Self-consistent correlations of randomly coupled rotators in the asynchronous state (Supplemental Material)

Alexander van Meegen[1,2] and Benjamin Lindner[1,2]

[1] *Bernstein Center for Computational Neuroscience Berlin, Philippstr. 13, Haus 2, 10115 Berlin, Germany*
[2] *Physics Department of Humboldt University Berlin, Newtonstr. 15, 12489 Berlin, Germany*

(Dated: November 17, 2018)

I. DERIVATION OF EQUATIONS (2) AND (3) IN THE MAIN MANUSCRIPT

To derive the equations of the stochastic mean-field approximation, equations (2) and (3) in the main manuscript, we use the path integral formalism for many particle systems [1] and specifically the treatment of spin glass systems in [2]. In order to keep the presentation self-contained, we rederive and state a few well-known results in our own notation.

I.1 Characteristic Functional of the Deterministic System

The starting point is the characteristic functional for the vector-valued process $\mathbf{\Theta}(t) = (\Theta_1(t), \Theta_2(t), \dots)^T$ with probability functional $P_{\mathbf{\Theta}}[\mathbf{\Theta}(t)]$

$$\Phi_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)] = \int \mathcal{D}\mathbf{\Theta}(t)\ P_{\mathbf{\Theta}}[\mathbf{\Theta}(t)] \exp\left(i \int \mathrm{d}t\ \mathbf{\Theta}(t) \cdot \boldsymbol{\vartheta}(t)\right). \tag{S1}$$

Here, $\mathbf{\Theta}(t) \cdot \boldsymbol{\vartheta}(t) = \sum_{m=1}^{N} \Theta_m(t)\vartheta_m(t)$ and $\int \mathcal{D}\mathbf{\Theta}(t)$ denotes an N-dimensional path integral, i.e.

$$\int \mathcal{D}\mathbf{\Theta}(t) \ = \prod_{m=1}^{N} \int \mathcal{D}\Theta_m(t) \ = \lim_{N_t \to \infty} \prod_{n=1}^{N_t} \prod_{m=1}^{N} \int \mathrm{d}\Theta_m(t_n)$$

where $t_0, t_1, \dots, t_{N_t}$ is a reasonable discretization of the time interval of interest. Expressing the exponential in eq. (S1) as a series, we see that functional derivatives of the characteristic functional evaluated at $\boldsymbol{\vartheta}(t) = 0$ yield the moments of $\mathbf{\Theta}(t)$.

If $\mathbf{\Theta}(t)$ is governed by a deterministic dynamics, $\dot{\mathbf{\Theta}} = \boldsymbol{\omega} + \boldsymbol{K} f(\mathbf{\Theta})$ with $[\boldsymbol{K} f(\mathbf{\Theta})]_m \equiv \sum_{n \neq m} K_{mn} f[\exp(i\Theta_n)]$, we can formally relate the difference of the two sides to a random variable $\mathbf{\Gamma} = \dot{\mathbf{\Theta}} - \boldsymbol{\omega} - \boldsymbol{K} f(\mathbf{\Theta})$ that is δ-distributed, $P_{\mathbf{\Gamma}}[\mathbf{\Gamma}(t)] = \delta[\mathbf{\Gamma}(t)]$, and possesses the characteristic functional $\Phi_{\mathbf{\Gamma}}[\boldsymbol{\gamma}(t)] = \left\langle \exp[i \int \mathrm{d}t\ \mathbf{\Gamma}(t) \cdot \boldsymbol{\gamma}(\mathbf{t})] \right\rangle_{\mathbf{\Gamma}} = 1$. $P_{\mathbf{\Gamma}}[\mathbf{\Gamma}(t)]$ and $P_{\mathbf{\Theta}}[\mathbf{\Theta}(t)]$ are related by the transformation

$$P_{\mathbf{\Theta}}[\mathbf{\Theta}(t)] = P_{\mathbf{\Gamma}}[\dot{\mathbf{\Theta}}(t) - \boldsymbol{\omega} - \boldsymbol{K} f(\mathbf{\Theta}(t))]$$

where the Jacobian can be set to unity by choosing a non-anticipating discretization of the ordinary differential equation [3]. Using the inverse transform $P_{\mathbf{\Gamma}}[\mathbf{\Gamma}(t)] = \int \mathcal{D}\boldsymbol{\gamma}(t)\ \exp\left(-i \int \mathrm{d}t\ \boldsymbol{\gamma}(t) \cdot \mathbf{\Gamma}(t)\right) \Phi_{\mathbf{\Gamma}}[\boldsymbol{\gamma}(t)]$ with $\int \mathcal{D}\boldsymbol{\gamma}(t) = \lim_{N_t \to \infty} \prod_{n=1}^{N_t} \prod_{m=1}^{N} \frac{1}{2\pi} \int \mathrm{d}\gamma_m(t_n)$, we get

$$P_{\mathbf{\Theta}}[\mathbf{\Theta}(t)] = \int \mathcal{D}\boldsymbol{\gamma}(t)\ \exp\left(-i \int \mathrm{d}t\ \boldsymbol{\gamma}(t) \cdot [\dot{\mathbf{\Theta}}(t) - \boldsymbol{\omega} - \boldsymbol{K} f(\mathbf{\Theta}(t))]\right).$$

Inserting $P_{\mathbf{\Theta}}[\mathbf{\Theta}(t)]$ into eq. (S1) results in an explicit expression for the characteristic functional:

$$\Phi_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)] = \int \mathcal{D}\mathbf{\Theta}(t)\ \int \mathcal{D}\boldsymbol{\gamma}(t)\ \exp\left(-i \int \mathrm{d}t\ \boldsymbol{\gamma}(t) \cdot [\dot{\mathbf{\Theta}}(t) - \boldsymbol{\omega} - \boldsymbol{K} f(\mathbf{\Theta}(t))] + i \int \mathrm{d}t\ \mathbf{\Theta}(t) \cdot \boldsymbol{\vartheta}(t)\right). \tag{S2}$$

I.2 Disorder Average

We recall that $\Phi_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)]$ is the characteristic functional of a *deterministic* N-dimensional system with randomly chosen but temporally frozen parameters in $\boldsymbol{\omega}$ and $\boldsymbol{K}$. We assume that, with respect to the single-rotator statistics, the system in the limit $N \to \infty$ is self-averaging: averages over the elements of the network (population averages) are equal to averages over realizations of the frozen parameters (disorder averages). Hence, if we are interested in an effective *population-averaged* description of the single element, we can consider an effective characteristic functional

$$\hat{\Phi}_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)] = \langle\langle \Phi_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)]\rangle_{\boldsymbol{\omega}}\rangle_{\boldsymbol{K}}\,. \tag{S3}$$

Functional derivatives of $\hat{\Phi}_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)]$ evaluated at $\boldsymbol{\vartheta}(t)=0$ yield the disorder-averaged moments of $\mathbf{\Theta}(t)$, which becomes again apparent when we write the exponential in eq. (S1) as a series (note that $\Phi_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)]$ is already properly normalized, $\Phi_{\mathbf{\Theta}}[0]=1$, such that no renormalization for $\hat{\Phi}_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)]$ is needed [4]).

Since we assume $\boldsymbol{\omega}$ and $\boldsymbol{K}$ to be independently distributed, the two averages act on two different factors in eq. (S2). The first is

$$\left\langle e^{i\int \mathrm{d}t\ \boldsymbol{\gamma}(t)\cdot\boldsymbol{\omega}}\right\rangle_{\boldsymbol{\omega}} = \prod_{m=1}^{N} \phi_\omega\left(\int \mathrm{d}t\ \gamma_m(t)\right) = \exp\left[\sum_{m=1}^{N} \ln\phi_\omega\left(\int \mathrm{d}t\ \gamma_m(t)\right)\right] \tag{S4}$$

where we have used the characteristic function $\phi_\omega(k) = \langle \exp(ik\omega)\rangle_\omega$ of the natural frequencies and the fact that all ω_m are independent and identically distributed (i.i.d.). For the second factor involving $\boldsymbol{K}$, we use the assumption of i.i.d. Gaussian coefficients K_{mn} with mean $\bar{K}/N$ and variance K^2/N (and thus $\langle\exp(ixK_{mn})\rangle_{K_{mn}} = \exp[i\frac{\bar{K}}{N}x - \frac{K^2}{2N}x^2]$):

$$\begin{aligned}
\left\langle e^{i\int \mathrm{d}t\ \boldsymbol{\gamma}(t)\cdot\boldsymbol{K}f(\mathbf{\Theta}(t))}\right\rangle_{\boldsymbol{K}} &= \prod_{m=1}^{N}\prod_{\substack{n=1\\ n\neq m}}^{N} \exp\left[i\frac{\bar{K}}{N}\int \mathrm{d}t\ \gamma_m(t) f(\Theta_n(t)) - \frac{K^2}{2N}\left(\int \mathrm{d}t\ \gamma_m(t) f(\Theta_n(t))\right)^2\right] \\
&= \exp\left(i\frac{\bar{K}}{N}\sum_{m=1}^{N}\sum_{\substack{n=1\\ n\neq m}}^{N}\int \mathrm{d}t\ \gamma_m(t)f(\Theta_n(t)) - \frac{K^2}{2N}\sum_{m=1}^{N}\sum_{\substack{n=1\\ n\neq m}}^{N}\int \mathrm{d}t_1\int \mathrm{d}t_2\ \gamma_m(t_1)\gamma_m(t_2) f(\Theta_n(t_1))f(\Theta_n(t_2))\right).
\end{aligned}$$

We introduce the population-averaged moments of the time-dependent input,

$$\mu(t) = \frac{\bar{K}}{N}\sum_{n=1}^{N} f(\Theta_n(t)), \qquad C(t_1,t_2) = \frac{K^2}{N}\sum_{n=1}^{N} f(\Theta_n(t_1))f(\Theta_n(t_2)), \tag{S5}$$

as new variables which can be included by means of additional integrals:

$$\begin{aligned}
\left\langle e^{i\int \mathrm{d}t\ \boldsymbol{\gamma}(t)\cdot\boldsymbol{K}f(\mathbf{\Theta}(t))}\right\rangle_{\boldsymbol{K}} =& \int \mathcal{D}\mu(t)\ \delta\left[\mu(t) - \frac{\bar{K}}{N}\sum_{n=1}^{N} f(\Theta_n(t))\right]\int \mathcal{D}C(t_1,t_2)\ \delta\left[C(t_1,t_2) - \frac{K^2}{N}\sum_{n=1}^{N} f(\Theta_n(t_1))f(\Theta_n(t_2))\right] \\
&\times \exp\left(i\sum_{m=1}^{N}\int \mathrm{d}t\ \gamma_m(t)\mu(t) - \frac{1}{2}\sum_{m=1}^{N}\int \mathrm{d}t_1\int \mathrm{d}t_2\ \gamma_m(t_1)C(t_1,t_2)\gamma_m(t_2)\right).
\end{aligned}$$

Here we neglected an $O(1/N)$ terms in the exponent due to the constraint $n \neq m$ of the sums. We can express the delta-functionals by additional inverse transforms, the Fourier-arguments of which we scale with the system size:

$$\begin{aligned}
\left\langle e^{i\int \mathrm{d}t\ \boldsymbol{\gamma}(t)\cdot\boldsymbol{K}f(\mathbf{\Theta}(t))}\right\rangle_{\boldsymbol{K}} =& \int \mathcal{D}\mu(t)\ \int \mathcal{D}(Nm(t))\ e^{-i\int \mathrm{d}t\, m(t)[N\mu(t) - \bar{K}\sum_{n=1}^{N} f(\Theta_n(t))]} \\
&\times \int \mathcal{D}C(t_1,t_2)\ \int \mathcal{D}(Nc(t_1,t_2))\ e^{-i\int \mathrm{d}t_1\int \mathrm{d}t_2\, c(t_1,t_2)[NC(t_1,t_2) - K^2\sum_{n=1}^{N} f(\Theta_n(t_1))f(\Theta_n(t_2))]} \\
&\times \exp\left(i\sum_{m=1}^{N}\int \mathrm{d}t\ \gamma_m(t)\mu(t) - \frac{1}{2}\sum_{m=1}^{N}\int \mathrm{d}t_1\int \mathrm{d}t_2\ \gamma_m(t_1)C(t_1,t_2)\gamma_m(t_2)\right).
\end{aligned} \tag{S6}$$

Inserting eq. (S4) and eq. (S6) into the disorder-averaged characteristic functional, eq. (S3), yields

$$\begin{aligned}
\hat{\Phi}_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)] =& \int \mathcal{D}\mu(t)\ \int \mathcal{D}(Nm(t))\ \int \mathcal{D}C(t_1,t_2)\ \int \mathcal{D}(Nc(t_1,t_2))\ e^{-iN\int \mathrm{d}t\, m(t)\mu(t) - iN\int \mathrm{d}t_1\int \mathrm{d}t_2\, c(t_1,t_2)C(t_1,t_2)} \\
&\times \prod_{m=1}^{N}\left(\int \mathcal{D}\Theta_m(t)\ \int \mathcal{D}\gamma_m(t)\ \exp(S[\Theta_m(t),\gamma_m(t),\vartheta_m(t)])\right)
\end{aligned} \tag{S7}$$

where

$$S[\Theta(t),\gamma(t),\vartheta(t)] = -i\int \mathrm{d}t\ \gamma(t)\dot{\Theta}(t) + i\int \mathrm{d}t\ \Theta(t)\vartheta(t) + \ln\left[\phi_\omega\left(\left(\int \mathrm{d}t\ \gamma(t)\right)\right)\right]$$
$$+i\int \mathrm{d}t\ \gamma(t)\mu(t) - \frac{1}{2}\int \mathrm{d}t_1\int \mathrm{d}t_2\ \gamma(t_1)C(t_1,t_2)\gamma(t_2)$$
$$+i\bar{K}\int \mathrm{d}t\ m(t)f(\Theta(t)) + iK^2\int \mathrm{d}t_1\int \mathrm{d}t_2\ f(\Theta(t_1))c(t_1,t_2)f(\Theta(t_2)).$$

At this point, the dynamics are only coupled through the newly introduced fields $\mu(t)$, $m(t)$, $C(t,t')$ and $c(t,t')$. The decoupling of the N integrals at the price of introducing two additional integrals is equivalent to a Hubbard-Stratonovich (sometimes also called Gaussian) transformation [1].

I.3 Saddle-Point Approximation

We aim at a *stochastic* description of the system valid for $N \gg 1$. Making use of the large system size, we evaluate the integrals over $C(t,t')$ and $c(t,t')$ in a saddle-point approximation. To this end, we recast eq. (S7) into

$$\hat{\Phi}_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)] = \int \mathcal{D}\mu(t)\int \mathcal{D}(Nm(t))\int \mathcal{D}C(t_1,t_2)\int \mathcal{D}(Nc(t_1,t_2))\ e^{-iN\int \mathrm{d}t\, m(t)\mu(t) - iN\int \mathrm{d}t_1\int \mathrm{d}t_2\, c(t_1,t_2)C(t_1,t_2)}$$
$$\times \exp\left(N\ln\left\{\int \mathcal{D}\Theta(t)\ \int \mathcal{D}\gamma(t)\ \exp(S[\Theta(t),\gamma(t),\vartheta(t)])\right\}\right)$$

which admits a saddle-point approximation that becomes exact in the limit $N\to\infty$. At the saddle-point, the first variations of the exponent with respect to $\mu(t)$, $m(t)$, $C(t,t')$ and $c(t,t')$ have to vanish. The corresponding saddle-point equations read

$$im(t_0) = \frac{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ \frac{\delta S[\Theta(t),\gamma(t),\vartheta(t)]}{\delta\mu(t_0)}e^{S[\Theta(t),\gamma(t),\vartheta(t)]}}{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ e^{S[\Theta(t),\gamma(t),\vartheta(t)]}} = i\frac{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ \gamma(t_0)e^{S[\Theta(t),\gamma(t),\vartheta(t)]}}{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ e^{S[\Theta(t),\gamma(t),\vartheta(t)]}},$$
$$i\mu(t_0) = \frac{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ \frac{\delta S[\Theta(t),\gamma(t),\vartheta(t)]}{\delta m(t_0)}e^{S[\Theta(t),\gamma(t),\vartheta(t)]}}{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ e^{S[\Theta(t),\gamma(t),\vartheta(t)]}} = i\bar{K}\frac{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ f(\Theta(t_0))e^{S[\Theta(t),\gamma(t),\vartheta(t)]}}{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ e^{S[\Theta(t),\gamma(t),\vartheta(t)]}},$$
$$ic(t_1,t_2) = \frac{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ \frac{\delta S[\Theta(t),\gamma(t),\vartheta(t)]}{\delta C(t_1,t_2)}e^{S[\Theta(t),\gamma(t),\vartheta(t)]}}{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ e^{S[\Theta(t),\gamma(t),\vartheta(t)]}} = -\frac{1}{2}\frac{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ \gamma(t_1)\gamma(t_2)e^{S[\Theta(t),\gamma(t),\vartheta(t)]}}{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ e^{S[\Theta(t),\gamma(t),\vartheta(t)]}},$$
$$iC(t_1,t_2) = \frac{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ \frac{\delta S[\Theta(t),\gamma(t),\vartheta(t)]}{\delta c(t_1,t_2)}e^{S[\Theta(t),\gamma(t),\vartheta(t)]}}{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ e^{S[\Theta(t),\gamma(t),\vartheta(t)]}} = iK^2\frac{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ f(\Theta(t_1))f(\Theta(t_2))e^{S[\Theta(t),\gamma(t),\vartheta(t)]}}{\int\mathcal{D}\Theta(t)\ \int\mathcal{D}\gamma(t)\ e^{S[\Theta(t),\gamma(t),\vartheta(t)]}}.$$

Since $S[\Theta(t),\gamma(t),\vartheta(t)]$ depends on $\mu(t)$, $m(t)$, $C(t,t')$ and $c(t,t')$, the saddle-point equations constitute a coupled set of self-consistent equations for these four quantities.

A possible solution to the first and third saddle-point equations resulting in a properly normalized characteristic functional is $m_\xi(t) = c_\xi(t_1,t_2) = 0$; in a similar problem it can be shown that this is the only solution that ensures causality of the dynamics [2]. Denoting the (so far undetermined) solutions of the second and fourth saddle-point equations by $\mu_\xi(t)$ and $C_\xi(t_1,t_2)$, the disorder-averaged characteristic functional factorizes in the saddle-point approximation into

$$\hat{\Phi}_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)] = \left(\int\mathcal{D}\Theta(t)\int\mathcal{D}\gamma(t)\ e^{-i\int \mathrm{d}t\,\gamma(t)\dot{\Theta}(t)+\ln[\phi_\omega(\int \mathrm{d}t\,\gamma(t))]+i\int \mathrm{d}t\,\gamma(t)\mu_\xi(t)-\frac{1}{2}\int \mathrm{d}t_1\int \mathrm{d}t_2\,\gamma(t_1)C_\xi(t_1,t_2)\gamma(t_2)+i\int \mathrm{d}t\,\Theta(t)\vartheta(t)}\right)^N.$$

Here, we see that the dynamics decouple entirely in the saddle-point approximation. The sole remainder of the interaction are the network-averaged statistics of the input $\mu_\xi(t)$, $C_\xi(t_1,t_2)$ that were introduced in eq. (S5). In order to make sense of this result, we rewrite it as $\hat{\Phi}_{\mathbf{\Theta}}[\boldsymbol{\vartheta}(t)] = \prod_{m=1}^N \langle\Phi_{\Theta_m}[\vartheta_m(t)]\rangle_{\omega_m}$ with the one-particle characteristic functional

$$\Phi_{\Theta_m}[\vartheta_m(t)] = \int\mathcal{D}\Theta_m(t)\ P_\xi[\dot{\Theta}_m(t)-\omega_m]\exp\left(i\int \mathrm{d}t\ \Theta_m(t)\vartheta_m(t)\right)$$

and the probability functional

$$P_\xi[\xi(t)] = \int \mathcal{D}\gamma(t)\; e^{-i\int \mathrm{d}t\gamma(t)\xi(t)} \exp\left(i\int \mathrm{d}t\; \gamma(t)\mu_\xi(t) - \frac{1}{2}\int \mathrm{d}t_1 \int \mathrm{d}t_2\; \gamma(t_1)C_\xi(t_1,t_2)\gamma(t_2)\right). \tag{S8}$$

Using $P_{\Theta_m}[\Theta_m(t)] = P_\xi[\dot{\Theta}_m(t) - \omega_m]$ as in the beginning (again, the Jacobian can be set to unity with a non-anticipating discretization scheme), we see that $\Phi_{\Theta_m}[\vartheta_m(t)]$ is the characteristic functional of the dynamical equation

$$\dot{\Theta}_m(t) = \omega_m + \xi_m(t) \tag{S9}$$

driven by the stochastic process $\xi_m(t)$. According to eq. (S8), the characteristic functional of $\xi_m(t)$ is $\Phi_\xi[\gamma(t)] = \exp\left(i\int \mathrm{d}t\; \gamma(t)\mu_\xi(t) - \frac{1}{2}\int \mathrm{d}t_1 \int \mathrm{d}t_2\; \gamma(t_1)C_\xi(t_1,t_2)\gamma(t_2)\right)$, hence $\xi_m(t)$ is a Gaussian process with mean $\mu_\xi(t)$ and correlation function $C_\xi(t_1,t_2)$ [5].

The mean $\mu_\xi(t)$ and correlation function $C_\xi(t_1,t_2)$ are determined self-consistently by the second and fourth saddle-point equations. Setting $\vartheta(t) = 0$ on the r.h.s., they can be simplified to

$$\begin{aligned} \mu_\xi(t_0) &= \bar{K}\frac{\left\langle \int \mathcal{D}\Theta(t)\; P_\xi[\dot{\Theta}(t)-\omega] f(\Theta(t_0))\right\rangle_\omega}{\left\langle \int \mathcal{D}\Theta(t)\; P_\xi[\dot{\Theta}(t)-\omega]\right\rangle_\omega} = \bar{K}\left\langle \langle f(\Theta(t_0))\rangle_\xi \right\rangle_\omega, \\ C_\xi(t_1,t_2) &= K^2\frac{\left\langle \int \mathcal{D}\Theta(t)\; P_\xi[\dot{\Theta}(t)-\omega] f(\Theta(t_1))f(\Theta(t_2))\right\rangle_\omega}{\left\langle \int \mathcal{D}\Theta(t)\; P_\xi[\dot{\Theta}(t)-\omega]\right\rangle_\omega} = K^2\left\langle \langle f(\Theta(t_1))f(\Theta(t_2))\rangle_\xi \right\rangle_\omega \end{aligned} \tag{S10}$$

where we used that the denominator equals unity because $P_\xi[\xi(t)]$ is properly normalized ($\Phi_\xi[0] = 1$). This completes the derivation of the sought-for stochastic description of the dynamics, i.e. eqs. (2) and (3) in the main text.

Briefly recapitulating the above calculation, we showed that the disorder-averaged characteristic functional of the original dynamics $\langle\langle\Phi_{\boldsymbol{\Theta}}[\boldsymbol{\vartheta}(t)]\rangle_{\boldsymbol{\omega}}\rangle_{\boldsymbol{K}}$ equals the product of N independent characteristic functionals of the reduced dynamics eq. (S9) averaged over the natural frequencies $\langle\Phi_{\Theta_m}[\vartheta_m(t)]\rangle_{\omega_m}$ in the limit $N \to \infty$. Thus, using the assumption of self-averaging, the population-averaged moments of the original dynamics equal the moments of the stochastic dynamics averaged over the natural frequencies. In particular, this is true for the correlation function of the input — in other words, eqs. (S9) and (S10) can be used to obtain an effective stochastic input $\xi_m(t)$.

II. DERIVATION OF EQUATION (4) IN THE MAIN MANUSCRIPT

We consider the self-consistency conditions for the network correlations, eq. (S10), in the stationary state

$$\mu_\xi = \lim_{t_0\to\infty}\mu_\xi(t_0) = \bar{K}\lim_{t_0\to\infty}\langle f(\Theta(t_0))\rangle_{\omega,\xi}, \qquad C_\xi(\tau) = \lim_{t_0\to\infty}C_\xi(t_0,t_0+\tau) = K^2\lim_{t_0\to\infty}\langle f(\Theta(t_0))f(\Theta(t_0+\tau))\rangle_{\omega,\xi}.$$

Since $f(\Theta)$ is periodic in Θ, it admits a Fourier series expansion $f(\Theta) = \sum_{\ell=-\infty}^{\infty} A_\ell e^{i\ell\Theta}$. Inserting this on the r.h.s. results in

$$\mu_\xi = \bar{K}\sum_{\ell=-\infty}^{\infty} A_\ell \lim_{t_0\to\infty}\left\langle e^{i\ell\Theta(t_0)}\right\rangle_{\omega,\xi}, \qquad C_\xi(\tau) = K^2\sum_{\ell=-\infty}^{\infty}\sum_{\ell'=-\infty}^{\infty} A_\ell A_{\ell'}\lim_{t_0\to\infty}\left\langle e^{i\ell\Theta(t_0)}e^{i\ell'\Theta(t_0+\tau)}\right\rangle_{\omega,\xi}.$$

The invariance of the stationary state under rotation of the reference frame $\Theta(t) \to \Theta(t) + \alpha$ yields

$$\lim_{t_0\to\infty}\left\langle e^{i\ell\Theta(t_0)}\right\rangle_{\omega,\xi} = e^{i\ell\alpha}\lim_{t_0\to\infty}\left\langle e^{i\ell\Theta(t_0)}\right\rangle_{\omega,\xi} \quad\Rightarrow\quad \lim_{t_0\to\infty}\left\langle e^{i\ell\Theta(t_0)}\right\rangle_{\omega,\xi} = 0 \;\; \forall\; \ell \neq 0, \tag{S11}$$

$$\begin{aligned} &\lim_{t_0\to\infty}\left\langle e^{i\ell\Theta(t_0)}e^{i\ell'\Theta(t_0+\tau)}\right\rangle_{\omega,\xi} = e^{i(\ell+\ell')\alpha}\lim_{t_0\to\infty}\left\langle e^{i\ell\Theta(t_0)}e^{i\ell'\Theta(t_0+\tau)}\right\rangle_{\omega,\xi} \\ \Rightarrow\;& \lim_{t_0\to\infty}\left\langle e^{i\ell\Theta(t_0)}e^{i\ell'\Theta(t_0+\tau)}\right\rangle_{\omega,\xi} = 0 \;\; \forall\; \ell+\ell' \neq 0. \end{aligned} \tag{S12}$$

Using eq. (S11), eq. (S12), $A_\ell = A^*_{-\ell}$ and rewriting $\Theta(t_0+\tau) - \Theta(t_0) = \int_{t_0}^{t_0+\tau} \mathrm{d}t\ \dot{\Theta}(t)$, we obtain

$$\mu_\xi = \bar{K} A_0, \qquad C_\xi(\tau) = K^2 \sum_{\ell=-\infty}^{\infty} |A_\ell|^2 \lim_{t_0\to\infty} \left\langle e^{i\ell \int_{t_0}^{t_0+\tau} \mathrm{d}t\ \dot{\Theta}(t)} \right\rangle_{\omega,\xi}. \tag{S13}$$

Within the assumption of rotation invariance in the stationary state, the self-consistency equation for the stationary mean value of the network noise becomes trivial. Thus, we can simply account for $\mu_\xi = \bar{K}A_0$ by rescaling $\omega_0 + \mu_\xi \to \omega_0$ and $\xi(t) - \mu_\xi \to \xi(t)$.

To further simplify the self-consistency equation for the correlation function we insert eq. (S9) into the integral, thereby ensuring that $\Theta(t)$ fulfills its dynamical equation. This decouples the average values:

$$C_\xi(\tau) = K^2 \sum_{\ell=-\infty}^{\infty} |A_\ell|^2 \left\langle e^{i\ell\omega\tau} \right\rangle_\omega \lim_{t_0\to\infty} \left\langle e^{i\ell \int_{t_0}^{t_0+\tau} \mathrm{d}t\ \xi(t)} \right\rangle_\xi.$$

The average over the natural frequency ω results in the characteristic function $\left\langle e^{i\ell\omega\tau} \right\rangle_\omega = \phi_\omega(\ell\tau)$, the average over the Gaussian process $\xi(t)$ in the characteristic functional $\left\langle e^{i\int \mathrm{d}t\ \gamma(t)\xi(t)} \right\rangle_\xi = e^{-\frac{1}{2}\int \mathrm{d}t_1\ \int \mathrm{d}t_2\ \gamma(t_1) C_\xi(t_1,t_2)\gamma(t_2)}$. Explicitly inserting the test function $\gamma(t) = \ell \mathbb{1}_{[t_0,t_0+\tau]}(t)$ (where $\mathbb{1}_X(t)$ denotes the indicator function of a set X) into the characteristic functional leads to

$$C_\xi(\tau) = K^2 \sum_{\ell=-\infty}^{\infty} |A_\ell|^2 \phi_\omega(\ell\tau) \lim_{t_0\to\infty} e^{-\frac{1}{2}\ell^2 \int_{t_0}^{t_0+\tau} \mathrm{d}t_1\ \int_{t_0}^{t_0+\tau} \mathrm{d}t_2\ C_\xi(t_1,t_2)}.$$

In the exponent, the stationary auto-correlation function of the network noise appears. Using its time-translation invariance and its symmetry under time-reversal, we obtain

$$C_\xi(\tau) = K^2 \sum_{\ell=-\infty}^{\infty} |A_\ell|^2 \phi_\omega(\ell\tau) e^{-\ell^2 \int_0^\tau \mathrm{d}u\ (\tau-u) C_\xi(u)} \tag{S14}$$

which provides a non-linear integral equation for $C_\xi(\tau)$, i.e. equation (4) in the main manuscript.

III. DERIVATION OF EQUATIONS (10) AND (14) IN THE MAIN MANUSCRIPT

III.1 Derivation of equations (10)

If the interaction function is a simple sine function and all oscillators possess the same natural frequency ω_0, the differential equation for Λ attains the simple form [eq. (9) in the main manuscript]

$$\ddot{\Lambda} = \frac{K^2}{2} \cos(\omega_0 \tau) e^{-\Lambda}. \tag{S15}$$

For sufficiently small ω_0, the r.h.s. of eq. (S15) becomes essentially zero before the cosine function deviates significantly from one. Consequently, we can solve eq. (S15) for $\omega_0 = 0$, yielding

$$\Lambda(\tau) = \ln[\cosh^2(K\tau/2)],$$

and use this solution in equation (7) from the main manuscript (with $\omega_0 \neq 0$) to determine the correlation function

$$C_x(\tau) \approx \frac{\exp(i\omega_0\tau)}{\cosh^2(K\tau/2)}, \qquad S_x(\omega) \approx \frac{4\pi(\omega-\omega_0)/K^2}{\sinh(\pi(\omega-\omega_0)/K)} \tag{S16}$$

where the power spectrum follows by Fourier transform of $C_x(\tau)$.

III.2 Derivation of equations (14)

The starting point are the equations for the autocorrelation and power spectrum of a single rotator

$$C_x(\tau) = \exp\left(i\omega_0\tau - \frac{K^2}{2}\int_0^\tau \mathrm{d}t\ (\tau - t)\mathrm{Re}\left(C_x(t)\right)\right), \tag{S17}$$

$$S_x(\omega) = \frac{K^2/4}{\omega - \omega_0}\int \frac{\mathrm{d}\omega'}{2\pi}\frac{S_x(\omega')}{\omega'}(S_x(\omega - \omega') - S_x(\omega + \omega')). \tag{S18}$$

Here we combined equations (6), (7) and (9) from the main manuscript to obtain the first equation and the latter follows from the first by temporal differentiation and Fourier transform.

For weak coupling $K \ll \omega_0$, we expect the power spectrum to be of the form $S_x(\omega) \approx \tilde{f}(\omega - \omega_0)$ where $\tilde{f}(\omega)$ is a strongly peaked, symmetric function. Inserting this Ansatz into eq. (S18), we get two non-vanishing contributions on the r.h.s. around $\omega = 0 + \Delta\omega$ and $\omega = 2\omega_0 + \Delta\omega$:

$$S_x(0 + \Delta\omega) \approx \frac{K^2/4}{\omega_0 - \Delta\omega}\int \frac{\mathrm{d}\omega'}{2\pi}\frac{\tilde{f}(\omega' - \omega_0)}{\omega'}\tilde{f}(\Delta\omega - \omega_0 + \omega') \approx \frac{K^2}{4\omega_0^2}\int \frac{\mathrm{d}\omega'}{2\pi}\tilde{f}(\omega' - \omega_0)\tilde{f}(\Delta\omega - \omega_0 + \omega'),$$

$$S_x(2\omega_0 + \Delta\omega) \approx \frac{K^2/4}{\omega_0 + \Delta\omega}\int \frac{\mathrm{d}\omega'}{2\pi}\frac{\tilde{f}(\omega' - \omega_0)}{\omega'}\tilde{f}(\Delta\omega + \omega_0 - \omega') \approx \frac{K^2}{4\omega_0^2}\int \frac{\mathrm{d}\omega'}{2\pi}\tilde{f}(\omega' - \omega_0)\tilde{f}(\Delta\omega + \omega_0 - \omega').$$

In the first step, we neglected one of the two convolutions which hardly contributes to the r.h.s. because one of the factors in its integrand is always far away from the peak for sufficiently small $\Delta\omega$. In the second step, we approximated $1/\omega' \approx 1/\omega_0$ in the integrand (i.e. we assume $\tilde{f}(\omega' - \omega_0)$ to vanish much faster than $1/\omega'$ varies) and $1/(\omega_0 \pm \Delta\omega) \approx 1/\omega_0$. Including these two contributions at $\omega \approx 0$ and $\omega \approx 2\omega_0$, a successive Fourier transform yields the ansatz

$$C_x(\tau) \approx f(\tau)e^{i\omega_0\tau}\left[1 + \tfrac{K^2}{4\omega_0^2}f(\tau)\left(e^{-i\omega_0\tau} + e^{i\omega_0\tau}\right)\right] \tag{S19}$$

for the autocorrelation function where $f(\tau)$ denotes the inverse Fourier transform of $\tilde{f}(\omega)$. Particularly, the convolution in Fourier space becomes a product in the temporal domain.

To determine $f(\tau)$, we insert eq. (S19) into eq. (S17) and retain only the leading term on the l.h.s.:

$$f(\tau) \approx \exp\left(-\frac{K^2}{2}\int_0^\tau \mathrm{d}t\ (\tau - t)\{\cos(\omega_0 t)f(t) + \tfrac{K^2}{4\omega_0^2}f^2(t)[1 + \cos(2\omega_0 t)]\}\right).$$

Assuming $f(t) = f(\varepsilon t)$ with $\varepsilon = O(K/\omega_0) \ll 1$ (i.e. it varies slowly), the oscillatory integrals in the exponent on the r.h.s. are $O((K/\omega_0)^2)$ while the non-oscillatory integral is $O(1)$. Neglecting the oscillatory integrals yields the integral equation

$$f(\tau) \approx \exp\left(-\frac{K^4}{8\omega_0^2}\int_0^\tau \mathrm{d}t\ (\tau - t)f^2(t)\right).$$

Introducing $I(\tau) = \int_0^\tau \mathrm{d}t\ (\tau - t)f^2(t)$, we get the differential equation $\ddot{I} = \exp\left(-\frac{K^4}{4\omega_0^2}I\right)$ with initial conditions $I(0) = \dot{I}(0) = 0$. Its solution $I(\tau) = 8\omega_0^2\ln[\cosh(K^2\tau/\sqrt{8}\omega_0)]/K^4$ leads to $f(\tau) = \sqrt{\ddot{I}(\tau)} = 1/\cosh(K^2\tau/\sqrt{8}\omega_0)$ and thus

$$\begin{aligned} C_x(\tau) &\approx \frac{e^{i\omega_0\tau}}{\cosh(K^2\tau/\sqrt{8}\omega_0)}\left(1 + \frac{K^2}{2\omega_0^2}\frac{\cos(\omega_0\tau)}{\cosh(K^2\tau/\sqrt{8}\omega_0)}\right), \\ S_x(\omega) &\approx \frac{\sqrt{8}\pi\omega_0/K^2}{\cosh(\sqrt{2}\pi\omega_0(\omega - \omega_0)/K^2)} + \frac{2\pi(\omega - 2\omega_0)/K^2}{\sinh(\sqrt{2}\pi\omega_0(\omega - 2\omega_0)/K^2)} + \frac{2\pi\omega/K^2}{\sinh(\sqrt{2}\pi\omega_0\omega/K^2)} \end{aligned} \tag{S20}$$

where the corresponding power spectrum follows by Fourier transform. Note that this solution accounts for the peaks at $\omega = \omega_0$, $\omega = 0$ and $\omega = 2\omega_0$, see Fig. S1. In eq. (14) in the main manuscript, we presented only the solution for the main peak and neglected the latter two contributions (peaks at $\omega = 0$ and $\omega = 2\omega_0$). In the exponentially small gaps between the three peaks we see deviations from the theory and, by construction, the theory does not account for further peaks, cf. Fig. S1.

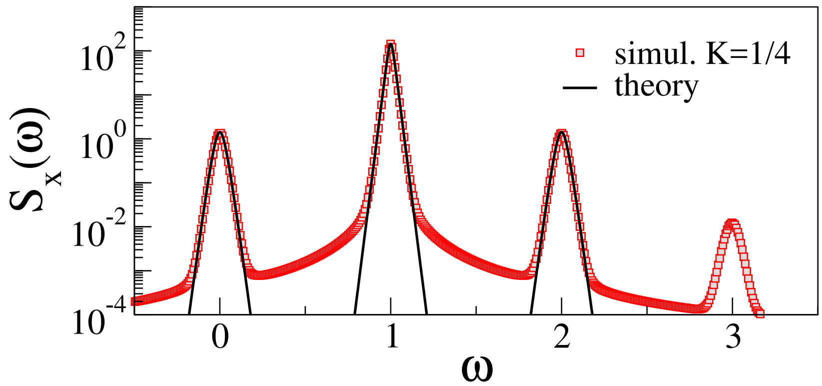

FIG. S1. **Theory accounts for multiple peaks in weak coupling limit**. Equivalent data to Fig. 3a in the main manuscript but with larger x-axis range and logarithmic y-axis scale. The three peaks correspond to the three terms in eq. (S20).

IV. MAPPING SPIKING INTEGRATE-AND-FIRE NETWORK TO ROTATOR NETWORK

To apply our theory to the self-consistent autocorrelation statistics of spiking neurons in sparse recurrent networks, we first extend the rotator theory itself and then describe how to map the parameters of the spiking network to it.

We consider the modified rotator network where the coupling function mimics a sequence of incoming spikes, $\sum_i \delta(t-t_n^i)$ with t_n^i denoting the ith spike time of the nth presynaptic neuron. If these spikes result from a presynaptic variable reaching a threshold, or, in terms of a phase variable, the nth presynaptic phase crossing a multiple of 2π from below, we can approximate the input stream as follows:

$$\sum_i \delta(t-t_n^i) \approx \sum_{k=-\infty}^{\infty} \dot{\Theta}_n \delta(\Theta_n - 2\pi k).$$

Let us briefly motivate the rationale behind this approximation: Using the transformation property $\delta(x-x_0) = |g'(x_0)|\delta(g(x))$ of the Dirac delta function, it is in principle possible to transform the argument from time to phase. However, the absolute value of the phase velocity $|\dot{\Theta}|$ appears in the formula. We deliberately excluded the absolute value because threshold crossings from below occur with a positive velocity while crossings from above occur with a negative one. Without the absolute value in the coupling, a rapid series of successive crossings from above and below averages out over short time intervals such that only one net crossing from below remains, mimicking the fire-and-reset rule in integrate-and-fire neuron models. Clearly, this crucially depends on the individual rotator being in a mean-driven regime, i.e. $\omega_0 > 0$.

IV.1 Derivation of equation (17)

If we incorporate the derivative into the phase dynamics, we obtain the modified dynamics

$$\dot{\Theta}_m = \omega_m + \sum_{n\neq m} K_{mn} \dot{\Theta}_n f(\Theta_n).$$

Instead of a Dirac comb $f(\Theta) = \sum_{k=-\infty}^{\infty} \delta(\Theta - 2\pi k)$, we use for the coupling function a regularized version of it, given by a sum of narrow-centered Gaussians $f(\Theta) = \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{2\pi\lambda^2}} e^{-(\Theta-2\pi k)^2/2\lambda^2}$ with $\lambda \ll 2\pi$. The reason for this regularization is twofold: First, the Dirac comb leads to a diverging term in eq. (17) [for $\tau = 0$, $\vartheta_3(0, e^{-\lambda^2}) \to \infty$ for $\lambda \to 0$] that is numerically not treatable. Second, we assume the finite width to support the averaging of successive crossing from above and below. The regularization affects only the high frequency behavior of the spectra, in particular the spectra do not saturate. For numerical solutions of eq. (17), we use the smallest value for λ that is numerically stable.

Adding the velocity term $\dot{\Theta}_n$ to the coupling does not alter the technical derivation of the stochastic mean-field approximation such that we obtain again eq. (S9) for the stochastic dynamics while the self-consistency equations change to

$$\mu_\xi(t) = \bar{K}\langle \dot{\Theta}(t) f(\Theta(t))\rangle_{\xi,\omega}, \qquad C_\xi(t,t') = K^2 \langle \dot{\Theta}(t) f(\Theta(t)) \dot{\Theta}(t') f(\Theta(t'))\rangle_{\xi,\omega}. \tag{S21}$$

Writing the coupling function as a Fourier series $f(\Theta)=\sum_\ell A_\ell e^{i\ell\Theta}$ and using the assumption of rotation invariace (eq. (S11) and eq. (S12) still hold because $\frac{d}{dt}(\Theta(t)+\alpha)=\dot{\Theta}(t)$ for any constant α), we get for the stationary mean and correlation function

$$\mu_\xi=\bar{K}A_0\lim_{t\to\infty}\left\langle\dot{\Theta}(t)\right\rangle_{\omega,\xi},\qquad C_\xi(t'-t)=K^2\sum_{\ell=-\infty}^{\infty}|A_\ell|^2\lim_{t,t'\to\infty}\left\langle\dot{\Theta}(t)\dot{\Theta}(t')e^{i\ell[\Theta(t')-\Theta(t)]}\right\rangle_{\omega,\xi}.$$

The equation for the mean network noise is straightforward to solve: Using eq. (S9), we get $\lim_{t\to\infty}\left\langle\dot{\Theta}(t)\right\rangle_{\omega,\xi}=\omega_0+\mu_\xi$ and thus

$$\mu_\xi=\frac{\bar{K}A_0}{1-\bar{K}A_0}\omega_0. \tag{S22}$$

With μ_ξ known, we rescale $\omega_0+\mu_\xi\to\omega_0$ and $\xi(t)-\mu_\xi\to\xi(t)$ to account for it.

For the correlation function we get rid of the velocities terms by temporal differentiation, leading to

$$C_\xi(t'-t)=K^2|A_0|^2\lim_{t,t'\to\infty}\left\langle\dot{\Theta}(t)\dot{\Theta}(t')\right\rangle_{\omega,\xi}+2K^2\mathrm{Re}\sum_{\ell=1}^{\infty}\frac{|A_\ell|^2}{\ell^2}\lim_{t,t'\to\infty}\frac{d}{dt}\frac{d}{dt'}\left\langle e^{i\ell[\Theta(t')-\Theta(t)]}\right\rangle_{\omega,\xi}.$$

Using eq. (S9), it is straightforward to evaluate $\lim_{t,t'\to\infty}\left\langle\dot{\Theta}(t)\dot{\Theta}(t')\right\rangle_{\omega,\xi}$ because $\xi(t)$ and ω are independent and $\langle\xi(t)\rangle_\xi=0$ due to the rescaling of ω_0. For the second term, we use eq. (S9) in the exponent, the characteristic function of the natural frequencies $\left\langle e^{i\omega x}\right\rangle_\omega=\phi_\omega(x)$ and the characteristic functional of the network noise $\left\langle e^{i\int \mathrm{d}t\ \gamma(t)\xi(t)}\right\rangle_\xi=e^{-\frac{1}{2}\int\mathrm{d}t_1\int\mathrm{d}t_2\gamma(t_1)C_\xi(t_1,t_2)\gamma(t_2)}$ to get (compare section II.)

$$C_\xi(\tau)=K^2|A_0|^2[C_\xi(\tau)+\sigma_\omega^2+\omega_0^2]-2K^2\mathrm{Re}\sum_{\ell=1}^{\infty}\frac{|A_\ell|^2}{\ell^2}\frac{d^2}{d\tau^2}\phi_\omega(\ell\tau)e^{-\ell^2\int_0^\tau \mathrm{d}t\,(\tau-t)C_\xi(t)}.$$

Here, we also defined $\tau=t'-t$ and used $\frac{d}{dt}\frac{d}{dt'}f(t'-t)=-\frac{d^2}{d\tau^2}f(\tau)$. In terms of $\Lambda(\tau)=\int_0^\tau\mathrm{d}t\ (\tau-t)C_\xi(t)$, we get the ordinary differential equation

$$\ddot{\Lambda}=K^2|A_0|^2[\ddot{\Lambda}+\sigma_\omega^2+\omega_0^2]-2K^2\mathrm{Re}\sum_{\ell=1}^{\infty}\frac{|A_\ell|^2}{\ell^2}\frac{d^2}{d\tau^2}\phi_\omega(\ell\tau)e^{-\ell^2\Lambda} \tag{S23}$$

that determines the self-consistent correlation function of the network noise.

The Fourier coefficients of the regularized Dirac comb are given by

$$A_\ell=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{dx}{\sqrt{2\pi\lambda^2}}\exp\left(-\frac{x^2}{2\lambda^2}\right)e^{-i\ell x}\approx\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{dx}{\sqrt{2\pi\lambda^2}}\exp\left(-\frac{x^2}{2\lambda^2}\right)e^{-i\ell x}=\frac{1}{2\pi}e^{-\frac{1}{2}\lambda^2\ell^2}$$

where we used $\lambda\ll 2\pi$ to extend the domain of integration to infinity. For Gaussian distributed natural frequencies, their characteristic function reads $\phi_\omega(x)=\exp\left(i\omega_0x-\frac{1}{2}x^2\sigma_\omega^2\right)$. Inserting both A_ℓ and $\phi_\omega(x)$ into eq. (S23), we obtain a closed form in terms of Jacobi Theta functions $\vartheta_3(z,q)=\sum_{\ell=-\infty}^{\infty}e^{2i\ell z}q^{\ell^2}$ and its derivatives $\vartheta_3'(z,q)=\partial_z\vartheta_3(z,q)$:

$$\begin{aligned}\ddot{\Lambda}=\frac{K^2}{4\pi^2}\{&(\ddot{\Lambda}+\sigma_\omega^2+\omega_0^2)\vartheta_3[\tfrac{1}{2}\omega_0\tau,e^{-\frac{1}{2}\sigma_\omega^2\tau^2-\lambda^2-\Lambda}]+\omega_0(\sigma_\omega^2\tau+\dot{\Lambda})\vartheta_3'[\tfrac{1}{2}\omega_0\tau,e^{-\frac{1}{2}\sigma_\omega^2\tau^2-\lambda^2-\Lambda}]\\&+\tfrac{1}{4}(\sigma_\omega^2\tau+\dot{\Lambda})^2\vartheta_3''[\tfrac{1}{2}\omega_0\tau,e^{-\frac{1}{2}\sigma_\omega^2\tau^2-\lambda^2-\Lambda}]\}.\end{aligned} \tag{S24}$$

With $A_0=1/2\pi$, eq. (S22) determines the self-consistent mean value of the network noise as $\mu_\xi=\bar{K}\omega_0/(2\pi-\bar{K})$.

IV.2 Integrate-and-fire network

We consider a sparse heterogeneous network of N_E excitatory and N_I inhibitory exponential integrate-and-fire (EIF) neurons [6] which are connected with probability ε and strength J_E ($J_I=-gJ_E$) for excitatory (inhibitory) connections. The single neuron dynamics obeys eq. (16) in the main manuscript (here recasted into a slightly different form):

$$\tau_m \dot{V}_i = -(V_i - V_{\rm leak}) + \Delta_T e^{-(V_i - V_{\rm th})/\Delta_T} + RI_{{\rm ext},i} + RI_{{\rm rec},i}(t). \tag{S25}$$

The model is supplemented by the fire-and-reset rule as outlined in the main text. The total recurrent input $RI_{{\rm rec},i}(t) = RI_{E,i}(t) + RI_{I,i}(t)$ consists of the excitatory/inhibitory recurrent contributions given by

$$RI_{E/I,i}(t) = \tau_m J_{E/I} \sum_{j \in N_{E/I}} C_{ij} x_j(t - \tau_D). \tag{S26}$$

Here, $x_i(t) = \sum_k \delta(t - t_i^k)$ is the spike train of neuron i and τ_D is the transmission delay. The random connectivity matrix C_{ij} possesses i.i.d. distributed entries that are one with probability ε and zero otherwise. The external input current $RI_{{\rm ext},i}$ is assumed to be Gaussian distributed across neurons. Note that all neurons (no matter whether excitatory or inhibitory) are statistically equivalent in their cellular dynamics. However, because of the subdivision into two subpopulations of E and I cells, the network obeys Dale's law, i.e. excitatory (inhibitory) neurons have an exclusively excitatory (inhibitory) effect on all their postsynaptic partners.

For the simulation of the network shown in Fig. 4 of the main manuscript, we used $N_E = 10000$ excitatory and $N_I = 2500$ inhibitory neurons with a connection probability of $\varepsilon = 0.1$. The strength of excitatory connections was $J_E = 0.05$mV and the relative strength of inhibitory ones $g = 4.5$. We chose a transmission delay of $\tau_D = 1.5$ms and $RI_{\rm ext} = 30$mV for the mean value of the external input current with a standard deviation of $\sigma_{\rm ext} = 0.1 RI_{\rm ext}$. The single neuron parameters were given by $V_{\rm th} = 20$mV, $V_{\rm leak} = 0$mV, $V_{\rm reset} = 0$mV, $V_{\rm peak} = 21$mV, $\Delta_T = 2$mV, $\tau_m = 20$ms and $t_{\rm ref} = 2$ms.

IV.3 Mapping integrate-and-fire network parameters to rotator theory

As a first step towards the rotator model, we reduce the EIF network obeying Dale's law to a fully connected network of perfect integrate-and-fire (PIF) neurons with Gaussian distributed weights; this network does not obey Dale's law anymore. After rescaling time $t/\tau_m \to t$ and setting the transmission delay τ_D to zero, the dynamical equations of the PIF neurons and their recurrent inputs are

$$\dot{V}_i = RI_{{\rm ext},i} + RI_{{\rm rec},i}(t), \qquad RI_{{\rm rec},i}(t) = \sum_j K_{ij} x_j(t). \tag{S27}$$

The fire-and-reset rule is modified to produce a spike and reset the voltage to $V_{\rm reset}$ whenever the voltage crosses $V_{\rm th}$; the refractory period is discarded. The first two cumulants of the fully connected Gaussian topology are set to match those of the sparse network obeying Dale's law, i.e.

$$\bar{K} = J_E(N_E - gN_I)\varepsilon, \qquad K^2 = J_E^2(N_E + g^2 N_I)\varepsilon(1-\varepsilon). \tag{S28}$$

The simplified dynamics eq. (S27) and connectivity eq. (S28) readily suggest to identify $V_i \to \Theta_i$, $RI_{{\rm ext},i} \to \omega_i$ and $RI_{{\rm rec},i}(t) \to \xi_i(t)$.

However, the strong alteration of the dynamics from eq. (S25) to eq. (S27) causes changes already in the mean activity of the neurons, i.e. in the mean firing rate. To preserve the mean activity, we change the mean external input current to

$$\langle RI_{{\rm ext},i} \rangle_{I_{\rm ext}} = (V_{\rm th} - V_{\rm reset} - \bar{K})\nu_0 \tag{S29}$$

while keeping its relative standard deviation. Here, ν_0 denotes the targeted average firing rate which could be obtained from a simulation of the sparse EIF network. To keep the mapping entirely independent of simulations, we use instead a Poisson-noise-approximation to calculate the mean firing rate of a sparse network of leaky integrate-and-fire (LIF) neurons [7]. This is possible because the simplification of the single neuron dynamics from EIF to LIF ($\Delta_T \to 0$ in eq. (S25)) affects the mean firing rate only weakly. With the modified external input current, eq. (S29), we are able to perform the mapping $V_i \to \Theta_i$, $RI_{{\rm ext},i} \to \omega_i$ and $RI_{{\rm rec},i}(t) \to \xi_i(t)$ to obtain the results shown in figure 4 in the main manuscript.

[1] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets* (World Scientific, New Jersey, 2009).

[2] H. Sompolinsky and A. Zippelius, Phys. Rev. B **25**, 6860 (1982).
[3] R. L. Stratonovich, in *Noise in nonlinear dynamical systems*, Vol. 1 (Cambridge Univ. Press, U. K., 1989) pp. 16–71.
[4] C. De Dominicis, Phys. Rev. B **18**, 4913 (1978).
[5] P. Hänggi, Z. Phys. B - Condensed Matter **75**, 275 (1989).
[6] L. Badel, S. Lefort, R. Brette, C. C. H. Petersen, W. Gerstner, and M. J. E. Richardson, J. Neurophysiol. **99**, 656 (2008).
[7] N. Brunel, J. Comput. Neurosci. **8**, 183 (2000).