manuscript. This is precisely the regime in which slow oscillations exhibit increased frequency and reduced amplitude as I_e^v is increased, as we saw in the last section.

Why is this the relevant regime? Recall that in order for the Hopf bifurcation to persist as f_D is increased we require that excitatory firing rates not increase too much. This is because the instability only occurs if the excitatory gain is sufficiently strong and the gain actually decreases as rates increase (saturation). One way to avoid this is to put the inhibitory population in the Q regime, where the gain is very high; this keeps excitatory rates under control. It can be shown that, in general, if the inhibition is linear or saturating, then the region of bistability would actually increase as f_D increases. Only the SQ regime generically shows the opposite trend. We leave a detailed analysis of the different regimes for future work.

4 Why does an oscillatory instability destroy the bistability?

In general, an oscillatory instability leads to a stable oscillatory solution, e.g. a limit cycle. In our model this is not the case. Is this a general phenomenon? Fig.S.7 shows what happens when the full 4D is simulated with an initial condition very close to the unstable branch (dashed orange line) in Fig.S.3, for $I_e^v = -0.07$. There is clearly an oscillatory instability, but instead of approaching a limit cycle solution, the activity eventually reaches the stable node at low rates. In fact, this behavior is generic when a Hopf bifurcation is in the vicinity of a SN bifurcation. The ensuing phenomenology can be completely characterized qualitatively by assuming the Hopf and SN coincide in a so-called Takens-Bogdanov bifurcation. We refer the interested reader to [3], where a detailed analysis is provided.

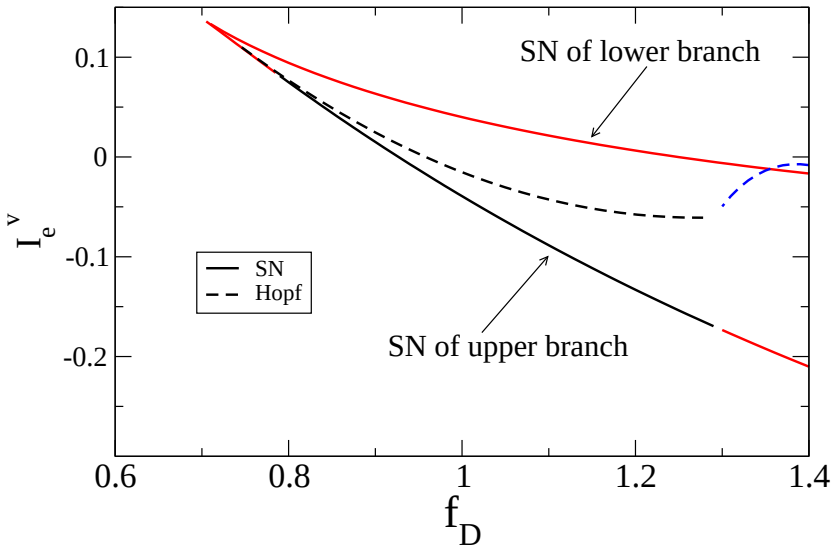

Fig. S.6: Phase diagram for the same parameters as in Fig.S.3. Saddle-node bifurcations occur along the solid lines and Hopf bifurcations along the dashed lines. The color of the line indicates the fixed-point regime: QQ (red), SQ (black) and SS (blue).

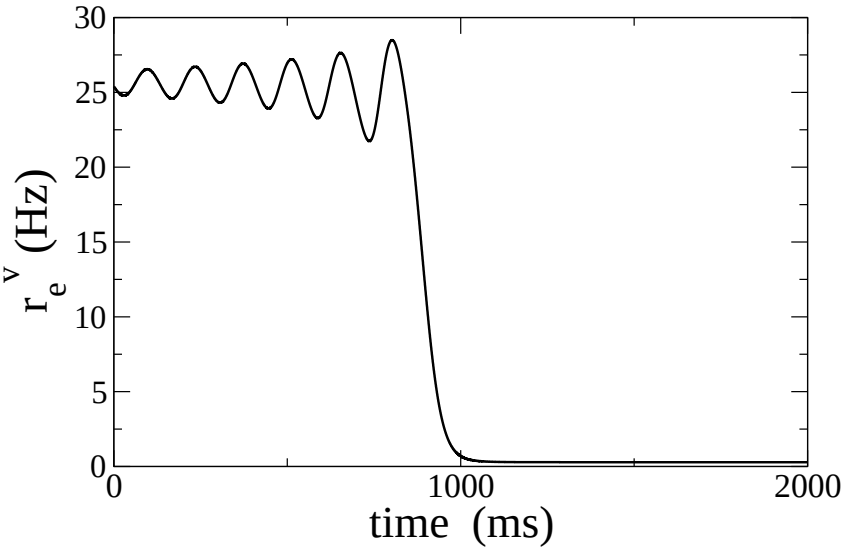

Fig. S.7: Unstable oscillatory activity for $I_e^v = -0.07$. All other parameters are the same as in Fig.S.3.

References

[1] J. Guckenheimer and P. Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Springer Verlag, 1983.

[2] A. Roxin and A. Ledberg. Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Comp. Biol.*, 4:e1000046, 2008.

[3] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer, 2nd edition, 2003.