Analysis and extension of exact mean-field theory with dynamic synaptic currents

Giulio Ruffini[a]

[a]Neuroelectrics, Barcelona

This preprint was compiled on September 3, 2021

Neural mass models such as the Jansen-Rit system provide a practical framework for representing and interpreting electrophysiological activity (1–6) in both local and global brain models (7). However, they are only partly derived from first principles. While the post-synaptic potential dynamics in NMM are inferred from data and can be grounded on diffusion physics (8–10), Freeman's "wave to pulse" sigmoid function (11–13), used to transduce mean population membrane potential into firing rate, rests on a weaker theoretical standing. On the other hand, Montbrió et al (14, 15) derive an exact mean-field theory from a quadratic integrate and fire neuron model under some simplifying assumptions (MPR), connecting microscale neural mechanisms and meso/macroscopic phenomena. The MPR model can be seen to replace Freeman's sigmoid function with a pair of differential equations for the mean membrane potential and firing rate variables, providing a mechanistic interpretation of NMM semi-empirical sigmoid parameters. In doing so, we show it sheds light on the mechanisms behind enhanced network response to weak but uniform perturbations: in the exact mean-field theory, intrinsic population connectivity modulates the steady-state firing rate transfer function in a monotonic manner, with increasing self-connectivity leading to higher firing rates. This provides a plausible mechanism for the enhanced response of densely connected networks to weak, uniform inputs such as the electric fields produced by non-invasive brain stimulation. Next, we complete the MPR model by adding the second order equations for delayed post-synaptic currents, bringing together the MPR and the usual NMM formalisms into a unified exact mean-field theory (NMM2) displaying rich dynamical features. As an example, we analyze the dynamics of a single population model, and a model of two coupled populations with a simple excitation-inhibition (E-I) architecture, showing it displays rich dynamics with limit cycles, period doubling, bursting behavior and enhanced sensitivity to external inputs.

mean-field theory | Neural mass model | Brain stimulation | tES

In their seminal paper, Montbrió et al (14) provide a rigorous derivation of population dynamics from first principles using statistical mechanics methods. Their framework (MPR) provides a derivation of the neuron equations of state relating membrane potential and firing rate, in contrast to the prescription of a sigmoid function as is done in earlier work. This allows to properly interpret the meaning and limitations of the effective theory. In particular, it can help define effective electrical interaction models of neural masses with weak external fields. In their derivation, Montbrió et al make several assumptions, some by necessity and others for simplicity. In particular, they use instantaneous synaptic currents, which makes the model easier to analyze but also unrealistic and hard to relate to neural mass modeling, where synaptic delays play a key role. This has been recently expanded to include first order synapse equations in (16, 17). We

Significance

Several decades of research suggest that weak electric fields influence neural processing. A long-standing question in the field is how networks of neurons process spatially uniform weak inputs that barely affect a single neuron but that produce measurable effects in networks. Answering this can help implement electric field coupling mechanisms in neural mass models of the whole-brain, and better represent the impact of electrical stimulation or ephaptic communication. This issue can be studied using local detailed computational models, but the use of statistical mechanics methods can deliver "mean-field models" to simplify the analysis. Following the steps of Montbrió et al (14, 15), we show that the sensitivity to inputs can be modulated by the intrinsic self-connectivity of a neural population, and produce a more grounded set of equations for neural mass modeling to guide further work.

[2]To whom correspondence should be addressed. E-mail: giulio.ruffini@neuroelectrics.com

complement here the MPR equations with second order post-synaptic kinetics equations as used in NMM (4, 8–10, 18) to represent post-synaptic potentials (PSPs). By bringing these theories together, a new, extended model emerges (NMM2) providing a strong link between microscopic mechanisms and parameters and the mesoscopic population phenomena described by semi-empirical neural mass models.

Neural mass models. Neural mass (semi-empirical) models (NMM), first developed in the early seventies by W. Freeman (19) and F. Lopez de Silva (2), provide a physiologically grounded description of the average synaptic activity and firing rate of a neural population (1–6). NMMs are increasingly used for local and whole-brain modeling in neurology (e.g., epilepsy (20, 21) or Alzheimer's disease (22)) and for understanding and optimizing the effects of transcranial electrical brain stimulation (tES) (6, 7, 23, 24). The central conceptual elements in this framework are the synapse, which is seen to transduce incoming activity (quantified by firing rate) into a mean membrane potential perturbation in the receiving neuron population, and the sigmoid function transforming population membrane potential to output (mean) firing rate with due account for threshold and saturation effects (see (18) for a nice introduction to the Jansen-Rit model). The synaptic filter is instantiated by a second-order system coupling the mean firing rate of arriving signals φ to the mean post-synaptic voltage perturbation u,

$$L[u] = C\varphi \quad [1]$$

$$L[\ ] = \frac{1}{\gamma a}\left(\frac{1}{a}\frac{d^2}{dt^2} + 2\frac{d}{dt} + a\right) \quad [2]$$

or, in the equivalent first order form,

$$\dot{u} = z \quad [3]$$

$$\dot{z} = C\gamma a^2\varphi - 2az - a^2 u \quad [4]$$

where $u(t)$ is the mean post-synaptic voltage perturbation produced by the synapse, $\varphi(t)$ is the mean input firing rate and C is the dimensionless connectivity constant quantifying the mean number of synapses per neuron in the receiving population. The action of the L operator can also be represented by its inverse, a convolution with its Green's function $h(t) = H(t)\gamma a^2 t e^{-at}$ (with $H(t)$ the Heaviside step function) (18).

These linear ordinary differential equations can be generalized to account for different rising and decay times (8, 10). The parameters a and γ describe the delay time scale $1/a$ in seconds and the amplification factor in V/Hz. Traditionally (18), the parameter A (V) is used instead of γ, with $A = \gamma a$. The parameter C is dimensionless and quantifies the average number of synapses per neuron in the receiving population. These synaptic transmission equations need to be complemented by a relationship between the membrane potential of the neuron and its firing rate (the neuron "equation of state"). Freeman proposed the sigmoid as a simple model capturing essential properties of the response of neuron populations from inputs (11–13), based on modeling insights and empirical observations. Note this is a static relation between membrane potential and firing rate. Each neuronal population $\mathcal{P}_n$ converts the sum v_n of the membrane perturbations from each of the incoming synapses or external perturbations (e.g., an external electric field) to an output firing rate (φ_n) in a non-linear manner using Freeman's sigmoid function,

$$\begin{aligned} v_n(t) &= \lambda_n \cdot E(t) + \hat{\eta}_n + \sum_{s \rightsquigarrow \mathcal{P}_n} u_s(t) \\ \varphi_n(t) &= \sigma_n(v_n(t)) = \frac{2\varphi_0}{1 + e^{r(v_0 - v_n(t))}} \end{aligned} \quad [5]$$

where φ_0 is half of the maximum firing rate of each neuronal population, v_0 is the value of the potential when the firing rate is φ_0 and r determines the slope of the sigmoid at the central symmetry point (v_0, φ_0) (we have omitted an n subscript in these population dependent parameters for simplicity). The sum is over synapses *onto* population $\mathcal{P}_n$, which we express as $\sum_{s \rightsquigarrow \mathcal{P}_n}$. The effect of an electric field is represented by a

membrane potential perturbation, $\delta v_n = \lambda_n \cdot E$, with λ_n an effective, semi-empirical vector parameter that represents the first order dipole coupling of field and neuron.

The presence of semi-empirical "lumped" quantities such as φ_0 and λ highlight the fact that NMMs are not currently derived from microscopic models, but bring together a mix of theory and observational relations. This means, for example, that even if we have access to a detailed single neuron compartment model providing an electric dipole coupling constant, it is not obvious how to use it in an NMM. To better understand the effective coupling to external inputs, we consider a small perturbation of the population mean membrane potential around an equilibrium point $u_s(t) = u_s^0 + \delta u_s(t)$, that is, $v_n = v_n^0 + \delta v_n + \lambda_n \cdot E(t)$ with $\lambda_n \cdot E(t)$ a weak field perturbation and $v_n^0 = \sum_{s \rightsquigarrow \mathcal{P}_n} u_s^0, \quad \delta v_n = \sum_{s \rightsquigarrow \mathcal{P}_n} \delta u_s$. Then, we can linearize the sigmoid around the fixed point (Taylor expansion to first order) to obtain

$$\hat{L}_s[\delta u_s(t)] \quad = \quad C_{m \leftarrow n} \Gamma_n \left(\lambda_n \cdot E(t) + \sum_{s' \rightsquigarrow \mathcal{P}_n} \delta u_{s'} \right) \qquad [6]$$

where $C_{m \leftarrow n}$ is connectivity matrix from population n to m and $\Gamma_n = 2r\varphi_0/(1 + e^{r(v_o - v_n^0)})^2$. In the regime of small perturbations around an equilibrium point, the effective electric coupling impact is given by the product of $\lambda_n \, C_{m \leftarrow n} \, \Gamma_n$. However, it is not obvious how the parameters involved should be changed as a function of the neural mass properties (e.g, cortical path size it represents or self-connectivity), because in NMM they are semi-empirical quantities.

Extended MPR model (NMM2)

MPR model. In their uniform, mean-field derivation for a population of N quadratic integrate and fire (QIF) neurons, Monbrió et al (14) start from the equations for the neuron membrane potential perturbation from baseline,

$$\dot{V}_j = V_j^2 + \eta_j + Js(t) + I(t), \text{ if } V_j \geq V_p, \text{ then } V_r \leftarrow V_j \qquad [7]$$

In this equation, the total input current in neuron j is $I_j = \eta_j + Js(t) + I(t)$ and includes a quenched noise constant component η_j, the input from other neurons $s(t)$ per connection received (the mean synaptic activation) with uniform coupling J and a common input $I(t)$. The common input $I(t)$ can represent both a common external input or the effect of an electric field, i.e.,

$$I(t) = p(t) + P \cdot E(t) \qquad [8]$$

Here $p(t)$ is an external uniform current, and P is the dipole conductance term in the spherical harmonic expansion of the response of the neuron to an external, uniform electric field. This is a good approximation if the neuron is in its subthreshold, linear regime (see Supplementary Information), and can be computed using realistic compartment models of the (see, e.g., (25) and (26)).

The mean synaptic activation is given by

$$s(t) = \frac{1}{N} \sum_{j=1}^{N} \sum_{k | t_j^k < t} \int_{-\infty}^{t} dt' \, a_\tau(t - t') \, \delta(t' - t_j^k) \qquad [9]$$

where t_j^k is arrival time of the kth spike from the jth neuron, and $a(t)$ the synaptic activation function, e.g., $a(t) = e^{-t/\tau}/\tau$. Note that we can write $J = j \cdot N$, where j is the synapse coupling strength (charge delivered to the neuron per action potential at the synapse) of each synapse the cell receives from the network (there are N of them in a fully connected architecture with N neurons).

We assume here for simplicity that all neurons are equally oriented with respect to the electric field. If the electric field is constant, variations in orientation can be absorbed by the quenched noise term. The total input $p(t) + P \cdot E(t) + Js(t)$ is thus homogeneous across the population (does not depend on the neuron).

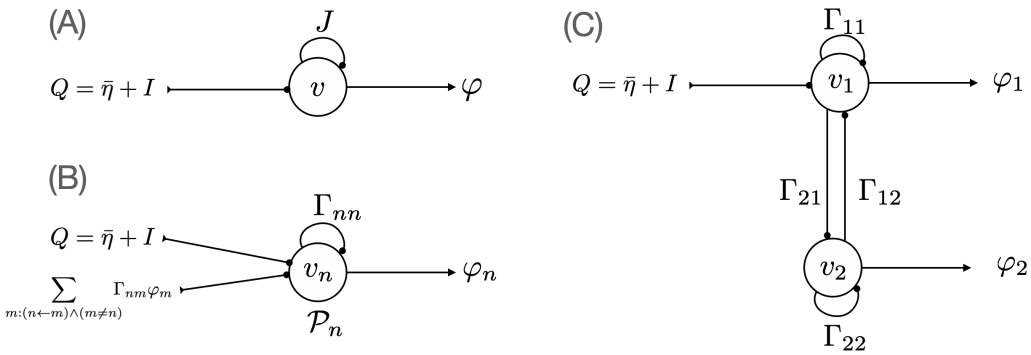


Fig. 1. NMM2 diagrams. (A): Diagram for self-coupled population with connectivity J receiving and external input Q. (B): generalization for multiple populations. (C): A generic two population model.

Starting from these, Montbrió et al derive an effective theory in the large N limit (Eq 12 in (14)),

$$\dot{\varphi} = \Delta/\pi + 2\varphi v \quad [10]$$
$$\dot{v} = v^2 + \bar{\eta} + Js + I(t) - \pi^2\varphi^2 \quad [11]$$

The new parameters $\bar{\eta}$ and Δ refer to the mean and half-width of the quenched noise distribution of η_j. The analysis in (14) hinges on the assumptions of all to all uniform connectivity (with synaptic weight J) and common input $I(t)$. In addition, a set of closed equations can be derived on the limit of instantaneous synaptic transmission ($\tau \to 0$ or $a_\tau(t-t') \to \delta(t-t')$), which implies $s \to \varphi$ to produce a closed, simple set of equations for the single population model,

$$\dot{\varphi} = \Delta/\pi + 2\varphi v \quad [12]$$
$$\dot{v} = v^2 + \bar{\eta} + J\varphi + I(t) - \pi^2\varphi^2 \quad [13]$$

This zero delay model is the one analyzed in (14). Figure 1 provides a diagram of the self-coupled population.

Adding dimensional constants. We can bring these equations closer to experiment by re-inserting dimensional constants. Our starting point becomes the QIF equation,

$$c\dot{V}_j = kV_j^2 + \eta_j + Js(t) + I(t), \text{ if } V_j \geq V_p, \text{ then } V_r \leftarrow V_j \quad [14]$$

where c is the membrane capacitance, $[c]$ =Q/V, and k has the units $[k]$= A/V^2. Here we see that $[J]$ = Coulomb, since $[s]$ = Hz, and $I(t), \eta_j$ both have current units of (A). Following the derivation of the mean-field equations in (14) now leads to

$$\dot{\varphi} = \frac{k}{c}\left[\frac{\Delta}{\pi c} + 2\varphi v\right] \quad [15]$$
$$\dot{v} = \frac{k}{c}v^2 - \pi^2\frac{c}{k}\varphi^2 + \frac{1}{c}\left[\bar{\eta} + Js(t) + I(t)\right] \quad [16]$$

with units $[r,\ s]$=Hz, $[J]$=C, $[v]$=V, $[k]$=A/V^2, $[c]$=C/V=A/(V Hz), $[\bar{\eta}, \Delta, Js(t), I]$ =A, and $[k/c]$=Hz/V.

Extended equations. Montbrió et al mention the use of an exponential delay function to relate firing rate and synaptic activity, but for simplicity set $s = \varphi$ in their final derivation. We extend this analysis here to make a connection with the neural mass modeling formalism.

When an action potential wave reaches the synapse, it is transformed into the influx of some ionic species into the cell. This transforms the shape of the action potential signal, low-pass filtering it. This process is

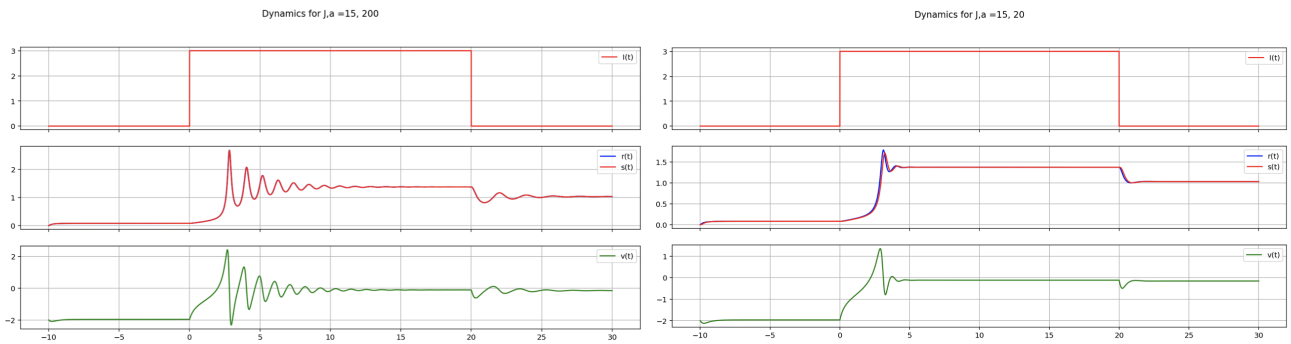


Fig. 2. Dynamics of NMM2 single population model. Left: Fast synaptic dynamics (with rate constant $a = 200$) are equivalent to the original MPR model where $s = \varphi$. Right: Slower dynamics ($a = 20$) with effect of synaptic filtering ($J = 15$, $\eta = -5$, $\Delta = 1$).

governed by the Poisson-Nernst-Planck equation (see, e.g., (9)), but can be described in a simplified manner by the dynamics of conductance $g(t)$ through the underlying ohmic relation $I_s = g(t)\,(V - V_R)$. As with the membrane perturbation, the dynamics of conductance can be modeled by a second-order linear operator (2) or its generalization to different rise and decay times (8, 10). To include synaptic dynamics we use a convolution operator as in (14), but redefine the synaptic activation function to include a Heaviside function (implementing causality), $h_\tau(t) = H(t)a_\tau(t)$. Then we can write (the convolution operator is linear)

$$\begin{aligned} s(t) &= \frac{1}{N}\sum_{j=1}^{N}\sum_{k}\int_{-\infty}^{\infty} dt'\, h_\tau(t-t')\,\delta(t'-t_j^k) \\ &= \int_{-\infty}^{\infty} dt'\, h_\tau(t-t')\left[\frac{1}{N}\sum_{j=1}^{N}\sum_{k}\delta(t'-t_j^k)\right] \\ &= \hat{L}^{-1}\left[\frac{1}{N}\sum_{j=1}^{N}\sum_{k}\delta(t'-t_j^k)\right] \equiv \hat{L}^{-1}\left[\varphi(t)\right] \end{aligned} \qquad [17]$$

with t_j^k the arrival time of the kth spike from the jth neuron, and $\varphi(t)$ the instantaneous firing rate. Thus, we have $\hat{L}[s(t)] = \varphi(t)$ to account for input gate delay. We will use the second-order operator discussed in Equation 2 to represent post-synaptic current kinetics (the generalized version with different rise and decay times can be equally used). The full set of equations for a single population becomes

$$\dot{\varphi} = \frac{k}{c}\left[\frac{\Delta}{\pi c} + 2\varphi v\right] \qquad [18]$$

$$\dot{v} = \frac{k}{c}v^2 - \pi^2\frac{c}{k}\varphi^2 + \frac{1}{c}\left[\bar{\eta} + Js + I(t)\right] \qquad [19]$$

$$\dot{s} = z \qquad [20]$$

$$\dot{z} = a^2\varphi - 2az - a^2 s \qquad [21]$$

These are readily generalized for interacting populations. Letting n denote the population index and nm denote a synapse from population m to n, the uniform input received by a population is

$$I_n(t) = p_n(t) + P_n \cdot E_n(t),$$

$$\dot{\varphi}_n = \frac{k_n}{c_n}\left[\frac{\Delta_n}{\pi c_n} + 2\varphi_n v_n\right] \quad [22]$$

$$\dot{v}_n = \frac{k_n}{c_n}v_n^2 - \frac{c_n}{k_n}\pi^2\varphi_n^2 + \frac{1}{c}\left[\bar{\eta}_n + I_n(t) + \sum_{m:n\leftarrow m}\Gamma_{nm}s_{nm}\right] \quad [23]$$

$$\dot{s}_{nm} = z_{nm} \quad [24]$$

$$\dot{z}_{nm} = a_{nm}^2\varphi_m - 2a_{nm}z_{nm} - a_{nm}^2 s_{nm} \quad [25]$$

This can be further simplified by assuming that the synapses associated to a source neuron are all equivalent, then $s_{nm} = s_m$ and $a_{nm} = a_m$. The fixed points of these equations are independent of a_{nm} and are essentially the same for the "unextended" equations

$$0 = \frac{k_n}{c_n}\left[\frac{\Delta_n}{\pi c_n} + 2\varphi_n v_n\right]$$

$$0 = \frac{k_n}{c_n}v_n^2 - \frac{c_n}{k_n}\pi^2\varphi_n^2 + \frac{1}{c}\left[\bar{\eta}_n + I_n(t) + \sum_{m:n\leftarrow m}\Gamma_{nm}\varphi_m\right] \quad [26]$$

with $z_{nm} = 0$ and $s_{nm} = \varphi_m$. The stability properties of these fixed points do depend on a_{nm}, see, e.g., Figure 4 or Figure SI-12 and others in SI.

Single population model

We consider here a single self-coupled population with constant input I and in steady state. Let Q represent the total input, $Q = I + \bar{\eta}$. Then, the fixed point equations become

$$0 = \frac{k}{c}\left[\frac{\Delta}{\pi c} + 2r\varphi\right], \quad 0 = \frac{k}{c}v^2 - \pi^2\frac{c}{k}\varphi^2 + \frac{1}{c}J\varphi + \frac{1}{c}Q \quad [27]$$

Substituting the first equation into the second we get

$$0 = \frac{k}{c}v^2 - \pi^2\frac{c}{k}(\frac{\Delta}{2\pi c v})^2 - \frac{J\Delta}{2\pi c v} + \frac{1}{c}Q$$

Canceling factors and multiplying by v^2 we end with

$$0 = kv^4 + Qv^2 - \frac{J\Delta}{2\pi}v - \frac{\Delta^2}{4k} \quad [28]$$

together with the supralinear neuron transfer function

$$\varphi = -\frac{\Delta}{2\pi c v} \quad [29]$$

which requires $v < 0$ since $\varphi \geq 0$. Equation 29 is a direct relation between voltage and firing rate and replaces the sigmoid in NMMs. Differentiating Equation 28 w.r.t. J, and using Equation 28 in the last term in the denominator, we find (since $v < 0$)

$$v'(J) = \frac{\frac{1}{2\pi}\Delta v^2}{kv^4 + Qv^2 + \frac{\Delta^2}{4k}} > 0 \quad [30]$$

since $k > 0$. This in turn implies, using Equation 29,

$$\varphi'(J) = \frac{\Delta}{2\pi c\, v^2}v'(J) > 0$$

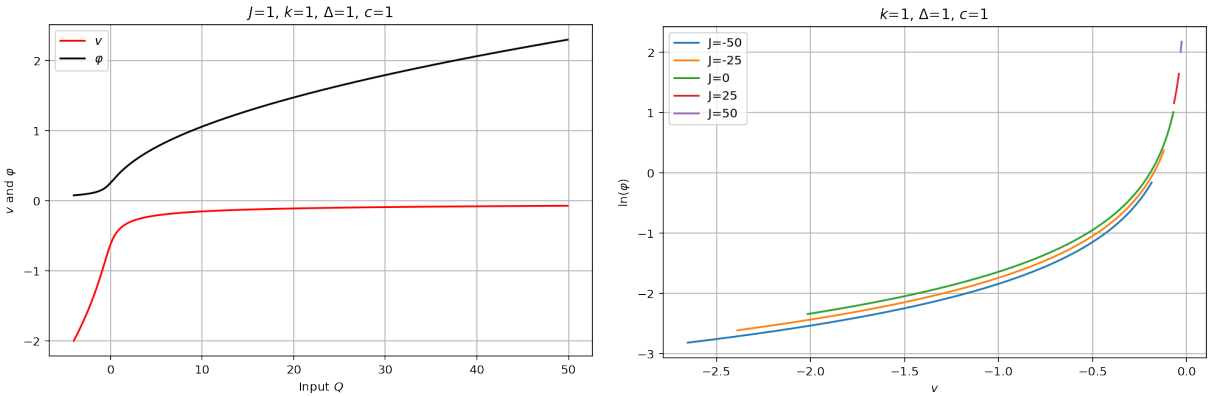


Fig. 3. Transfer function of single population model MPR/NMM2 as a function of connectivity J. Left: sample response to varying total input for a specific value of J. Observe the initial sigmoidal response shape of the firing rate φ. Right: Relation of mean firing rate and membrane potential for a set of total input values $Q \in (-4, 50)$, as a function of intrinsic connectivity J (different color traces). The underlying neuronal transfer function remains the same, but changing J moves the working point. The traces for each value of J are sequentially displaced slightly upward for clarity, but they all lie on the same line, Equation 29.

that is, membrane potential and firing rate monotonically increase with intrinsic connectivity, all other parameters being equal. Figure 3 displays the relation between firing rate and membrane potential as a function of connectivity J for a set of inputs. As can be observed, intrinsic connectivity modulates the operating point of the neuronal transfer function.

In summary, increasing connectivity leads to an increase of firing rate and membrane potential, all other things being equal. Thus, the intrinsic connectivity J acts like a (nonlinear but monotonic) gain modulator and affects the sensitivity of the population to external inputs. A J increases, so does the response of the system to an input. Translating this into the NMM framework means the sigmoid parameters such as φ_0 (Hz) in NMMs should be adjusted to reflect intrinsic connectivity. To wit, J, the signed (negative if the population is inhibitory) average number of synapses per neuron in the population, is a scaling factor of the firing rate. As a consequence of the above, the sensitivity of dynamics to input depends greatly on the self-connectivity strength of each neural mass. Figure 4 displays this relationship. The impact on dynamics of this change in operating point of the transfer function will depend, in general, on the state of the system.

Stability properties of the single population model. As saw, the fixed points of the extended model are the same as in the original MPR (they are not affected by the addition of the rate constants). However, the stability properties of fixed points are affected by the a parameter. For example, unlike with a traditional neural mass model, depending on the value of a, the single population model exhibits oscillations (a Hopf Bifurcation) when the coupling constant is negative—an Interneuron Network Gamma (ING) interaction (see Figures SI-3 and SI-4). This has also been pointed out using first order synaptic dynamics in (17).

Two population excitation-inhibition (E-I) model

Excitation-inhibition models such as pyramidal-interneuronal network gamma (PING) (23, 27–29) are often used to produce fast oscillations through a single Hopf bifurcation and are the simplest circuits in NMM capable of spontaneously oscillating under a constant input (30). They consist of an excitatory and an inhibitory population coupled together that can represent, for example, pyramidal and fast inhibitory cells in the cortex. We provide here an analog circuit in the NMM2 formalism. For the purposes of analysis, we simplify this system to have a simple constant input Q, and common parameters a, Δ, and we neglect

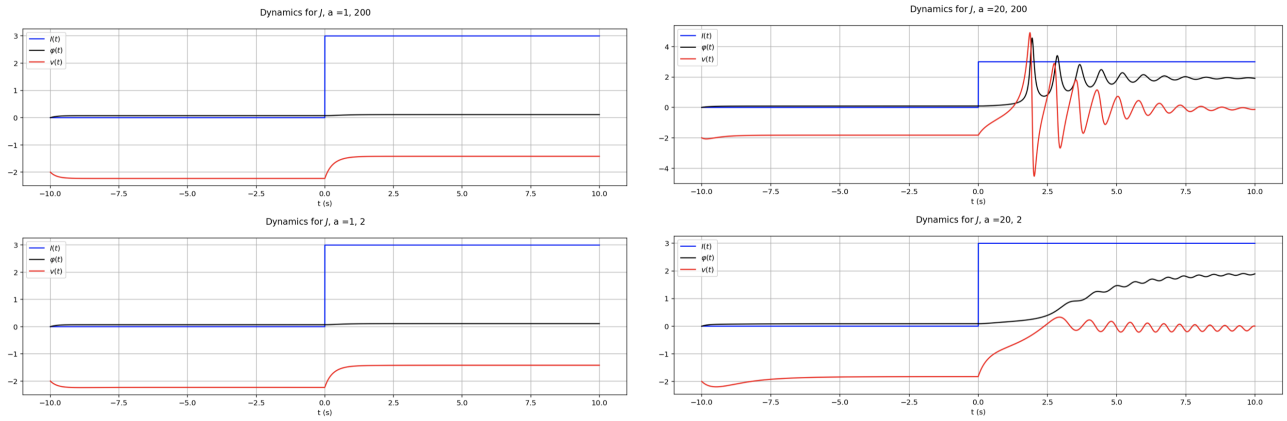


Fig. 4. Sensitivity to input depends on J and rate constant in single population NMM2 model. Left: response to input with $J = 1$ for fast (top) and slow (bottom) synaptic dynamics. Right: response with stronger self-connectivity $J = 20$ for fast (top) and slow (bottom) synaptic dynamics. In both cases $\eta = -5$, $\Delta = 1$. See Figures SI-5 and SI-6 for a more detailed analysis of changes in dynamics as a function of J.

self-coupling terms ($\Gamma_{nn} = 0$, see Figure 1). The equations for this system are

$$\begin{aligned}
\dot{\varphi}_1 &= \Delta/\pi + 2\varphi_1 v_1 \\
\dot{v}_1 &= v_1^2 + s_{12} + Q - \pi^2\varphi_1^2 \\
\dot{\varphi}_2 &= \Delta/\pi + 2\varphi_2 v_2 \\
\dot{v}_2 &= v_2^2 + s_{21} - \pi^2\varphi_2^2 \\
\dot{s}_{12} &= z_{12} \\
\dot{z}_{12} &= \Gamma_{12} a^2 \varphi_2 - 2a z_{12} - a^2 s_{12} \\
\dot{s}_{21} &= z_{21} \\
\dot{z}_{21} &= \Gamma_{21} a^2 \varphi_1 - 2a z_{21} - a^2 s_{21}
\end{aligned} \qquad [31]$$

This system of equations exhibits rich dynamics. Figure SI-7 provides some sample dynamics, which now include bursting. Figure SI-12 displays simple bifurcation diagrams (starting from the same initial conditions (0,-2,0,0) and eliminating a transient of 10 seconds) to study the effect of changing the rate parameter a. The tails of these diagrams are probably noise from long transients. Figure SI-13 displays the increased sensitivity of the model to input when the coupling is increased while maintaining the E/I balance (here the ratio of the Γ_{21}/Γ_{12} constants).

Discussion

The extended, unified model NMM2 brings together descriptions of synaptic transmission and neuron function based on mechanistic principles. This exact mean-field theory provides a physiologically robust description with parameters that can be related to physiological features. NMM2 features a supra-linear transfer function, exhibiting an increasing gain with growing inputs (and equivalently output firing rate). This means that effective coupling increases with increasing network activation, and as we saw, also with increasing self-coupling.

With regard to the self-coupling parameter J, note that it has units of charge (see Equation 15). This reflects the fact that J encodes the effective connectivity of the average neuron to the rest of the network, i.e., the total number of connections/synapses the cell receives times the charge deposited by each during an action potential.

We have seen that in NMM2 the intrinsic population connectivity plays a key role in modulating the response of the system to its inputs. Translating this into the context of NMMs, this means that the lumped sigmoid parameters (φ_0, v_0 and r) should be adjusted to reflect the self-connectivity of the cortical patch the mass model is representing.

The analysis of the MPR framework also clarifies that the electrical coupling constant derived from neuron compartment models (26) can be directly used as is in neural mass models: the effective response to weak electric field perturbations can be modulated by intrinsic connectivity J, or if one wishes to work with NMMs, with the modified sigmoid function parameters.

Such modulation of the effective coupling provides a long-sought mechanism for the surprising sensitivity of networks to weak electric fields generated by brain stimulation or by neural tissue itself, a phenomenon known as ephaptic interaction (31). It may also explain why stronger electrics fields are needed to induce measurable effects in in-vitro (low, disrupted connectivity) or small animal studies as compared with humans.

A tightly coupled network in the right regime can display more sensitivity to weak external inputs (see the Supplementary Information section for examples), although, in general, this will depend on the state of the network and factors such as the type of E-I balance. With regard to this, current evidence favors *loose balance* in the sensory cortex (defined to be when net input remaining after cancellation of excitation and inhibition is comparable in size with the factors that cancel), for example, but in the motor or frontal cortex the case is less clear, with *tight balance* also having been proposed (when the net input is very small relative to the canceling factors) (32). Presumably, the weak perturbation effects of tES would be more important in a tight balance regime.

In transcranial electric stimulation (tES), the electric field generated on the cortex is of the order of 1 V/m (33), which is known to produce a sub-mV membrane perturbation (25, 34). Yet, it is of mesoscopic nature, with a spatial scale of several centimeters. The main characteristic of the exogenous macroscopic electric field generated by tES is its small magnitude, low temporal frequency, moderate spatial correlation scales (> 1 cm), and a long application time (typically twenty minutes to one hour) and repeated sessions when long-term plastic effects are desired. Endogenous fields are similarly characterized by weak field magnitudes over mesoscopic scales (31). As we have seen in the analysis of an exact mean-field theory, weak but spatially uniform fields can affect the dynamics of densely connected neuron populations exposed.

The NMM2 formalism introduces new dynamical features, even in simple models. The MPR model, which can be seen as an NMM2 in the limit of very high rate constant (fast synapses), already displays chaos (14). We have seen here that NMM2 can produce rich dynamics, with features such as oscillations in a single population with constant input, as well as bursting in a simple two population model. We leave the detailed analysis of these and other systems in the NMM2 framework for further work.

Conclusions

Based on a recently derived exact mean-field theory and decades of work on neural mass modeling, we provided an extension of first that brings together the features in each with a solid mechanistic basis. The resulting model is limited, as it is based on a simple neuron population model (QIF with uniform self-connectivity), but is an exact mean-field theory endowing with unambiguous meaning each of its parameters and variables. The analysis of the resulting system and comparison with classical NMMs sheds light on network phenomena mediated by self-connectivity and the physiological basis of semi-empirical parameters in NMMs.

ACKNOWLEDGMENTS. This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 855109) and from FET under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101017716).

1. HR Wilson, JD Cowan, Excitatory and Inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**, 1–24 (1972).
2. F Lopes da Silva, A Hoek, H Smits, LKn Zetterberg, Model of brain rhythmic activity: the alpha rhythm of the thalamus. *Kybernetik* **15**, 27–37 (1974).
3. F Lopes da Silva, vR A, P Barts, E Heusden, W van Burr, Model of neuronal populations: the basic mechanism of rhythmicity. *Prog Brain Res* **45** (1976).
4. BH Jansen, G Zouridakis, ME Brandt, A neurophysiologically-based mathematical model of flash visual evoked potentials. *Biol Cybern* **68**, 275–83 (1993).
5. BH Jansen, VG Rit, Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol Cybern* **73**, 357–66 (1995).
6. R Sanchez-Todo, et al., Personalization of hybrid brain models from neuroimaging and electrophysiology data. *BioRxiv* **00**, 1–35 (2018).
7. G Ruffini, F Wendling, R Sanchez-Todo, E Santarnecchi, Targeting brain networks with multichannel transcranial current stimulation (tcs). *Curr. Opin. Biomed. Eng.* (2018).
8. A Destexhe, ZF Mainen, TJ Sejnowski, *Kinetic models of synaptic transmission*, eds. C Koch, I Segev. (MIT Press, Cambridge, MA.), 2nd edition, pp. 1–25 (1998).
9. J Pods, J Schönke, P Bastian, Electrodiffusion models of neurons and extracellular space using the poisson-nernst-planck equations—numerical simulation of the intra- and extracellular potential for an axon model. *Biophys. J.* **105**, 242–254 (2013).

10. G Ermentrout, DH Bard, Terman, *Mathematical Foundations of Neuroscience*. (Springer-Verlag New York), (2010).
11. WJ Freeman, *Mass Action in the Nervous System*. (New York: Academic Press), (1975).
12. LM Kay, The physiological foresight in Freeman's work. *J Conscious Stud.* **25**, 50–63 (2018).
13. FH Eeckman, FW J, Asymmetric sigmoid non-linearity in the rat olfactory system. *Brain Res.* **557**, 13–21 (1991).
14. E Montbrió, D Pazó, A Roxin, Macroscopic description for networks of spiking neurons. *Phys. Rev. X* **021028** (2015).
15. E Montbrió, D Pazó, Exact mean-field theory explains the dual role of electrical synapses in collective synchronization. *Phys. Rev. Lett.* **125** (2020).
16. F Devalle, A Roxin, E Montbrió, Firing rate equations require a spike synchrony mechanism to correctly describe fast oscillations in inhibitory networks. *PLOS Comput. Biol.* **13** (2017).
17. G Dumont, B Gutkin, Macroscopic phase resetting-curves determine oscillatory coherence and signal transfer in inter-coupled neural circuits. *PLoS Comput. Biol.* **15** (2019).
18. F Grimbert, O Faugeras, Analysis of Jansen's model of a single cortical column. *INRIA* **RR-5597**, 34 (2006).
19. WJ Freeman, Neural modeling, eds. MAB Brazier, DO Walter, D Schneider. (U. of California, Los Angeles), pp. 41–62 (1973).
20. F Wendling, F Bartolomei, JJ Bellanger, P Chauvel, Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *Eur J Neurosci* **15**, 1499–508 (2002).
21. F Wendling, P Chauvel, Transition to Ictal Activity in Temporal Lobe Epilepsy: Insights From Macroscopic Models. *Comput. Neurosci. Epilepsy* pp. 356–386 (2008).
22. LS et al., Linking molecular pathways and large-scale computational modeling to assess candidate disease mechanisms and pharmacodynamics in alzheimer's disease. *Front. Comput. Neurosci.* (2019).
23. B Molaee-Ardekani, P Benquet, F Bartolomei, F Wendling, Computational modeling of high-frequency oscillations at the onset of neocortical partial seizures: from 'altered structure' to 'dysfunction'. *Neuroimage* **52**, 1109–22 (2010).
24. I Merlet, et al., From Oscillatory Transcranial Current Stimulation to Scalp EEG Changes: A Biophysical and Physiological Modeling Study. *PLoS ONE* **8**, 1–12 (2013).
25. AS Aberra, AV Peterchev, WM Grill, Biophysically realistic neuron models for simulation of cortical stimulation. *J. Neural Eng.* **15**, 066023 (2018).
26. A Galan, Master's thesis (Elite Master Program in Neuroengineering, Department of Electrical and Computer Engineering, Technical University of Munich) (2021).
27. MA Whittington, RD Traub, N Kopell, B Ermentrout, EH Buhl, Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int J Psychophysiol* **38**, 315–36 (2000).
28. C Borgers, S Epstein, NJ Kopell, Gamma oscillations mediate stimulus competition and attentional selection in a cortical network model. *Proc. Natl. Acad. Sci.* **105**, 18023–18028 (2008).
29. R Sanchez-Todo, et al, A physical neural mass modeling framework forlaminar cortical measurements. *PNAS (in preparation)* (2021).
30. M Guasch, G Ruffini, Analysis of dynamics of elementary neural mass model circuits. *PLOS Comput. Biol. (in preparation)* (2021).
31. G Ruffini, et al., Realistic modeling of mesoscopic ephaptic coupling in the human brain. *PLoS Comput. Biol.* (2020).
32. Y Ahmadian, KD Miller, What is the dynamical regime of cerebral cortex? *Neuron* **109** (2021).
33. G Ruffini, et al., Transcranial Current Brain Stimulation (tCS): Models and Technologies. *IEEE Transactions on Neural Syst. Rehabil. Eng.* **21**, 333–345 (2013).
34. M Bikson, et al., Effects of uniform extracellular DC electric fields on excitability in rat hippocampal slices in vitro. *J Physiol* **557**, 175–90 (2004).
35. G Ruffini, MD Fox, O Ripolles, PC Miranda, A Pascual-Leone, Optimization of multifocal transcranial current stimulation for weighted cortical pattern targeting from realistic modeling of electric fields. *Neuroimage* **89**, 216–225 (2014).

Supplementary Information

Dipole nature of electric field interaction in compartment model of a neuron in the linear regime. Here we show that in the linear regime (far from threshold) the effect of an electric field on a neuron compartment, or its average over a set of compartments, can be written in the form $\delta V = \lambda \cdot E$, where λ is a vector. This is the so-called "lambda-E" model used in transcranial electrical stimulation models (33, 35).

The effect of an external field can be represented in a neuron compartment model by an axial current that results from the potential difference induced by the field along the fiber associated with a compartment (see, e.g., (26) and references therein),

$$I = g^a \int_x^{x+\Delta x} E_x(s)\, ds \qquad \text{[SI-1]}$$

where E_x is the field along the fiber. We can rewrite this as $I = g^a \int_f \mathbf{E} \cdot d\mathbf{l}$ with the line integral along the fiber compartment. If the field is constant along the fiber we can express this simply as $I = \mathbf{E} \cdot \mathbf{u}$. The vector $\mathbf{u}$ is a vector parallel to the line that ends in the compartment of interest and originates in the connected one(s), and thus pointing into the compartment (see Figure SI-1). If the E field is aligned with this direction, we get a positive current into the compartment. If the compartment of interest is k and the connected compartment is j, we refer to the vector $\mathbf{u}_{k,j}$ parallel to $\mathbf{x}_k - \mathbf{x}_j$ (the coordinates of the compartments). The superscripts r and a refer to radial and axial components and conductivities.

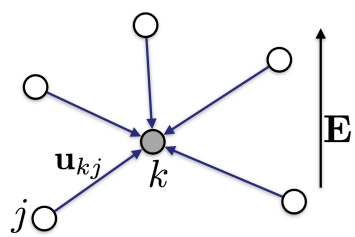


Fig. SI-1. Compartments and vectors for cable equation in E field. Here we represent a fiber by a node and and a set of directed edges.

From conservation of charge (the sum over j is over connected compartments, and i over membrane currents), and assuming that the field is uniform and thus independent of the compartment k, we obtain

$$C_k \frac{dV_k}{dt} + \sum_i I_{i,k} = \sum_{j \leftrightsquigarrow k} g_j^a \left(V_j - V_k\right) + \sum_{j \leftrightsquigarrow k} g_j^a\, \mathbf{E} \cdot \mathbf{u}_{k,j} \qquad \text{[SI-2]}$$

The sum $\sum_{j \leftrightsquigarrow k} Q_j \equiv \sum_j c_{kj} Q_j$ denotes the sum over compartments connected to the kth one, and can be expressed via the compartment connectivity matrix c_{kj}, with value of 1 if compartment j is connected to k, zero otherwise. The sum on the left hand side is over ionic membrane currents out of the compartment.

We can simplify this and also pull the constant electric field out the sum,

$$C_k \frac{dV_k}{dt} + I_k = \sum_{j \leftrightsquigarrow k} g_k^a (V_j - V_k) + \mathbf{p}_k \cdot \mathbf{E} \qquad \text{[SI-3]}$$

where $I_k = \sum_i I_{i,k}$, the k sum is over connected components to the kth compartment and where

$$\mathbf{p}_k = \sum_{j \leftrightsquigarrow k} g_j^a \mathbf{u}_{k,j} = \sum_j c_{kj} g_k^a \mathbf{u}_{k,j} \qquad \text{[SI-4]}$$

The equation for a single compartment in steady state is thus of the form

$$C_k \frac{dV_k}{dt} - \mathbf{p}_k \cdot \mathbf{E} + g_k^r V_k + I_k^0 + \sum_{j \leftrightsquigarrow k} g_j^a (V_k - V_j) = 0 \qquad \text{[SI-5]}$$

where $\mathbf{p}_k$ is the dipole vector corresponding to the kth compartment, and $I_k = V_k g_k^r + I_k^0$, with I^0 the sum of currents associated to the reversal potentials, and g_k^r the axial (transmembrane) conductivity.

Thus, we see that the behavior of each compartment is characterized by its own dipole and the interaction with other compartments.

Equation SI-5 is linear and can be expressed in matrix form, with the compartment array notation $V = (V_k) = (V_1, V_2, \ldots, V_N)$, $C = \mathrm{diag}(C_k) = \mathrm{diag}(C_1, C_2, \ldots, C_N)$, and $(\mathbf{p}_k \cdot \mathbf{E}) = (\mathbf{p}_1 \cdot \mathbf{E}, \ldots, \mathbf{p}_N \cdot \mathbf{E})$,

$$C\frac{\partial V}{\partial t} + AV = I^0 - (\mathbf{p}_k \cdot \mathbf{E}) \qquad \text{[SI-6]}$$

and with the matrix A given by

$$A_{kj} = \delta_{kj} g_j^r - g_j^a c_{k,j} + \delta_{kj} \sum_l g_l^a c_{l,j} \qquad \text{[SI-7]}$$

In steady state ($\dot{V} = 0$), the solution to this equation is

$$V = A^{-1}[I^0 - (\mathbf{p}_k \cdot \mathbf{E})] \qquad \text{[SI-8]}$$

or

$$V_k = \sum_j A_{kj}^{-1}[I_k^0 - \mathbf{p}_k \cdot \mathbf{E}] = \sum_j A_{kj}^{-1} - \mathbf{E} \cdot \sum_j A_{kj}^{-1} \mathbf{p}_j I_j^0 \qquad \text{[SI-9]}$$

where the second equality is a result of linearity of the A matrix, i.e., the fact that we are dealing with a linear equation. Finally, we can write the solution as

$$V_k = \sum_j A_{kj}^{-1} I_j^0 + P_k \cdot E \qquad \text{[SI-10]}$$

This equation shows that the response of the compartment to an electric field is of dipole form, with the total dipole P_k resulting from a superposition of dipoles,

$$\mathbf{P}_k = -\sum_j A_{kj}^{-1} \mathbf{p}_j = -\sum_{j,l} A_{kj}^{-1}\, c_{jl} g_j^a \mathbf{u}_{j,l} \qquad \text{[SI-11]}$$

Furthermore, we observe that any linear function of compartment potentials will also display, by further superposition, a dipole response. This includes, for example, the average membrane perturbation for apical dentrite compartments, or soma.

What happens in the **nonlinear case**? We can write

$$\mathbf{p}_k \cdot \mathbf{E} + I_k^r(V_k) + \sum_{j \leftrightsquigarrow k} g_j^a (V_k - V_j) = 0 \qquad \text{[SI-12]}$$

with $I_k^r(V_k)$ a nonlinear function, or, in a fashion analogous to the discussion above, in matrix form

$$\mathbf{p}_k \cdot \mathbf{E} + BV + I(V) = 0$$

or

$$F_k(V) = \mathbf{p}_k \cdot \mathbf{E}$$

for a nonlinear operator F. Unfortunately, this means we cannot pull out the E field out of the inverse, so the situation does not simplify.

Stability analysis for single population NMM2 model. We analyze here the fixed points and their stability of the NMM2 single population model, showing it displays a Hopf bifurcation for some parameters choices.

We start from the equations

$$
\begin{aligned}
\dot{\varphi} &= f_\varphi = \frac{k}{c}\left[\frac{\Delta}{\pi c} + 2\varphi v\right] \\
\dot{v} &= f_v = \frac{k}{c}v^2 - \pi^2\frac{c}{k}\varphi^2 + \frac{1}{c}\left[\bar{\eta} + Js + I(t)\right] \\
\dot{s} &= f_s = z \\
\dot{z} &= f_z = a^2\varphi - 2az - a^2 s
\end{aligned}
$$

The Jacobian is

$$
J = \begin{pmatrix} f_{\varphi,\varphi} & f_{\varphi,v} & f_{\varphi,s} & f_{\varphi,z} \\ f_{v,\varphi} & f_{v,v} & f_{v,s} & f_{v,z} \\ f_{s,\varphi} & f_{s,v} & f_{s,s} & f_{s,z} \\ f_{z,\varphi} & f_{z,v} & f_{z,s} & f_{z,z} \end{pmatrix} = \begin{pmatrix} \frac{2k}{c}v & \frac{2k}{c}\varphi & 0 & 0 \\ -\frac{2\pi^2 c}{k}\varphi & \frac{2k}{c}v & J/c & 0 \\ 0 & 0 & 0 & 1 \\ a^2 & 0 & -a^2 & -2a \end{pmatrix}
$$

To assess stability of a fixed point we compute the eigenvalues of the Jacobian at that point, and if any points have a positive real component they are marked as unstable.

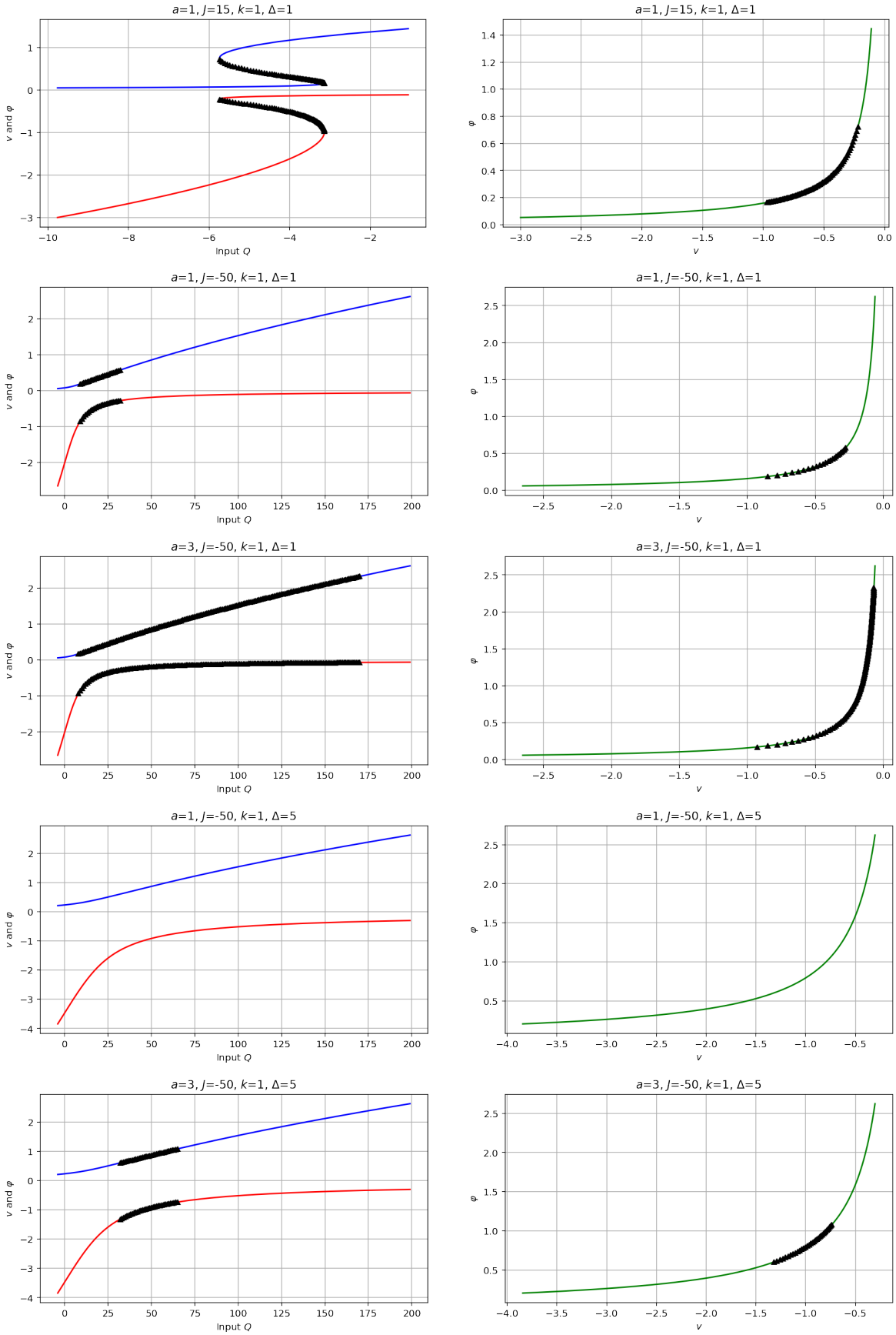


Fig. SI-2. Fixed points and their stability for different parameters. The first row corresponds to $J = 15$ as described in (14). The following ones are for the case $J = -50$. Triangles indicate unstable points that appear through a Hopf bifurcation in the negative self-coupling case.

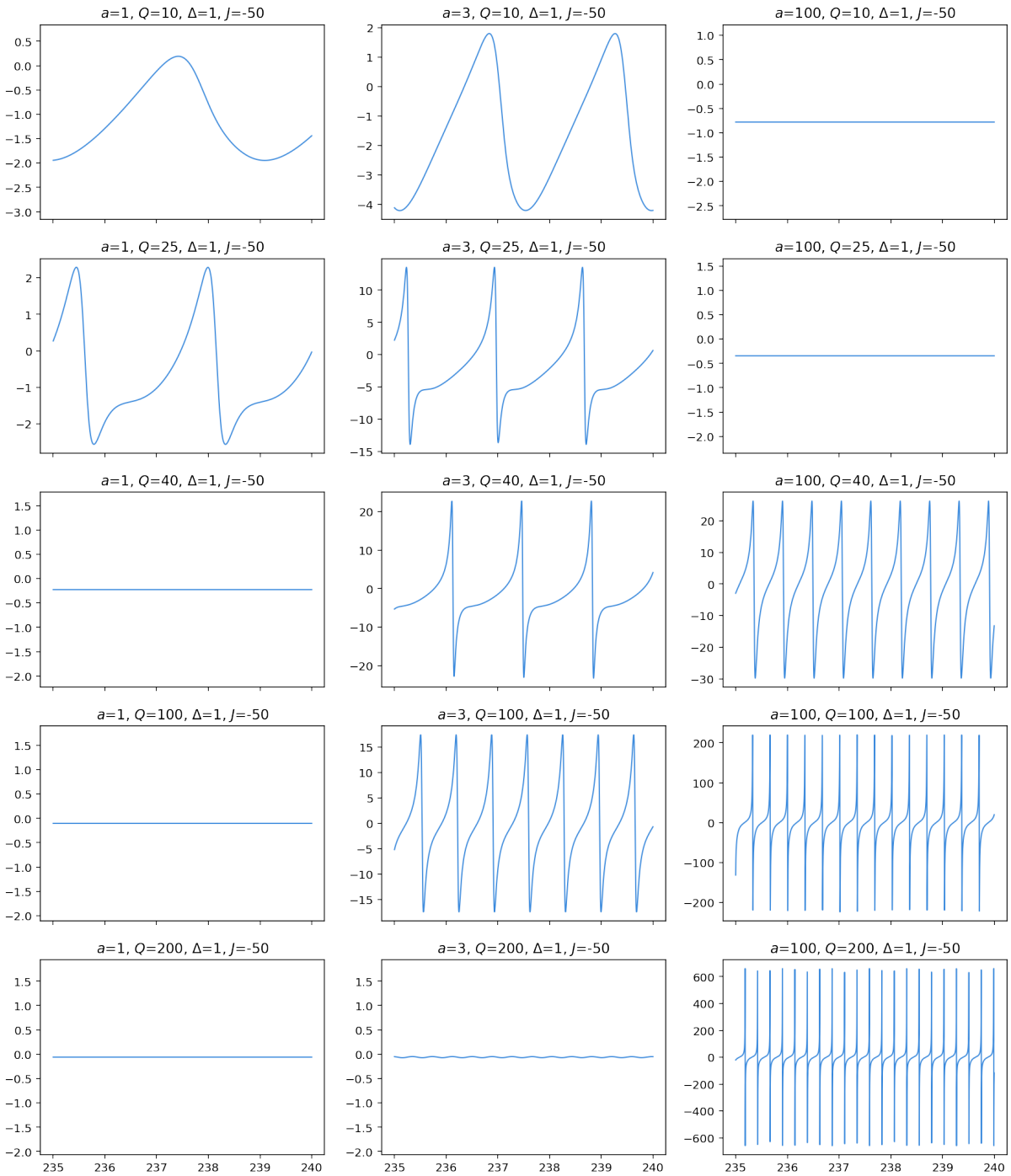


Fig. SI-3. **Sample dynamics for self-inhibitory single population system with $J = -50$ and $\Delta = 1$.**

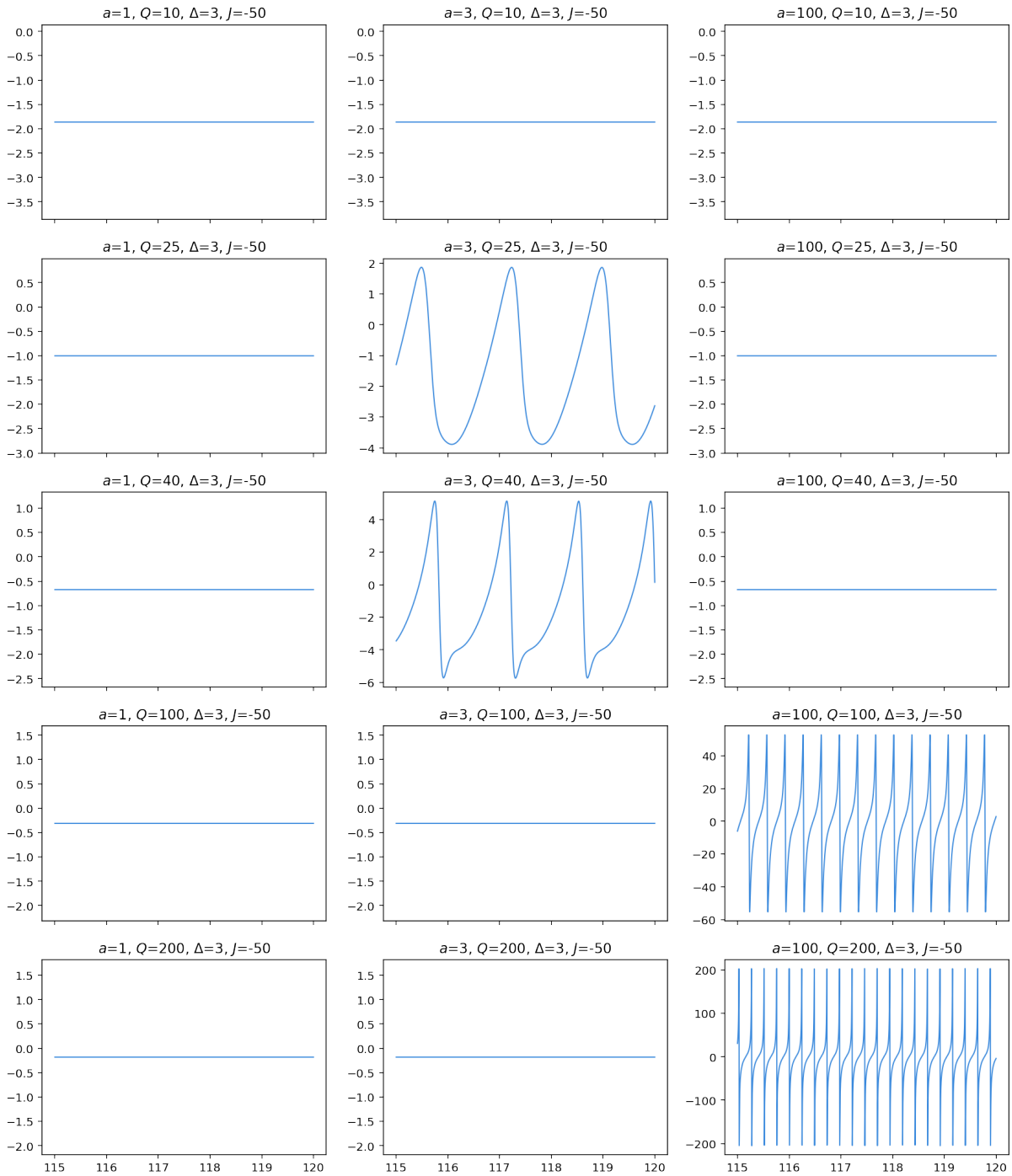


Fig. SI-4. **Sample dynamics for self-inhibitory single population system with $J = -50$ and $\Delta = 5$.**

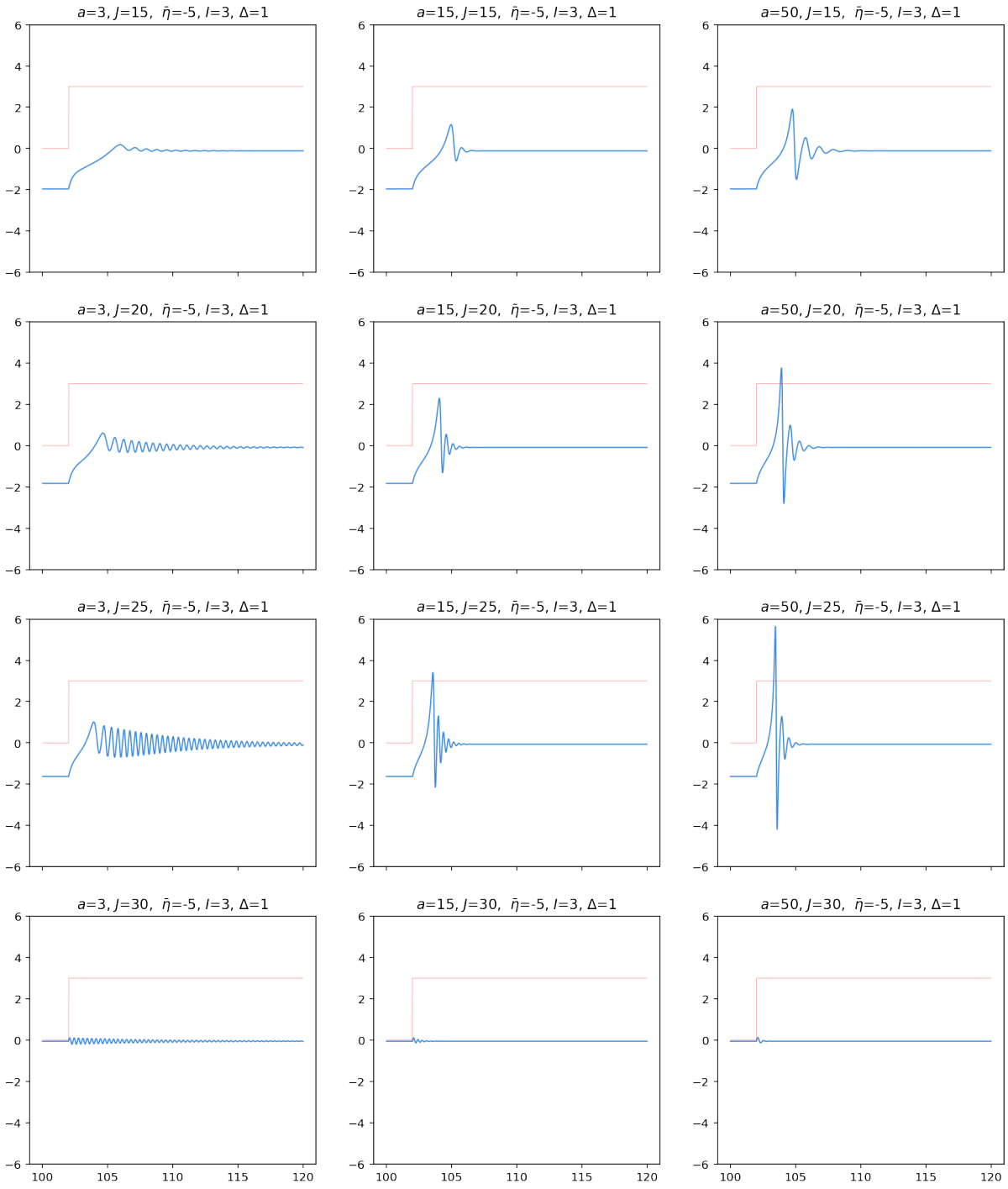


Fig. SI-5. Response of excitatory ($J > 0$) single population to step input function (red) as a function of self-coupling. Here the total input is $Q = \bar{\eta} + H(t - t_0)I$.

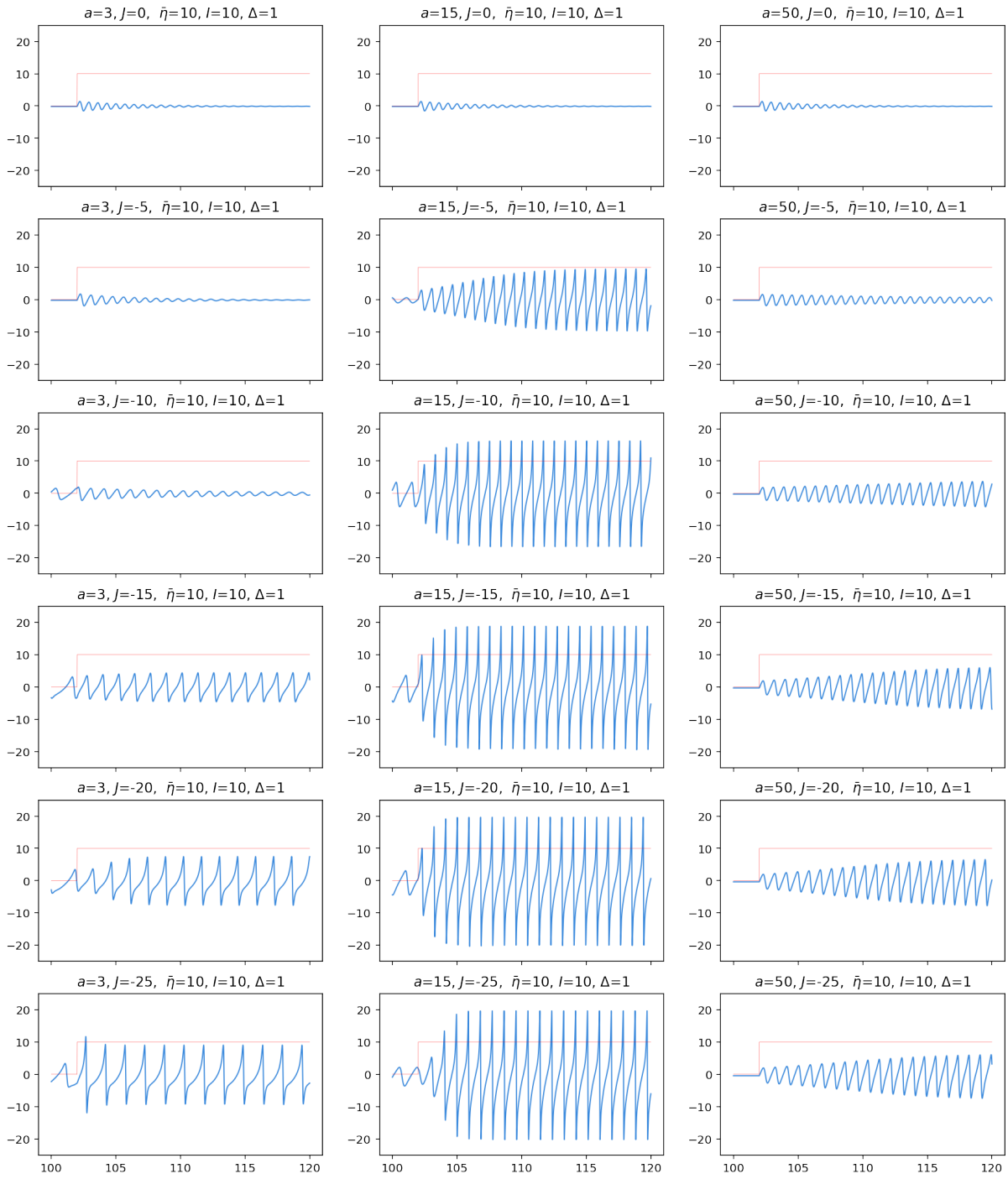


Fig. SI-6. Response of inhibitory ($J < 0$) single population to step input function (red) as a function of self-coupling. Here the total input is $Q = \bar{\eta} + H(t - t_0)I$.

Sample dynamics for simplified NMM2 E-I model.

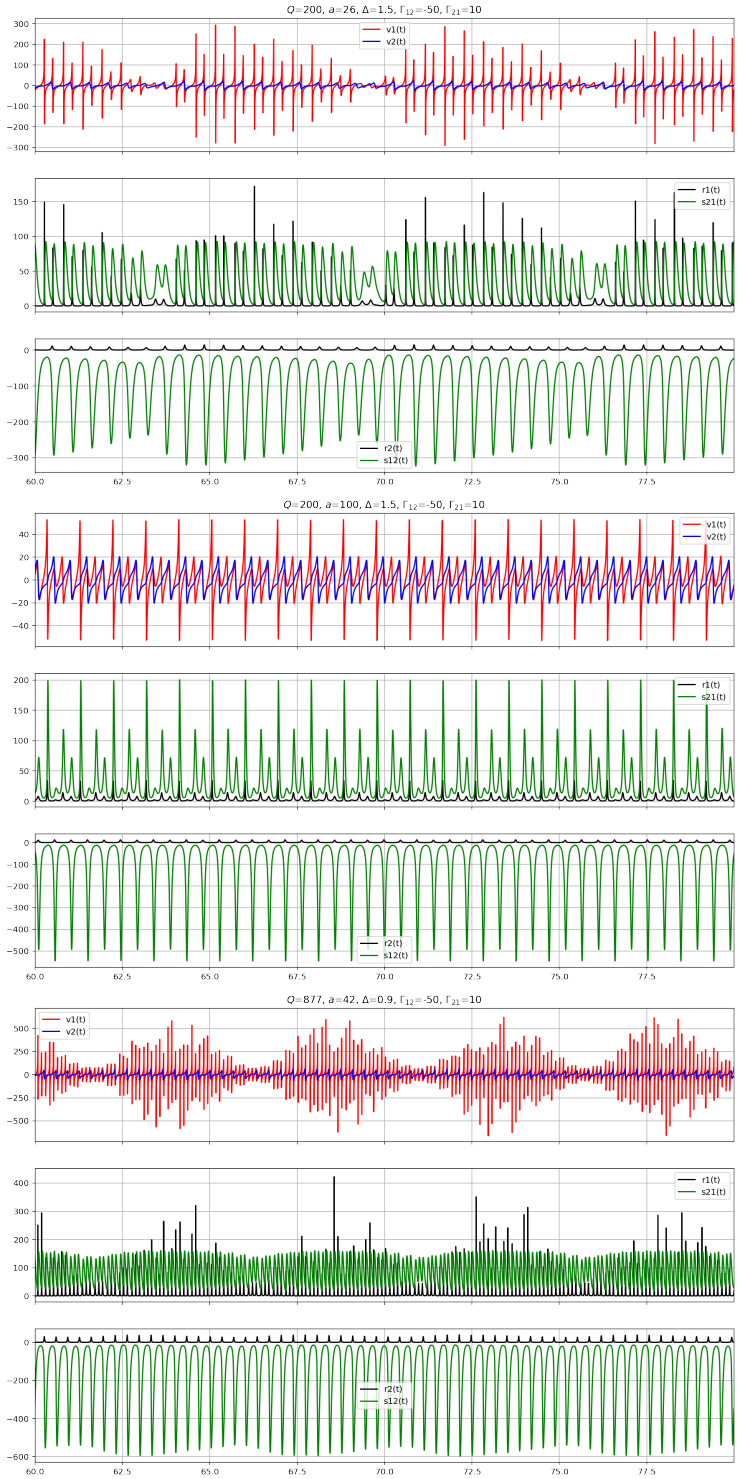


Fig. SI-7. Sample dynamics for the NMM2 two population E-I model.

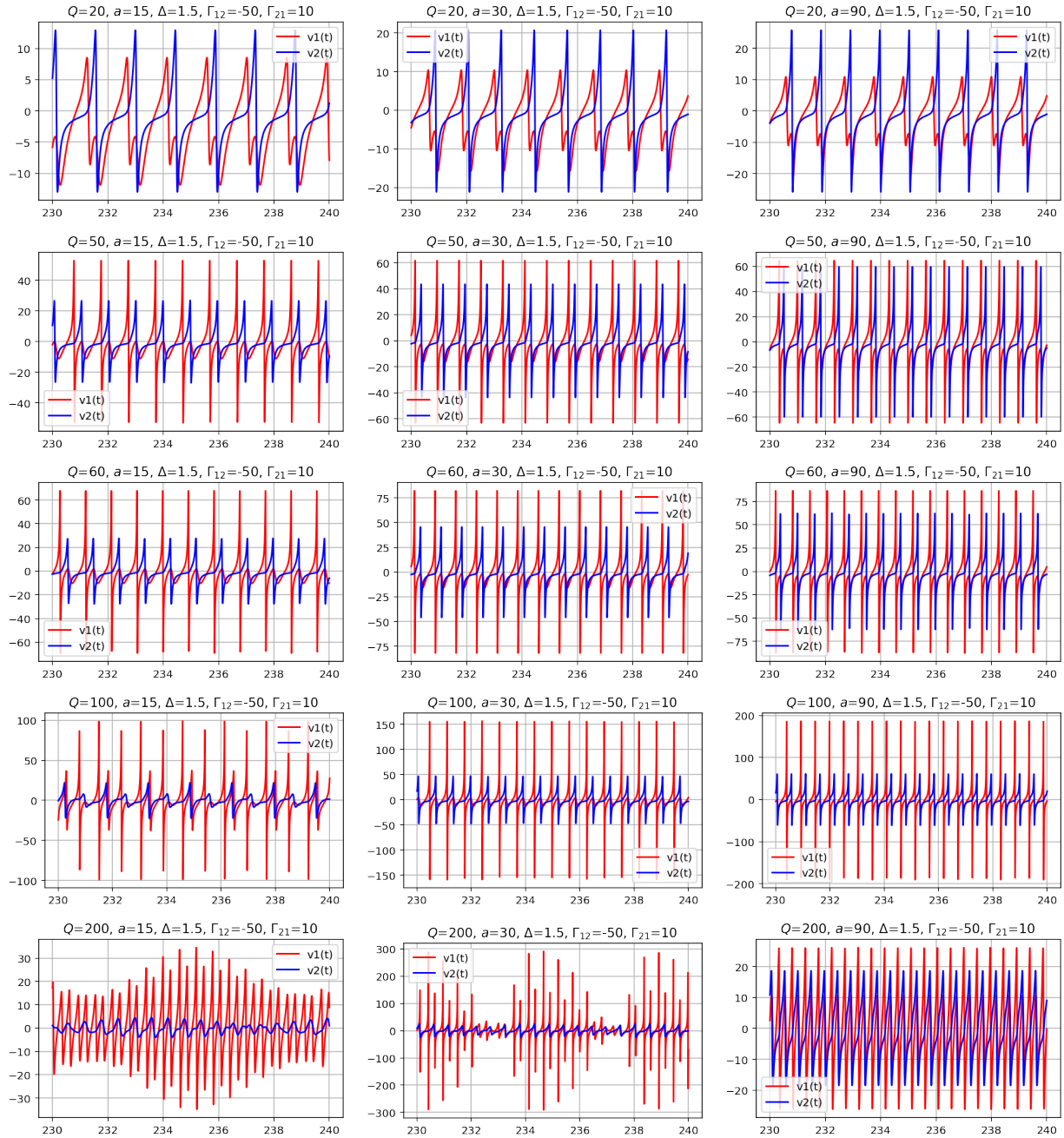


Fig. SI-8. Sample dynamics for the NMM2 two population E-I model.

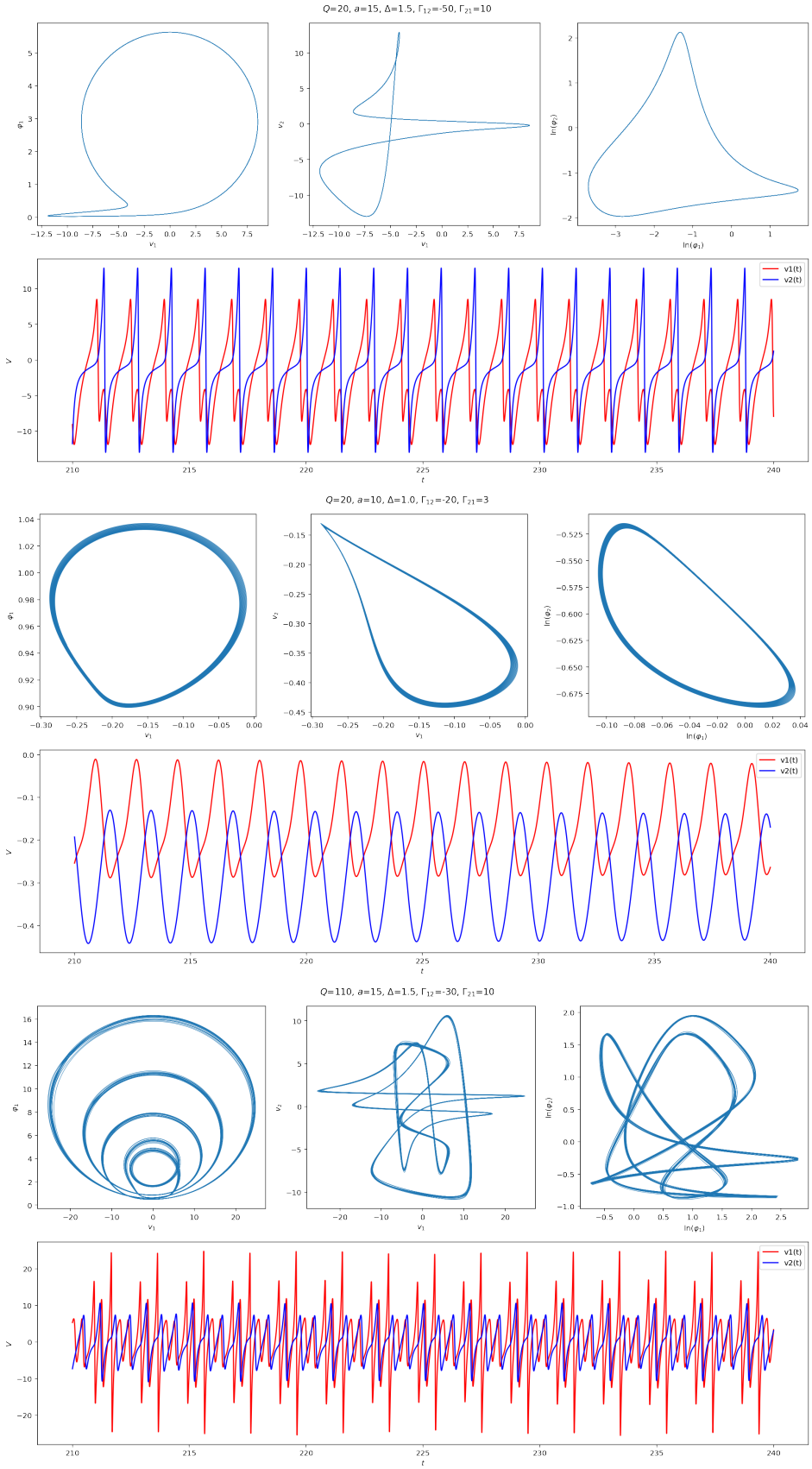


Fig. SI-9. Interesting dynamics for the NMM2 two population E-I model (I).

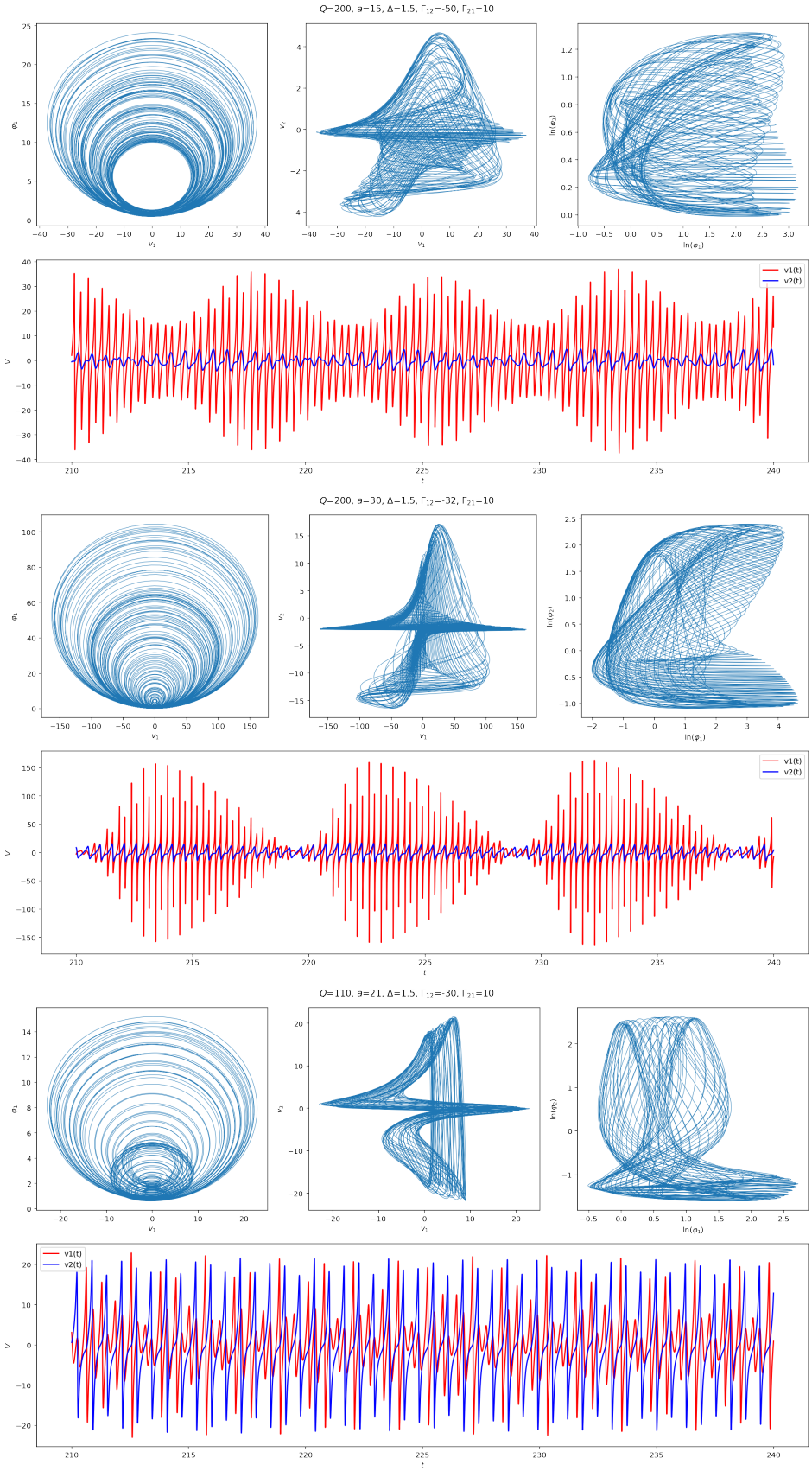
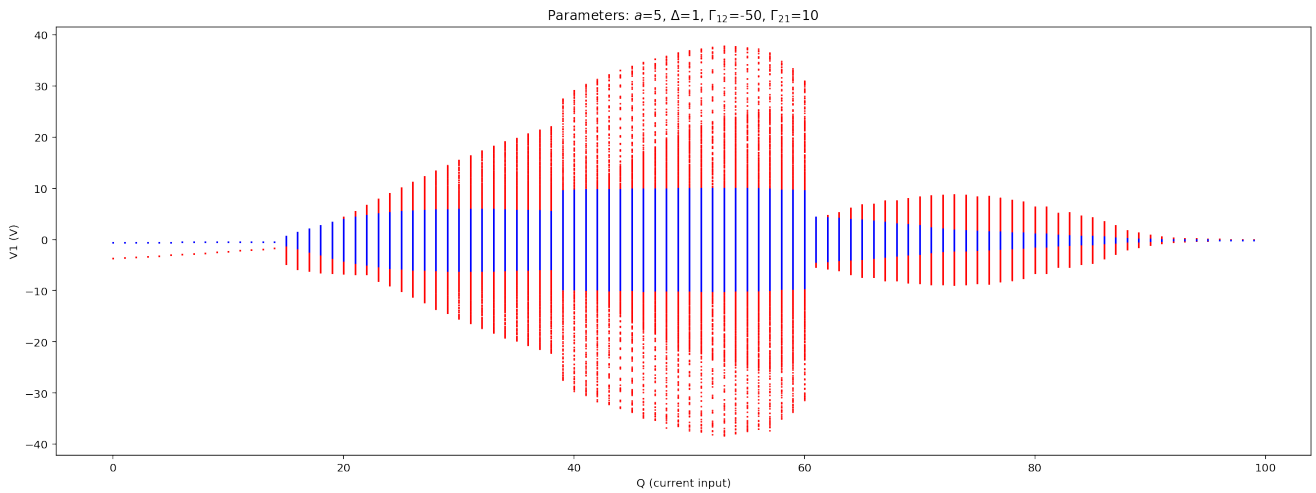


Fig. SI-10. Interesting dynamics for the NMM2 two population E-I model (II).

Bifurcation diagrams for simplified NMM2 E-I model. Here we provide some bifurcation diagrams for the simplified E-I model, by simply plotting the dynamics for each step of the bifurcation parameter (input current).



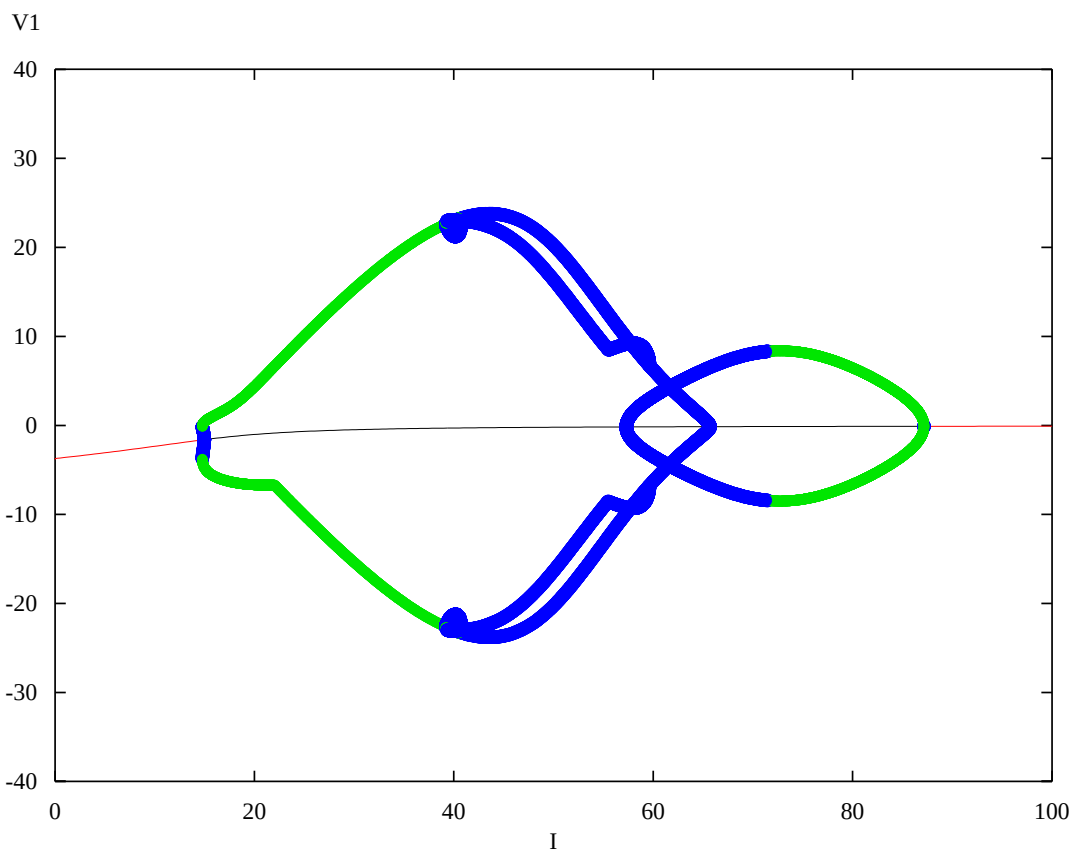


Fig. SI-11. Bifurcation analysis of NMM2 E-I model for $a = 5$ with XPP-AUTO (bottom, see 07.5 MPR bifurcation.ipynb). The diagram exhibits multiple bifurcation points, including Hopf and period doubling points.

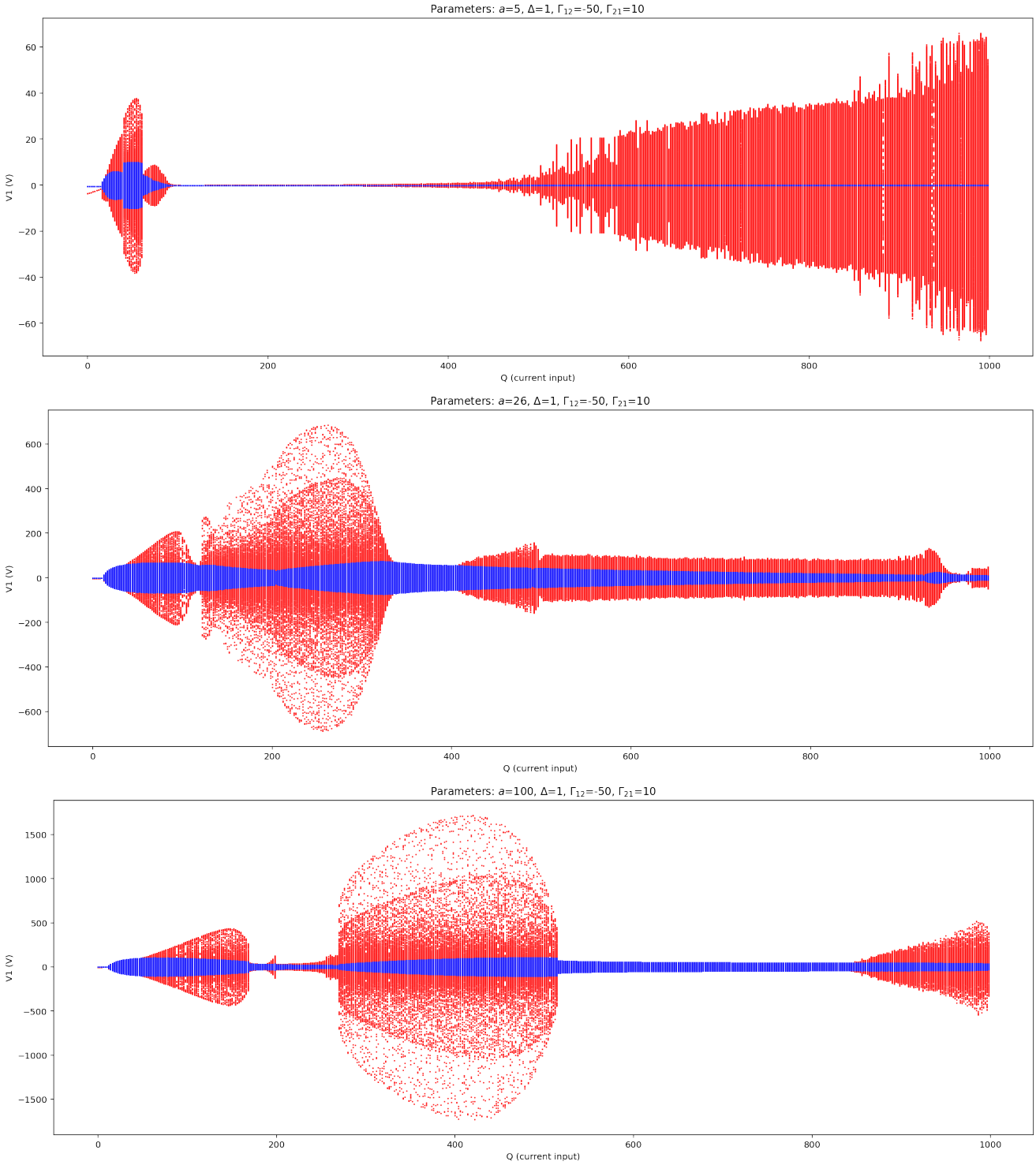


Fig. SI-12. Bifurcation analysis of NMM2 E-I model for different values of a (5,25,100).

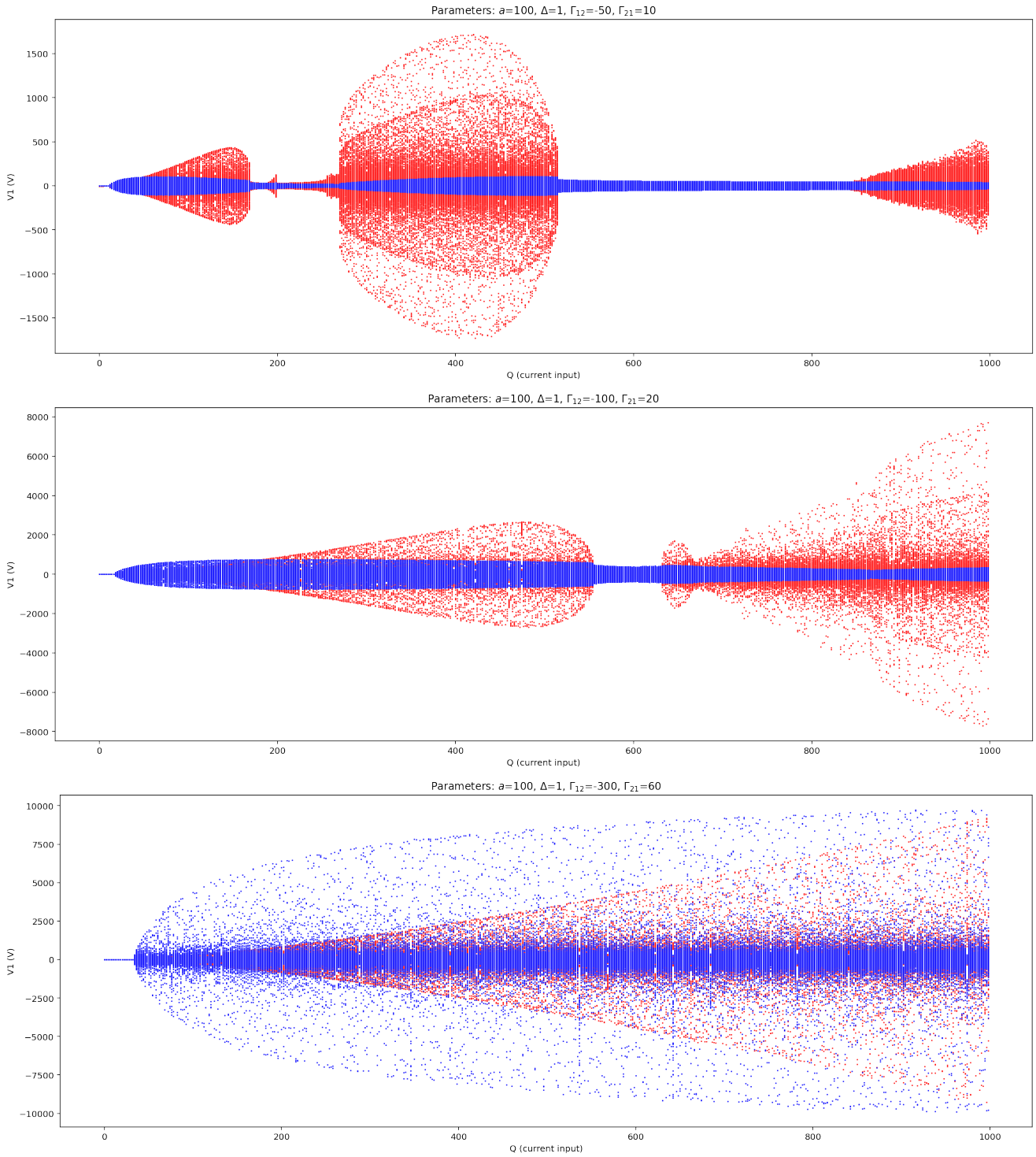


Fig. SI-13. Bifurcation analysis of NMM2 E-I model for different values of coupling gain (Γ increasing from top to bottom, with same E/I ratio). Note the increase in voltage as the gain scaling is increased.

More general E-I model. Here we retain more features of the full two population model. The system of equations becomes — first the firing rate equations,

$$\dot{\varphi}_1 = \Delta_1/\pi + 2\varphi_1 v_1 \quad [32]$$
$$\dot{v}_1 = v_1^2 + \Gamma_{12} s_2 + \Gamma_{11} s_1 + \bar{\eta}_1 + I - \pi^2 \varphi_1^2 \quad [33]$$
$$\dot{\varphi}_2 = \Delta_2/\pi + 2\varphi_2 v_2 \quad [34]$$
$$\dot{v}_2 = v_2^2 + \Gamma_{21} s_1 + \Gamma_{22} s_2 + \bar{\eta}_2 - \pi^2 \varphi_2^2 \quad [35]$$

and the synapse equations, assuming the type of synapse is fixed for each source neuron,

$$\dot{s}_1 = z_1 \quad [36]$$
$$\dot{z}_1 = a_1^2 \varphi_1 - 2a_1 z_1 - a_1^2 s_1 \quad [37]$$
$$\dot{s}_2 = z_2 \quad [38]$$
$$\dot{z}_2 = a_2^2 \varphi_2 - 2a_2 z_2 - a_2^2 s_2 \quad [39]$$

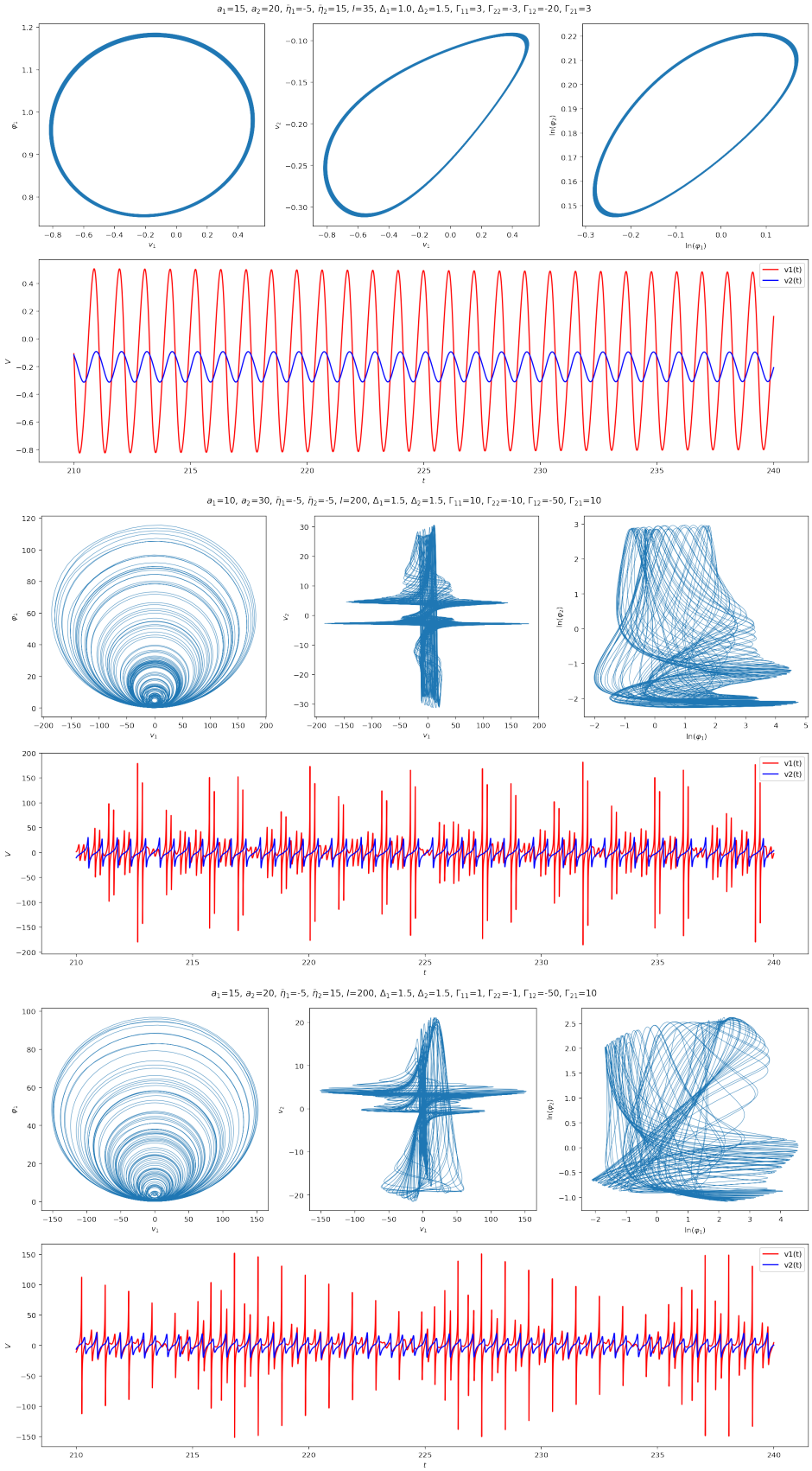


Fig. SI-14. Interesting dynamics for the NMM2 more general two population E-I model (III).