# Dynamical Theory for Adaptive Systems

**Tuan Minh Pham and Kunihiko Kaneko**

Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, Copenhagen, 2100-DK, Denmark

E-mail: tuan.pham@nbi.ku.dk

**Abstract.** The investigation of adaptive dynamics, involving many degrees of freedom on two separated timescales, one for fast changes of state variables and another for the slow adaptation of parameters controlling the former's dynamics is crucial for understanding feedback mechanisms underlying evolutionary and learning processes. We present an extension of the Martin–Siggia–Rose-De Dominicis–Janssen (MSRDJ) path-integral approach to the study of nonequilibrium phase transitions in such dynamical systems. As an illustration, we apply our framework to biological adaptation under the genotype-phenotype feedback: phenotypic variations are shaped by the fast stochastic gene-expression dynamics and are coupled to the slow evolution of the distribution of genotypes, each encoded by a gene-regulatory network architecture. We establish that under this coevolution, genotypes responsible for high fitness are selected, leading to the emergence of phenotypic robustness within an intermediate level of environmental noise in reciprocal genetic networks.

## 1. Introduction

In biological or neural systems parameters that control the dynamics of the state variables, such as the set of couplings among the system degrees of freedom (dofs), often change on a slower timescale compared to that of the dofs' dynamics. Moreover, such systems become *adaptive* if these parameters need to adjust to the long-time state of the fast dofs. Adaptive multiple-timescale dynamical systems include cellular adaptation [1–4], cell differentiation with epigenetic modifications [5–9], neural networks with synaptic plasticity [10, 11] or adaptation and synaptic filtering [12, 13], retrieval of sequences via interaction modulation [14], eco-evolutionary dynamics [15–17], as well as extensive adaptive network models [18, 19]. This class of systems has been analysed mostly in low-dimensional models using tools of dynamical systems [20], while dynamical mean-field theory (DMFT) [21–25], a powerful tool to study a large variety of systems with *quenched* disorder [26–51], does not consider adaptively *slow* changes in the parameters at all. Previous works on a related class of systems with *partially annealed* disorder and two temperatures, one for the fast dofs and another for the slow ones,

often assume relaxational dynamics toward thermodynamic equilibrium [52–60]. High-dimensional nonequilibrium adaptive systems have challenged these existing approaches and been so far much less studied (but see [61] and reference therein for a few exceptions).

Living systems belong to this class of stochastic adaptive dynamics with two well-separated timescales: phenotypes evolve under noise at cellular and molecular levels [62–67], while genotypes – the rule of such developmental dynamics – change under selection and mutation on a slower timescale, depending on the fitness advantage of the resulting phenotypes. A central question in evolutionary biology is which features of the genotype-phenotype maps give rise to phenotypic robustness against perturbations induced by noise and that against mutation [68–74]. The answer to this question together with the proof of a hypothesis about the existence of a correlation between these two different types of robustness under sufficient noise, as suggested by extensive numerical studies [74–79], nevertheless, remains elusive. A systematic approach is hence required to address this question. Such an approach needs to bridge a strong disconnection between the actual nonequilibrium underlying dynamics of development and their approximate treatment based on (quasi)potential landscape picture [80, 81].

In this paper, to account for the *slow* adaptation of the controlling parameters, we develop a so-called adaptive DMFT (ADMFT), thus leveraging the original use of DMFT in quenched disorder systems to adaptive ones. The key point of our approach which is based on the Martin–Siggia–Rose-De Dominicis–Janssen (MSRDJ) formalism [22, 23] is to derive effective dynamics for both fast and slow variables in the thermodynamic limit from the moment generating functional of their joint trajectories. We then apply our framework to gene-regulatory networks with genotype-phenotype feedback. Here we find three phases with regard to the phenotypic and genotypic states, among which of biological relevance is a region called *robust* phase, where both genotype and phenotype achieve high values. We show that in reciprocal networks this robust region exists at intermediate noise strength by describing the transitions leading to the emergence of robustness as the onset of instability of phenotypes with zero gene-expression levels.

## 2. Problem Formulation

We consider a system consisting of $N$ units whose states are characterised by a vector $\mathbf{x}$. The dynamics of $\mathbf{x}$ take place on a continuous time $t \in [0, \infty)$. Let $\mathbf{J}$ denote the set of parameters controlling the dynamics of $\mathbf{x}$, such as the set of interaction couplings among $\mathbf{x}$'s components. Elements of $\mathbf{J}$ are assumed to be updated over discrete generations $\tau = 0, 1, 2, \cdots, T_{\max}$. Let $\mathbf{J}(\tau)$ denote a configuration of $\mathbf{J}$ at generation $\tau$. Note that without adaptation, i.e. if there is no dynamics defined over the space of $\mathbf{J}$, one could interpret $\tau$ as the index of a $\mathbf{J}(\tau)$ configuration in the latter space. Here to stress that the fast dynamics of $\mathbf{x}$ depends on which set $\mathbf{J}(\tau)$ of parameters that is currently applied to it, we denote $\mathbf{x}$'s state by $x_k(t;\tau) \equiv x_k(t|\mathbf{J}(\tau))$, $k = 1, \cdots, N$, meaning that $x_k$ is an explicit function of $t$ only for any given $\tau$. We focus on adaptive dynamics with time-scale separation:

- (i) the fast dynamics of $\mathbf{x}(t;\tau)$ during which $\mathbf{J}(\tau)$ is kept *fixed* are assumed to relax toward a non-equilibrium steady state (NESS) with the corresponding distribution $P(\mathbf{x}|\mathbf{J}(\tau))$ as $t \to \infty$.
- (ii) once $\mathbf{x}$ reaches a NESS, all elements $J_{kj}(\tau)$ of $\mathbf{J}(\tau)$ evolve synchronously to $J_{kj}(\tau+1)$, each follows the direction set by a field $h_{kj}(\tau)$. In our problem, the fields $h_{kj}(\tau)$ represent feedback from the steady-state distribution of $\mathbf{x}$ to the adaptation dynamics of $\mathbf{J}(\tau)$. Therefore, these feedback fields need to be functions of some *fitness* $\Psi(\tau)$ – a scalar that contains the information about $\mathbf{x}$ and the derivatives of $\Psi(\tau)$ with respect to genetic variables ($J_{kj}$ in this case), i.e.,

$$h_{kj}(\tau) = h_{kj}(\Psi(\tau), \nabla_{\mathbf{J}}\Psi(\tau)) \tag{1}$$

  We shall consider fitness $\Psi(\tau)$ that depends only on the steady-state distribution of all the units $P(\mathbf{x}|\mathbf{J}(\tau))$:

$$\Psi(\tau) = N^{-1} \int d^N x \, P(\mathbf{x}|\mathbf{J}(\tau)) \sum_{k=1}^{N} \phi_k(x_k) \, . \tag{2}$$

  where $\phi_k$ is some (non)linear function. The general framework that we develop in the next section, however, does not depend on this particular choice of fitness.

To make the presentation clear, from now on we specify the elements of $\mathbf{J}(\tau)$ as the interactions among $\mathbf{x}$'s components, e.g. $J_{kj}(\tau)$ is the influence of $j$ on $k$. We assume that positive and negative influences are the most basic forms of interactions among the system units, so $J_{kj}(\tau)$ can take binary value: $J_{kj}(\tau) = 1$ ($J_{kj}(\tau) = -1$) if unit $j$ activates (inhibits) the activity of unit $k$. This can be ensured for the adaptation of $\mathbf{J}(\tau)$ in (ii) by adopting a discrete-time update with the sign function as shown below. When taking the thermodynamic limit on fully-connected graphs, proper scaling of $J_{kj}$ will be considered (detailed information on scaling properties of these interactions is provided in Appendix A). In summary, the unit-coupling dynamics of (i) and (ii) with $h_{kj}(\tau)$ and $\Psi(\tau)$ given in Eqs. (1) and (2), respective, are closed and obey:

$$\begin{cases} \left(\dfrac{\partial}{\partial t} + 1\right) x_k(t;\tau) = F\Big(\displaystyle\sum_j J_{kj}(\tau) x_j(t;\tau)\Big) + \xi_k(t;\tau) \\ J_{kj}(\tau+1) = \text{sign}\left[h_{kj}(\tau) + \beta^{-1}\tilde{\xi}_{kj}(\tau)\right] \end{cases} \tag{3}$$

where $J_{kk}(\tau) = 0$, $\forall\tau$, $F(\cdot)$ is a nonlinear function and $\xi_k(t;\tau)$ is a mean-zero white noise with $\langle \xi_k(t;\tau)\xi_j(t';\tau')\rangle = \sigma^2\delta_{kj}\delta_{\tau\tau'}\delta(t-t')$. Due to the non-linearity of $F(\cdot)$ the process $\mathbf{x}(t;\tau)$ in Eq. (3) does not obey detailed balance (not even for symmetric $\mathbf{J}(\tau)$ [31]). Here a parameter $\beta$ accounts for the strength of stochastic effects in the $\mathbf{J}$'s dynamics induced by a set of 'threshold noises' $\{\tilde{\xi}_{kj}\}$. $\{\tilde{\xi}_{kj}\}$ are independent and identically distributed random variables drawn from a distribution $p(\tilde{\xi})$ that fulfils $p(\tilde{\xi}) = p(-\tilde{\xi})$ [24]. Without the feedback $h_{kj}(\tau)$, $J_{kj}(\tau+1)$ is a mean-zero random variable due to the term $\beta^{-1}\tilde{\xi}_{kj}(\tau)$.

We shall implement the dynamics of Eq. (3) in a nested fashion. Specifically, at generation $\tau$, we integrate the first equation under the present interaction matrix $\mathbf{J}(\tau)$ until it reaches a steady-state solution $\mathbf{x}(\infty;\tau)$, and then we use this $\mathbf{x}(\infty;\tau)$ to

compute the feedback fields $h_{kj}(\tau)$ via Eqs. (1)–(2); once all the $h_{kj}(\tau)$ are known, we update $\mathbf{J}(\tau)$ to $\mathbf{J}(\tau+1)$ according to the second equation of Eq. (3). This procedure is again repeated at generation $\tau+1$. Underlying this implementation of Eq. (3) is our assumption that the duration of one generation must be long enough for the subsystem of fast variables to relax to its asymptotic attractor, which is a steady state defined by the interaction matrix at that generation. This is indeed the main assumption of many *genetic algorithms* [74–79] and has also been used recently in the evolutionary dynamics of ecological communities [17].

## 3. Adaptive Dynamical Mean-Field Theory (ADMFT)

The averaged coupling $\hat{\mu}(\tau)$ is defined as

$$\hat{\mu}(\tau) := \frac{1}{N(N-1)} \sum_{(i,j)} J_{ij}(\tau)\,. \tag{4}$$

In Appendix A, we derive

$$\hat{\mu}(\tau+1) = \tanh(\beta h(\tau))\,. \tag{5}$$

Moreover, to quantify the variance and the symmetry level of the couplings, we introduce parameters $\hat{\alpha}$ and $\nu$, respectively. The latter can be defined, for any pair of $i$ and $j$, as the covariance between $J_{ij}$ and $J_{ji}$‡:

$$\nu := \Big\langle [J_{ij} - \langle J_{ij}\rangle][J_{ji} - \langle J_{ji}\rangle] \Big\rangle \tag{6}$$

In general, $\nu \in [-1,1]$. In particular, the specific cases $\nu = \{1, 0, -1\}$ correspond to fully symmetric, asymmetric and and antisymmetric interactions. To simplify the analytical treatment we fix $\nu$ as a generation-independent constant, i.e. $\nu(\tau) = \nu(\tau+1) = \nu$, $\forall\,\tau$, by imposing an appropriate constraint on the relation between $h_{ij}(\tau)$ and $h_{ji}(\tau)$. While for binary couplings $\hat{\alpha}(\tau) = 1 - \mu^2(\tau)$, in deriving the ADMFT we shall consider the more general case in which $J_{ij}(\tau)$ are not necessarily binary, but have continuous values as detailed in Eq. (0.10). In this case, $\hat{\alpha}(\tau)$ is independent of $\hat{\mu}(\tau)$. Using the path integral formalism [21–25], we can derive from a saddle-point approximation that becomes exact as $N \to \infty$, an effective process of a single unit $x(t;\tau)$ whose distribution of trajectories is [see Appendix A for detailed derivation of Eqs. (7*a*)-(9)]

$$\mathcal{P}(\{x\}|\{\eta\},\{\xi\}) = \prod_{\tau=0}^{T_{\max}-1} \int dt \Big\langle \delta(\partial_t x + x - F(\kappa) - \xi) \Big\rangle_* \tag{7a}$$

$$\kappa(t;\tau) = \hat{\mu}(\tau) m(t;\tau) + \hat{\alpha}\nu \int_0^t dt' G(t,t';\tau) x(t';\tau) + \eta(t;\tau) \tag{7b}$$

where $\langle\cdot\rangle_*$ denotes the average taken wrt this effective measure $\mathcal{P}(\{x\}|\{\eta\},\{\xi\})$. The averaged activity of this variable $m(t;\tau)$, its autocorrelation $C(t,t';\tau)$, its response

‡ This average is taken wrt the distribution $\tilde{P}(\mathbf{J}(\tau))$ over the ensemble of networks at generation $\tau$.

function $G(t,t';\tau)$ and the effective noise $\eta(t';\tau)$, all are self-consistently defined from:

$$\begin{aligned} m(t;\tau) &:= \langle x(t;\tau)\rangle_* && (8a)\\ C(t,t';\tau) &:= \langle x(t;\tau)x(t';\tau)\rangle_* && (8b)\\ G(t,t';\tau) &:= \left\langle \frac{\delta x(t;\tau)}{\delta \eta(t';\tau)} \right\rangle_* , && (8c)\\ \langle \eta(t;\tau)\eta(t';\tau)\rangle_* &:= \hat{\alpha}\, C(t,t';\tau) && (8d) \end{aligned}$$

When $\nu = 0$, the term associated with the response function drops out, leading to dynamics similar to that of asymmetric networks [50]. If $\hat{\alpha} = 0$, then the system becomes effectively equilibrium. Once coupling the effective process of Eq. (7$a$) to the evolution of $\hat{\mu}(\tau)$, we arrive at the following *unit-coupling* system:

$$\begin{cases} \hat{\mu}(\tau+1) = \tanh(\beta h(\tau)) \\ \partial_t x(t;\tau) = -x(t;\tau) + F(\kappa(t;\tau)) + \xi(t;\tau) \end{cases} \tag{9}$$

In the $N \to \infty$ limit, the dynamics described by Eq. (9) yield exactly the same statistics as those of the original dynamics in Eq. (3), in an analogous manner to what has been rigorously proven for random neural networks [82]. Numerical solutions to the second equation of Eq. (9) can be found using Monte Carlo methods [83–86]. Once its asymptotic attractors $x(\infty;\tau)$ has been found, we can update $\hat{\mu}(\tau)$ using the first equation of Eq. (9), provided that a specific form of $h_{kj}(\tau)$ in Eq. (1) as a function of $\Psi(\tau)$ (and hence $x(\infty;\tau)$) is given. This set of closed dynamical equations in Eq. (9) then can be iterated over many generations, in a nested manner, i.e. in the same way as the aforementioned implementation of the full dynamics Eq. (3).

## 4. The evolution of genotype-phenotype relationship

### *4.1. Model description*

We now provide an illustration of how the general framework of Eq. (9) applies to the genotype-phenotype map – a fundamental component of evolutionary theories [73]. Specifically, we consider the model in [74] that represents the genotype by $\mathbf{J}$ and the phenotype by $\mathbf{x}$. In this model $\mathbf{J}$ is a gene regulatory network that has an all-to-all topology, while $\mathbf{x}$ are the gene expression levels of $N$ different genes on this network. Motivated by the observation that the functionality of genetic networks is often determined only by a fraction of genes, the model introduces the concept of *target* genes (units). Their set is denoted by $\mathcal{T}$ and their number – by $N_t$. All other units are called non-target, their set is denoted by $\mathcal{O}$ and their number – by $N_o = N - N_t$. A fixed ratio of target to non-target is considered in the thermodynamic limit:

$$\alpha = N_o/N_t = O(1)\,, \quad N \to \infty \tag{10}$$

The model assumes that non-target genes do not contribute to the *absolute* fitness $\Psi(\tau)$

and hence defines $\Psi(\tau)$ just as an average of $x_i$, for $i \in \mathcal{T}$ §

$$\Psi(\tau) = N_t^{-1} \int d^{N_t}x \, P(x_{i\in\mathcal{T}}|\mathbf{J}(\tau)) \Big( \sum_{i\in\mathcal{T}} x_i \Big) \,, \tag{11}$$

This definition of fitness reflects the fact that the overall expression of some target (important) genes is a good proxy for the organism's chance of survival and reproduction. In the context of neural networks with learning, such as FORCE learning [87], the readout neurons play the role of target units as their states determine the network performance in terms of some loss functions. For eco-evolutionary systems, target units represent the keystone (focal) species, i.e. those that have strong impacts on the communities in which they reside [15]. The partition into the target and non-target units results in 4 different types of interactions, namely, $J_{ij}^{(tt)}$ for $i \in \mathcal{T}$ and $j \in \mathcal{T}$; $J_{ij}^{(oo)}$ for $i \in \mathcal{O}$ and $j \in \mathcal{O}$; $J_{ij}^{(to)}$ for $i \in \mathcal{T}$ and $j \in \mathcal{O}$; $J_{ij}^{(ot)}$ for $i \in \mathcal{O}$ and $j \in \mathcal{T}$. A similar division into two groups has been formulated for spin-systems, where, under certain conditions, target spins are in an effective equilibrium, while non-target spins remain non-equilibrium [88].

In [74] the gene expression dynamics on continuous time $t$ are given by

$$\left( \frac{\partial}{\partial t} + 1 \right) x_k(t;\tau) = \tanh\Big( \sum_j J_{kj}(\tau) x_j(t;\tau) \Big) + \xi_k(t;\tau) \,, \tag{12}$$

and, also in a nested manner as explained in the previous section, at each generation $\tau$, the network adaptation follows a genetic algorithm that, (i) considers a population of individuals with different networks (genotypes) and (ii) keeps only a fraction of individuals carrying *advanced* genotypes $\mathbf{J}^*(\tau)$ for the production of offsprings at generation $\tau + 1$. Here, according to Fisher's fundamental theorem [89], $\mathbf{J}^*(\tau)$ are those that can give rise to a *relative* fitness $\Psi_*(\tau)/\sqrt{\hat{q}(\tau)}$ larger than the population average $\langle \Psi(\tau) \rangle_{\tilde{P}(\mathbf{J}(\tau))}$ [taken wrt the distribution $\tilde{P}(\mathbf{J}(\tau))$ over the ensemble of genetic networks at generation $\tau$], where $\Psi_*(\tau)$ is given in Eq. (11) and

$$\hat{q}(\tau) \;\; := \lim_{t\to\infty} \frac{1}{N_t^2} \int d^{N_t}x \, P_t(x_{k\in\mathcal{T}}|\mathbf{J}(\tau)) \sum_{k\in\mathcal{T}} x_k^2(t;\tau) \tag{13a}$$

$$x_k(\tau) := \lim_{t\to\infty} \int d^{N_t}x \, P_t(x_{k\in\mathcal{T}}|\mathbf{J}(\tau)) x_k(t;\tau) \,. \tag{13b}$$

Following the definition of the absolute fitness in Eq. (11), we have the relation:

$$\Psi(\tau) = \frac{1}{N_t} \sum_{k\in\mathcal{T}} x_k(\tau) \,. \tag{14}$$

Therefore, such a selection process by the relative fitness (instead of the absolute one) can be ensured by multiplying $x_k(\tau)$ by a factor $1/\sqrt{\hat{q}(\tau)}$, for $k \in \mathcal{T}$, i.e., $x_k(\tau) \to x_k(\tau)/\sqrt{\hat{q}(\tau)}$. Taking into account the symmetry of the attractors of Eq. (12) with respect to $\mathbf{x} \to -\mathbf{x}$, instead of $\Psi(\tau)$, we shall use its square $\Psi^2(\tau)$ from now

§ This corresponds to the choice of $\phi(x_k) = x_k(1+\alpha)\mathbf{1}_\mathcal{T}(k)$, where $\mathbf{1}_\mathcal{T}(k)$ is the indicator function: $\mathbf{1}_\mathcal{T}(k)$ equals 1 if $k \in \mathcal{T}$, and zero if $k \in \mathcal{O}$.

on. By excluding the terms with $k = j$ from $\Psi^2(\tau)$ as they are only subleading correction of order $O(N_t^{-1})$, we arrive at

$$\Psi^2(\tau) = \frac{1}{N_t^2} \sum_{k \neq j \in \mathcal{T}} x_k(\tau) x_j(\tau) \tag{15}$$

*4.2. Reformulation of the model in terms of Equation (3)*

To apply our ADMFT framework we first need to specify how $\mathbf{h}(\tau)$ depends on $\Psi(\tau)$ in Eq. (1) (more precisely, on $\Psi^2(\tau)$ in connection with the aforementioned attractor symmetry). Due to the partition into target and non-target groups, the specific form of Eq. (1) depends on which groups ($\mathcal{T}$ or $\mathcal{O}$) that $i$ and $j$ belong to. Accordingly, we denote the feedback fields that act on $J_{ij}^{(tt)}$, $J_{ij}^{(oo)}$, $J_{ij}^{(to)}$, $J_{ij}^{(ot)}$ by $h_{ij}^{(tt)}$, $h_{ij}^{(oo)}$, $h_{ij}^{(to)}$ and $h_{ij}^{(ot)}$, respectively. Since in the model [74] target genes are those that directly contribute to the fitness $\Psi^2(\tau)$, we expect the feedback fields $h_{ij}^{(tt)}(\tau)$ can be cast into such forms that, once taking the average over all the pairs $(i, j) \in \mathcal{T}$, yields $\Psi^2(\tau)$ in Eq. (15) (i.e. $\Psi^2(\tau) = N_t^{-2} \sum_{k \neq j \in \mathcal{T}} h_{kj}^{(tt)}(\tau)$). This leads to

$$h_{kj}^{(tt)}(\tau) := x_k(\tau) x_j(\tau)\,, \quad (k, j) \in \mathcal{T}\,. \tag{16}$$

This particular form of $h_{kj}^{(tt)}(\tau)$ is definitely not the unique way to implement the feedback of phenotype on genotype, but it captures the basic fact that to achieve a given fitness value, these fields need to depend on phenotypic covariances. This form is also known as the Hebbian learning rule [90], making the genotype-phenotype map in [74] closely related to a Hopfield model with evolving patterns [91]. For large $\beta$ in Eq. (3), this learning rule results in symmetric interactions among target genes, i.e. $J_{kj}^{(tt)}(\tau) = J_{jk}^{(tt)}(\tau)$, $\forall\, \tau$. Taking into account the use of relative fitness in [74] [that corresponds to $x_k(\tau) \to x_k(\tau)/\sqrt{\hat{q}(\tau)}$, as mentioned above], we deduce from Eq. (16):

$$h_{kj}^{(tt)}(\tau) := \frac{x_k(\tau) x_j(\tau)}{\hat{q}(\tau)}\,, \quad (k, j) \in \mathcal{T}\,. \tag{17}$$

Similarly, we can derive the following expression for $h_{k\ell}^{(to)}(\tau)$ and $h_{\ell k}^{(ot)}(\tau)$ from the gradient of the fitness, see Appendix C for the details:

$$h_{k\ell}^{(to)}(\tau) = \left(1 - \frac{x_k^2(\tau)}{\hat{q}(\tau)}\right) \frac{\Psi(\tau)}{\hat{q}(\tau)} \Big( N_t^{-1} \sum_{j \in \mathcal{T}} J_{k\ell}^{(to)}(\tau) J_{\ell j}^{(ot)}(\tau) x_j(\tau) \Big)\,. \tag{18}$$

In summary, we reformulated the model [74] whose fitness is defined as in Eq. (14) in terms of Eq. (3) with $F(\cdot) = \tanh(\cdot)$; and $h_{kj}^{(tt)}(\tau)$ and $h_{k\ell}^{(to)}(\tau)$ given in Eqs. (17) and (18), respectively. In the $N \to \infty$ limit, this *closed* system admits an ADMFT description in terms of Eq. (9). In the following subsection, we will present result concerning the fitness in Eq. (14), for which the maximal $\Psi^2_{\max} = 1$ can be achieved if either $x_k = 1$ or $x_k = -1$, $\forall\, k \in \mathcal{T}$. The generalisation to those cases, where other patterns of gene expression levels are the optimal configurations that maximise other fitness functions shall be discussed in Appendix E.

*4.3. Fixed-point solutions for fully symmetric intergroup couplings $\nu = 1$*

For any pair of $k \in \mathcal{T}$ and $j \in \mathcal{O}$, the symmetry of *intergroup* couplings between the target and non-target group is controlled by the reciprocity parameter $\nu$ defined as:

$$\nu := \left\langle [J_{kj}^{(to)} - \langle J_{kj}^{(to)} \rangle][J_{jk}^{(ot)} - \langle J_{jk}^{(ot)} \rangle] \right\rangle \tag{19}$$

In this section we only consider the case of $\nu = 1$, i.e. $J_{kj}^{(to)} = J_{jk}^{(ot)}$. We distinguish the averaged intragroup and intergroup couplings by introducing

$$\hat{\mu}(\tau) := \frac{1}{N(N-1)} \sum_{(i,j)\in\mathcal{T}} J_{(ij)}^{(tt)}(\tau) \tag{20a}$$

$$\hat{\lambda}(\tau) := \frac{1}{N_t N_o} \sum_{i\in\mathcal{T}, j\in\mathcal{O}} J_{ij}^{(to)}(\tau) \tag{20b}$$

In Appendix A, we shall derive that for the model [74] the variance $\hat{\alpha}$ of $J_{ij}^{(tt)}$ equals to $\alpha$, i.e. $\hat{\alpha} = \alpha$. Assuming that at generation $\tau$ the target unit $x$'s dynamics reach a stationary state with time-translational symmetry, then $\forall s > 0$, we have

$$\lim_{t'\to\infty} C(t = t' + s, t', \tau) = C(s, \tau) \tag{21a}$$

$$\lim_{t'\to\infty} G(t = t' + s, t', \tau) = G(s, \tau) \tag{21b}$$

$$\lim_{t\to\infty} G(t, t', \tau) = 0\,, \forall\,(t', \tau) \tag{21c}$$

As well-known from previous studies on static fully-connected networks [38], a large value of $\alpha$ would drive the system towards chaotic attractor, thus breaking this assumption that the fast-time dynamics converge to a well-defined stationary state. We will elaborate on the restriction of $\alpha \leq 0.5$ for a fully-connected system to reach a stable fixed-point solutions in Appendix D. We define the integrated response $\hat{\chi}$ (that should remain finite for fixed-point solutions) as follows

$$\hat{\chi}(\tau) = \int_0^\infty ds\, G(s, \tau) \tag{22}$$

When $x(t;\tau)$ fluctuates around a $t$-independent average at a given generation $\tau$, the steady-state autocorrelation $C(s,\tau)$ reaches a positive (possibly) plateau value as $s \to \infty$: $\hat{q}(\tau) := \lim_{s\to\infty} C(s,\tau)$. After taking this limit, we take the limit $\tau \to \infty$ to identify fixed points of the unit-coupling system Eq. (9) that are denoted by

$$\mu = \lim_{\tau\to\infty} \hat{\mu}(\tau)\,, \quad x_* = \lim_{\tau,t\to\infty} x(t,\tau)\,, \quad \chi = \lim_{\tau\to\infty} \hat{\chi}(\tau)\,, \quad q = \lim_{\tau\to\infty} \hat{q}(\tau) \tag{23}$$

and, similarly, in this limit

$$\lambda = \lim_{\tau\to\infty} \hat{\lambda}(\tau)\,. \tag{24}$$

Fixed points of Eq. (9) then satisfy:

$$x_* = \tanh\left(\mu \langle x_* \rangle_{\eta,\xi} + \alpha\chi x_* + \eta\right) + \xi \tag{25}$$

This equation, for given distributions of $\eta$ and $\xi$, defines the distribution of $x_*$. As $x$ reaches a stationary state, so does $\eta$: $\lim_{t,\tau\to\infty} \eta(t;\tau) = \eta_0 = J_0\sqrt{q}\tilde{z}$, where $\eta_0$ is a

Gaussian random number with mean zero and variance $J_0^2 q$, $J_0 = \sqrt{\alpha}$ and $\tilde{z} \sim \mathcal{N}(0,1)$, the standard normal distribution. Likewise, at stationary, the effect of the white noise $\xi$ is equivalent to adding a static random number $\xi_0 = \sigma z$ to the dynamics specified by the first two terms of the last line in Eq. (9), for $z \sim \mathcal{N}(0,1)$. Taking these facts into account, we can rewrite Eq. (25) as

$$x_*(\tilde{z}, z) = \tanh\left(\mu m_\infty + \alpha\chi x_* + J_0\sqrt{q}\tilde{z}\right) + \sigma z\,. \tag{26}$$

where $m_\infty = \langle x_* \rangle_{z,\tilde{z}}$. In general, Eq. (26) can have multiple solutions [86]. We, however, assume that $x_*(\tilde{z}, z)$ does not depend on initial conditions, i.e. it is unique for any given realisation of $z$ and $\tilde{z}$. This allows us to obtain the stationary state of the genotype-phenotype system as a solution to the following set of self-consistency equations for the averages taken over the ensemble of fixed points:

$$\begin{cases} \mu = \tanh(\beta m_\infty^2/q) \\ m_\infty = \int_{-\infty}^{\infty} Dz \int_{-\infty}^{\infty} D\tilde{z}\, x_*(\tilde{z}, z) \\ q = \displaystyle\int_{-\infty}^{\infty} Dz \int_{-\infty}^{\infty} D\tilde{z} (x_*(\tilde{z}, z))^2 \\ \chi = \displaystyle\int_{-\infty}^{\infty} Dz \int_{-\infty}^{\infty} D\tilde{z}\, \frac{1 - f_0^2}{1 - \alpha\chi(1 - f_0^2)} \end{cases} \tag{27}$$

where $f_0 := \tanh\left(\mu m_\infty + \alpha\chi x_* + J_0\sqrt{q}\tilde{z}\right)$ and $Dz := dz e^{-z^2/2}/\sqrt{2\pi}$ is the Gaussian measure; $m_\infty$ is the average of $x_*$; $q$ is the average of $x_*^2$ ; $\chi$ is the integrated response of $x_*$; $\mu$ is the steady-state mean value of $J^{(tt)}$. Solutions are obtained as stable attractors of iterative dynamics started from an initial condition with sufficiently large $m_\infty$, $q$, $\chi$, $\mu$. Finally, the steady-state average $\lambda$ of $J^{(to)}$ can be computed from [see Appendix C for detailed derivation]:

$$\lambda = \tanh(\beta\lambda^2 m_\infty^2/q(1 - m_\infty^2/q)) \tag{28}$$

Note $\lambda = 0$ is always a solution of this equation. We are particularly interested in the transition between this fixed-point solution and the non-trivial solutions that can only emerge if $m_\infty \neq 0$. Therefore, it is important to know when the system relaxes to a steady state with non-zero activity, $m_\infty \neq 0$. To this end, we need to analyse the local stability of $\mathbf{x} = \mathbf{0}$ (i.e. $x_i = 0, \forall\, i \in \mathcal{T}$) that can be quantified using the largest eigenvalue $\Lambda_1$ of the Jacobian matrix at $\mathbf{x} = \mathbf{0}$ (i.e. by analysing the linearised subsystem of target units). We obtain $\Lambda_1$ from $\mu$ and $\lambda$ (see Appendix B for the derivation):

$$\Lambda_1 = \frac{\lambda^4 + \mu^2}{\mu} - 2 \tag{29}$$

This equation shows that $\Lambda_1$ depends on the steady-state behavior of the genotypic values: $\Lambda_1 = -2$ if $\mu = \lambda = 0$, $\Lambda_1 = -1$ if $\lambda = 0$ but $\mu = 1$ and $\Lambda_1 = 0$ if $\mu = \lambda = 1$. Therefore, apart from being a measure of stability, $\Lambda_1$ is also a good indicator of the genotype's states. In showing results, to separate the effect of external noise on the system behaviour from that of the self-sustained fluctuations, instead of $q$, we chose to plot the so-called intrinsic variance $q_0$, defined as

$$q_0 = q - \sigma^2\,. \tag{30}$$

**Table 1.** The model steady-state phases and their features

| Phases | $m_\infty$ | $\lambda$ | $\mu$ |
|---|---|---|---|
| robust | $m_\infty > 0$ | $\lambda > 0$ | $\mu > 0$ |
| non-robust, | $m_\infty > 0$ | $\lambda = 0$ | $\mu > 0$ |
| para-attractor | $m_\infty = 0$ | $\lambda = 0$ | $\mu = 0$ |

We now present results for $\alpha = 0.5$, similar behaviour is observed for other $\alpha > 0.07$ (see Fig. 2 **(B)** below, where $\alpha = 0.07$ is the minimal value of $\alpha$ for which solution with non-zero $\lambda$ can be found). The particular value of $\alpha = 0.5$ is chosen for the sake of presentation only as in this case one can find a robust region over the largest window of noise. The steady-state behaviors and, in particular, the emergence of robustness can be quantified by the dependence of $\Lambda_1$ computed from Eq. (29) on $T_J = 1/\beta$ and $\sigma$. In figure 1 **(A)** we find three distinct regions corresponding to different values of $\Lambda_1$, namely, $\Lambda_1 = -2$, $\Lambda_1 \simeq -1$ and $\Lambda_1$ close to zero. This is because of a non-monotonic behaviour of $\lambda$ as a function of $\sigma$ at high enough selection pressure. Specifically, upon increasing $\sigma$, $\lambda$ first increases from zero to a plateau value close to 1 and then drops to zero again as observed in figure 1 **(B)** for $\beta = 10$. This figure demonstrates three possible solutions of Eqs. (27)-(28): the first solution is $\lambda = 0$ and $m_\infty, \mu > 0$, the second corresponds to $m_\infty, \lambda, \mu > 0$ and the last – to $m_\infty = \lambda = \mu = 0$. We call them non-robust, robust and para-attractor, respectively. Their characteristics are given in Table 1. Note that due to a positive feedback between $\mu$ and $m_\infty$ that is imposed by the first of Eq. (27), $m_\infty$ and $\mu$ always behave similarly.

Figures 1 **(C)**–**(F)** show detailed behaviours of the order parameters as functions of $T_J = 1/\beta$ and $\sigma$. While both $m_\infty$ and $\mu$ only undergo an ordered/disordered transition as $\sigma$ is increased at fixed $T_J = \beta^{-1}$, below some $T_J^{(\lambda)}$, the robust phase with $\lambda = 1$ emerges within an intermediate range $\sigma \in [\sigma_c^{(1)}(T_J), \sigma_c^{(2)}(T_J)]$. Importantly, $\sigma_c^{(1)}(T_J) > 0$. If noise is increased beyond $\sigma_c^{(2)}(T_J)$, a transition from the robust to para-attractor phase occurs due to the dominant effect of noise. In the para-attractor region, both phenotypic and genotypic values become zero, indicating the unique state is neither robust nor functional. These results are in agreement with the model [74] where a loss of the robustness to noise that evolved at intermediate noise is observed with decreasing noise.

Next, we plot the phase diagrams in terms of $\alpha$ and $\sigma$ in Figure 2. Here again, three phases are observed. At a given noise level $\sigma$, the system reaches a state with non-zero fitness $m_\infty > 0$ as $\alpha > \alpha_c(\sigma)$. Note that $\alpha_c(\sigma)$ grows with $\sigma$, indicating the necessity of having a sufficient fraction of non-target genes to achieve high-fitness values at large noise. This enhancement of robustness with the help of non-target genes is also observed in the behavior of the genotypic variable $\lambda$, where the robust region with $\lambda > 0$ broadens with increasing $\alpha$. Note that due to our choice of fully-connected networks, the fixed-point assumption only holds for $\alpha \leq 0.5$, as can be shown via a local stability

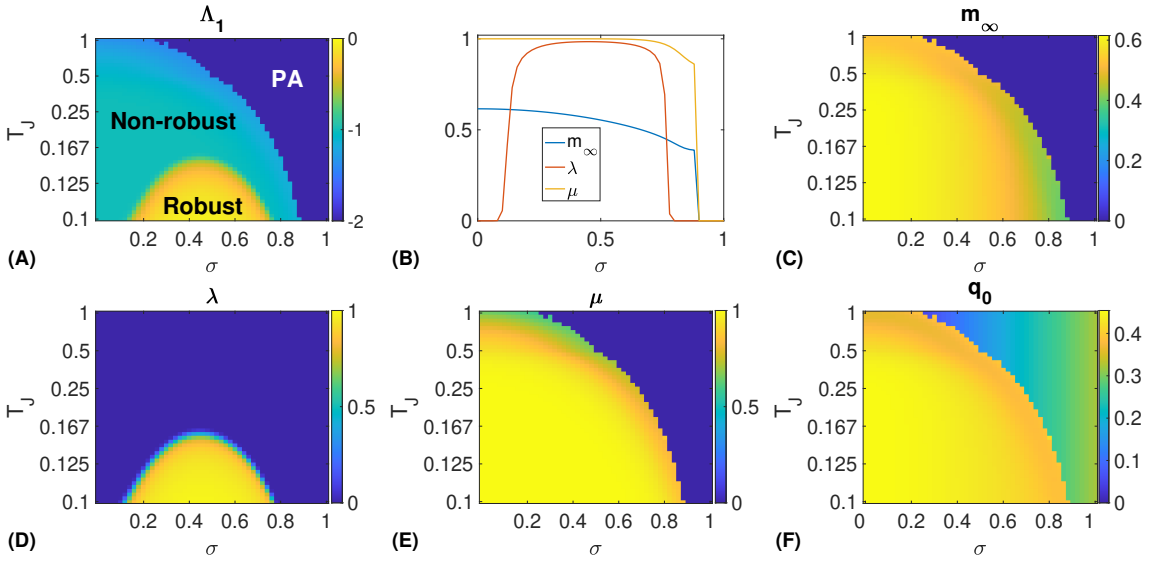


**Figure 1.** **(A)** The largest eigenvalue $\Lambda_1$ as function of $\sigma$ and $T_J = \beta^{-1}$. Here $\Lambda_1$ is computed by Eq. (0.56) of the linearised subsystem of target genes at $\mathbf{x} = \mathbf{0}$. Three distinct values of $\Lambda_1$ indicate three different phases that we call robust, non-robust, and para-attractor. **(B)** Order parameters as function of $\sigma$ for $\beta = 10$. In the non-robust phase $m_\infty, \mu > 0$ but $\lambda = 0$, while in the robust phase $m_\infty, \lambda, \mu > 0$. In the para-attractor phase **(PA)** $m_\infty = \lambda = \mu = 0$. **(C)** Averaged activity of target genes $m$. **(D)** The average of target vs non-target coupling $\lambda$. **(E)** The average of target vs target coupling $\mu$. **(F)** The intrinsic variance in the target gene's activity $q_0$. In all panels $\alpha = 0.5$ and $\nu = 1$.

analysis given in Appendix D. We expect that in realistic genetic networks that are *sparse*, such a restriction might not be necessary. Therefore, one can consider $\alpha > 0.5$ and even $\alpha \gg 1$ in that case. Finally, quantifying the system behaviour in terms of $\Lambda_1$ in Figure 2 **E**, we can see the existence of robustness phase corresponding to the values of $\Lambda_1$ very close to zero.

## 5. Discussion

In this paper, we constructed the ADMFT framework that is applicable to a wide range of adaptive systems, in which slow adaptation of one type of degrees of freedom occurs in response to fast changes in the state of the other. We demonstrated our approach within the context of genotype-phenotype evolution, where we found a transition from robust- to nonrobust phase with decreasing noise in networks of fully-symmetric intergroup couplings. This happens due to a trade-off between phenotype and genotype that leads to a strong non-monotonic behavior of the genotypic value $\lambda$. A comprehensive picture of the model behavior in the full range of $\nu \in [-1, 1]$ with possible limit-cycle and chaotic behaviours will be addressed in future work. At the moment, we can speculate that as the intergroup interactions become more non-reciprocal, such non-monotonicity is reduced, resulting in a robust phase at low noise. In this case, however,

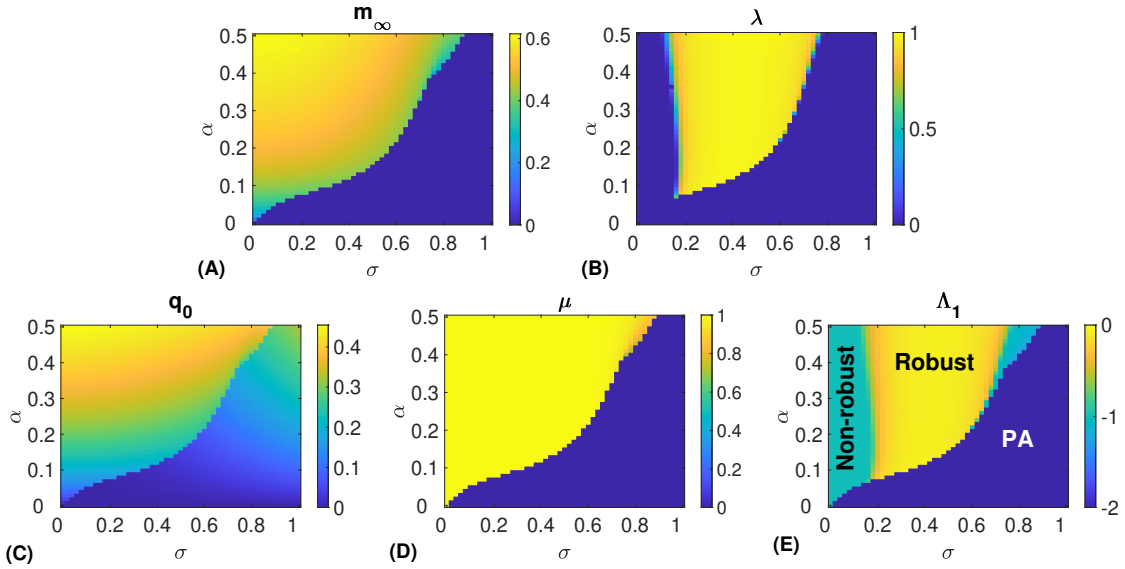


**Figure 2.** Order parameters as function of $\alpha = N_o/N_t$ and $\sigma$ at $\nu = 1$ and $\beta = 10$. **(A)** Averaged activity of target genes $m_\infty$. **(B)** The average of target vs non-target coupling $\lambda$. **(C)** The intrinsic variance in the activity of target genes $q_0$. **(D)** The average of target vs target coupling $\mu$. **(E)** $\Lambda_1$ as function of $\alpha = N_o/N_t$ and $\sigma$. The value of $\Lambda_1$ is used to distinguish the robust, non-robust, and para-attractor phases.

the system achieves a lower averaged activity than what is obtained for reciprocal intergroup interactions. In the fully symmetric case, we also found the emergence of an outlier eigenvalue from the genotype-phenotype feedback. Such eigenmode controls the long-time behavior of the gene-expression dynamics, which is consistent with the dimensional reduction of phenotypic dynamics resulting from genetic algorithms [92–94]. For a given environment, our work suggests the loss of robustness at low noise and confirms the beneficial role of noise in the evolution of robustness as previously discussed in numerical studies [74]. In particular, in the robust phase where $m_\infty > 0$, a proportionality between the response to environmental stochasticity and the response to mutation $G_\mu(\tau, \tau') \equiv \partial\langle x(\tau)\rangle/\partial\hat{\mu}(\tau')$ can be expected to arise as a consequence of the Hebbian-learning in Eq. (17). Such a proportionality shall imply a correlation between phenotypic changes due to genetic variation and those in response to environmental perturbations, as suggested by [68, 74, 76, 95–101]. It would be interesting to extend our approach to a fluctuating environment where a nonlinear dependence of the fitness on the expression patterns due to a trade-off between the cost and benefit is expected [102]. In comparison to other recent extensions of DMFT for neural dynamics that focus on local learning rules, such as activity-dependent plasticity [103] and pattern-based learning [104], our approach derives the adaptation rules for the coupling matrix $\mathbf{J}$ from a global fitness function. Nevertheless, persistent fluctuations are observed in the robust phase of the presented GRN model are similar to chaos with retrieval in [104]. In this regard, the ADMFT suggests the relevance of noise to shape robust memory by maintaining a finite overlap with the stored patterns.

**Acknowledgments**

We acknowledge support from Novo Nordisk Foundation (0065542) and would like to thank the anonymous reviewers, Albert Alonso, David Saad, Edo Kussell, Fernando Metz, Richardo Rao, Tarek Tohme and Valentina Ros for helpful comments.

**Appendices**

*A. Derivation of the effective dynamics in Equation (9)*

We use an integral representation of the probability $P(\xi)$ of the mean zero white noise $\xi(t)$:

$$P(\xi) \sim \exp\left[-\frac{1}{2}\int dtdt' \xi(t)[C_\xi(t,t')]^{-1}\xi(t')\right], \qquad C_\xi(t,t') = \langle \xi(t)\xi(t')\rangle = \sigma^2\delta(t-t')$$

to rewrite any SDE of the form $\partial_t x = \mathcal{F}(t) + \xi(t)$ as [$i$ is the imaginary unit $i^2 = -1$]

$$1 = \int DxD\hat{x} \exp\left[-\frac{1}{2}\int dtdt'\, \hat{x}(t)C_\xi(t,t')\hat{x}(t') + i\int dt\, \hat{x}(t)[\partial_t x - \mathcal{F}(t)]\right] \quad (.1)$$

Let $\langle\cdot\rangle$ denote the average taken wrt the measure $\mathbb{P}(\{\mathbf{x}\}_{0,T_{\max}})$ – the distribution over an ensemble of trajectories of $\mathbf{x}(t;\tau)$ for $t \in [0,\infty)$ over $T_{\max}$ generations:

$$\mathbb{P}(\{\mathbf{x}\}_{0,T_{\max}}) := \mathbb{P}\Big[\{\mathbf{x}([0,t_f],\tau=0)\},\{\mathbf{x}([0,t_f],\tau=1)\},\cdots,\{\mathbf{x}([0,t_f],\tau=T_{\max})\}\Big]_{t_f\to\infty}.$$

Since at each generation $\tau \in \{0,1,\cdots,T_{\max}\}$, $\mathbf{x}(t;\tau)$ obey the first of Eq. (3), we can represent $\mathbb{P}(\{\mathbf{x}\}_{0,T_{\max}})$ by rewriting the dynamics of $\mathbf{x}$ in the presence of an external field $\boldsymbol{\theta}(t;\tau)$ with the help of Eq. (.1) as an identity:

$$1 = \int D[x\hat{x}f\hat{f}] \exp\Big\{i\sum_{\tau=0}^{T_{\max}}\sum_{k=1}^{N}\int dt\Big[S_k^{(0)} - I_k\Big]\Big\} \quad (0.2a)$$

$$S_k^{(0)}[x,\hat{x},f,\hat{f}] = \hat{x}_k\Big[(\partial_t+1)x_k - F(f_k)\Big] + \hat{f}_k[f_k - \theta_k] + i\sigma^2\hat{x}_k^2/2 \quad (0.2b)$$

$$I_k[x,\hat{f}] = \sum_{j=1}^{N} J_{kj}(\tau)\hat{f}_k(t;\tau)x_j(t;\tau) \quad (0.2c)$$

where $D[x\hat{x}f\hat{f}] := \prod_{n=0}^{t_f/\Delta t}\prod_{k=1}^{N}\prod_{\tau=0}^{T_{\max}} Dx_k(n\Delta t,\tau)D\hat{x}_k(n\Delta t,\tau)Df_k(n\Delta t,\tau)D\hat{f}_k(n\Delta t,\tau)$, for $\Delta t \to 0$, denotes the functional measure over all possible paths. The moment generating functional of $\mathbb{P}(\{\mathbf{x}\}_{0,T_{\max}})$ with $\{\boldsymbol{\psi}(t;\tau)\}$ for $\tau = 0,\cdots,T_{\max}$ is:

$$Z[\boldsymbol{\psi}] = \left\langle \exp\left\{ i\sum_{\tau=0}^{T_{\max}}\sum_{k=1}^{N}\int dt\, \psi_k(t;\tau)x_k(t;\tau)\right\}\right\rangle \quad (0.3)$$

When the distribution of noise is $p(\tilde{\xi}) = [1-\tanh^2(\tilde{\xi})]/2$, we can rewrite the second dynamics in Eq. (3) as a master equation for the distribution $\tilde{P}(\mathbf{J}(\tau))$ of configurations $\mathbf{J}(\tau)$ [24]:

$$\tilde{P}(\mathbf{J}(\tau+1)) = \sum_{\{\mathbf{J}(\tau)\}}\tilde{P}(\mathbf{J}(\tau))\prod_{k\neq j}\frac{e^{\beta h_{kj}(\tau)J_{kj}(\tau+1)}}{2\cosh[\beta h_{kj}(\tau)]} \quad (0.4)$$

This means the actual choice of value of $J_{kj}(\tau+1)$ is probabilistic and drawn from a distribution over two possibilities, namely, $J_{kj}(\tau+1)=1$ and $J_{kj}(\tau+1)=-1$, with probability $\propto \exp\{\beta h_{kj}(\tau)J_{kj}(\tau+1)\}$. Without loss of generality, we take $\tilde{P}(\mathbf{J}(0))$ as a uniform distribution. In order to separate the feedback fields at any generation, let us introduce $\tilde{h}_{kj}(\tau)$ as the value that $h_{kj}(\tau)$ admits at the $\tau$-th generation. This can be done by inserting the following identity to Eq. (0.4):

$$1=\int D[h\hat{h}]\exp\Big\{i\sum_{\tau=0}^{T_{\max}-1}\sum_{k,j}\hat{h}_{kj}(\tau)[h_{kj}(\tau)-\tilde{h}_{kj}(\tau)]\Big\} \tag{0.5}$$

For example, in the case of the genotype-phenotype model in Section 4, $\tilde{h}_{kj}(\tau)$ is given by Equation (16) if $(k,j)\in\mathcal{T}$ and Equation (18) if $k\in\mathcal{T}$ and $j\in\mathcal{O}$. As seen from that example, in general, $\tilde{h}_{kj}(\tau)$ depends on the joint distribution of the fields $\mathbf{x}$ and $\mathbf{J}$.

Let $\mathbb{E}[\cdot]$ denote the ensemble average wrt the joint distribution of trajectories in the combined space of $\mathbf{x}$ and $\mathbf{J}$. This measure can be obtained by inverse Fourier transform of the moment-generating functional $Z[\boldsymbol{\psi},\boldsymbol{\Psi}]$, obtained by plugging Eqs. (0.2$a$)-(0.5), altogether into Eq. (0.3):

$$Z[\boldsymbol{\psi},\boldsymbol{\Psi}]=\int D[x\hat{x}f\hat{f}h\hat{h}]\sum_{\{\mathbf{J}(0)\}}\cdots\sum_{\{\mathbf{J}(T_{\max})\}}e^{\mathcal{L}} \tag{0.6$a$}$$

$$\mathcal{L}=i\sum_{\tau=0}^{T_{\max}}\sum_{k=1}^{N}\int dt\Big[S_k^{(\psi)}-I_k\Big]+\sum_{k\neq j}[I_{kj}+J_{kj}(0)\Psi_{kj}(0)] \tag{0.6$b$}$$

where

$$S_k^{(\psi)}=S_k^{(0)}+\psi_k(t;\tau)x_k(t;\tau)$$

$$L_{kj}(\tau)=i\hat{h}_{kj}(\tau)[h_{kj}(\tau)-\tilde{h}_{kj}(\tau)]-\ln\left(2\cosh\left[\beta h_{kj}(\tau)\right]\right)$$

$$I_{kj}=\sum_{\tau=0}^{T_{\max}-1}\Big\{L_{kj}(\tau)+J_{kj}(\tau+1)[\Psi_{kj}(\tau+1)+\beta h_{kj}(\tau)]\Big\}$$

Denoting the $J$-independent and $J$-dependent part of $\mathcal{L}$ by $\mathcal{L}_0$ and $\mathcal{L}_J:=\sum_{k\neq j}\mathcal{L}_{kj}^{(J)}$, respectively, the exponential in $Z[\boldsymbol{\psi},\boldsymbol{\Psi}]$ can be written as:

$$\mathcal{L}=\mathcal{L}_0+\mathcal{L}_J \tag{0.7$a$}$$

$$\mathcal{L}_0=i\sum_{\tau=0}^{T_{\max}}\sum_{k}\int dtS_k^{(\psi)}(t;\tau)+\sum_{\tau=0}^{T_{\max}-1}\sum_{k\neq j}L_{kj}(\tau) \tag{0.7$b$}$$

$$\begin{aligned}\mathcal{L}_{kj}^{(J)}=&+J_{kj}(0)\Big[\Psi_{kj}(0)-i\int dt\hat{f}_k(t,0)x_j(t,0)\Big]\\&+\sum_{\tau=1}^{T_{\max}}J_{kj}(\tau)\Big\{\Psi_{kj}(\tau)+\beta h_{kj}(\tau-1)-i\int dt\hat{f}_k(t;\tau)x_j(t,\tau)\Big\}\end{aligned} \tag{0.7$c$}$$

In the following we shall present a derivation for ADMFT in the general setting of $N$ fully-connected units with 2 groups of units: target ($i\in\mathcal{T}$) and non-target ($i\in\mathcal{O}$). From now on by $\mathbf{x}$ we mean only $x_i$, for $i\in\mathcal{T}$. To avoid confusion, the state vector of

all non-target units $x_i$, for $i \in \mathcal{O}$, is denoted by $\mathbf{y}$. The partition into the target and non-target units results in 4 different types of interactions, namely, $J_{ij}^{(tt)}$ for $i \in \mathcal{T}$ and $j \in \mathcal{T}$; $J_{ij}^{(oo)}$ for $i \in \mathcal{O}$ and $j \in \mathcal{O}$; $J_{ij}^{(to)}$ for $i \in \mathcal{T}$ and $j \in \mathcal{O}$; $J_{ij}^{(ot)}$ for $i \in \mathcal{O}$ and $j \in \mathcal{T}$. We have

$$\frac{\partial}{\partial t}x_k = -x_k + F\Big(\sum_{j=1,j\neq k}^{N_t} J_{kj}^{(tt)}(\tau)x_j + \sum_{\ell=1}^{N_o} J_{k\ell}^{(to)}(\tau)y_\ell\Big) + \xi_k\,. \tag{0.8}$$

The presence of a non-target unit $\ell \in \mathcal{O}$ with its dynamics

$$\frac{\partial}{\partial t}y_k = -y_k + F\Big(\sum_{j=1,j\neq k}^{N_o} J_{kj}^{(oo)}(\tau)y_j + \sum_{\ell=1}^{N_t} J_{k\ell}^{(ot)}(\tau)x_\ell\Big) + \xi_k\,, \tag{0.9}$$

modifies the dynamic of those target units $k \in \mathcal{T}$ connected to it. Therefore, to derive a closed set of dynamical equations for the subset of target variables exclusively, we consider the following ansatz for the effective interaction between any pair of target variables $k \in \mathcal{T}$ and $j \in \mathcal{T}$:

$$J_{kj}^{(e)}(\tau) = J_{kj}^{(tt)}(\tau) + \Delta J_{kj}(\tau) \tag{0.10}$$

where, apart from their direct interaction $J_{kj}^{(tt)}(\tau)$ ($J_{kj}^{(tt)}(\tau) \in \{-1, 1\}$), $\Delta J_{kj}(\tau)$ is the part of the effective coupling that is induced by all their common *non-target* neighbors $\ell \in \mathcal{O}$ via $J_{k\ell}^{(to)}(\tau)$ and $J_{\ell j}^{(ot)}(\tau)$. As we consider fully-connected networks, if $N_t \to \infty$, the couplings $J_{kj}^{(e)}$ should be rescaled by $1/\sqrt{N_t}$ to ensure a sensible thermodynamic limit. However, due to a low-rank structure that can emerge from the adaptation dynamics of $J_{kj}^{(tt)}$, a proper scaling of $J_{kj}^{(e)}$ is $J_{ij}^{(e)} \to J_{ij}^{(e)}/N_t$. To determine $\Delta J_{kj}(\tau)$ in Eq. (0.10), we assume that the steady state solutions of the following *approximate* dynamics for the subset of target units:

$$\frac{\partial}{\partial t}x_k(t;\tau) = -x_k(t;\tau) + F\Big(\sum_{j=1,j\neq k}^{N_t} J_{kj}^{(e)}(\tau)x_j(t;\tau)\Big) + \xi_k(t;\tau) \tag{0.11}$$

is the same as those of the original dynamics Eq. (0.8). As at generation $\tau$, both $\mathbf{x}$ and $\mathbf{y}$ reach their corresponding attractors defined component-wise from Eqs. (0.8)-(0.9) as

$$x_k(\tau) := \lim_{t\to\infty} x_k(t;\tau) = F\Big(\sum_{j\in\mathcal{T}} J_{kj}^{(tt)}(\tau)x_j(\tau) + \sum_{\ell\in\mathcal{O}} J_{k\ell}^{(to)}(\tau)y_\ell(\tau)\Big) \tag{0.12}$$

$$y_\ell(\tau) = F\Big(\zeta_\ell^{(o)} + \zeta_\ell^{(t)}\Big)\,, \quad \zeta_\ell^{(o)} \equiv \sum_{\ell'\in\mathcal{O}} J_{\ell\ell'}^{(oo)}(\tau)y_{\ell'}(\tau)\,, \quad \zeta_\ell^{(t)} \equiv \sum_{j\in\mathcal{T}} J_{\ell j}^{(ot)}(\tau)x_j(\tau)$$

where $\zeta_\ell^{(o)}$ and $\zeta_\ell^{(t)}$ are the contributions of $\ell' \in \mathcal{O}$ and $j \in \mathcal{T}$, respectively, to $y_\ell(\tau)$. If there is no feedback of phenotype on the couplings among the non-target units, then $h_{\ell\ell'}^{(oo)}(\tau) = 0$, $\forall\tau$, and as a result, $J_{\ell\ell'}^{(oo)}$ remains random over generations. Therefore, we have $\zeta_\ell^{(o)} \ll \zeta_\ell^{(t)} \ll 1$, and $y_\ell(\tau) \simeq \zeta_\ell^{(t)}$. Substituting this $y_\ell(\tau)$ into Eq. (0.12), we get the condition for the asymptotic attractor of Eq. (0.11) equals to that of Eq. (0.8) $x_k(\tau) \simeq F(\sum_{j\in\mathcal{T}}[J_{kj}^{(tt)}(\tau) + \Delta J_{kj}(\tau)]x_j(\tau))$, if and only if

$$\Delta J_{kj}(\tau) = \sum_{\ell\in\mathcal{O}} J_{k\ell}^{(to)}(\tau)J_{\ell j}^{(ot)}(\tau)\,. \tag{0.13}$$

From now on, we shall compute $Z[\boldsymbol{\psi}, \boldsymbol{\Psi}]$, for the above dynamics of Eq. (0.11). This means that we need to insert $J_{kj}(\tau) = J_{kj}^{(e)}(\tau)$ (with $\Delta J_{kj}$ from Eq. (0.13)) into the expression of $I_k[x, \hat{f}]$ in Equation (0.2$c$). We define the order parameters:

$$w_{kj}(\tau) = \frac{1}{N_t} \int dt \hat{f}_k(t;\tau) x_j(t;\tau) \tag{0.14a}$$

$$m(t;\tau) = \frac{1}{N_t} \sum_{j\in\mathcal{T}} x_j(t;\tau) \tag{0.14b}$$

$$g(t;\tau) = \frac{1}{N_t} \sum_{k\in\mathcal{T}} \hat{f}_k(t;\tau) \tag{0.14c}$$

$$q(t,t',\tau) = \frac{1}{N_t} \sum_{k\in\mathcal{T}} x_k(t;\tau) x_k(t',\tau) \tag{0.14d}$$

$$Q(t,t',\tau) = \frac{1}{N_t} \sum_{k\in\mathcal{T}} \hat{f}_k(t;\tau) \hat{f}_k(t',\tau) \tag{0.14e}$$

$$K(t,t',\tau) = \frac{1}{N_t} \sum_{k\in\mathcal{T}} x_k(t;\tau) \hat{f}_k(t',\tau)\,, \tag{0.14f}$$

Let $[a]_{t,\tau}$ denote $a(t;\tau)$. We remark that $\sum_{k\neq j} w_{kj}(\tau) = N_t \int dt\, m(t;\tau) g(t;\tau)$. Thus,

$$1 = \int D[m\hat{m}g\hat{g}\hat{\mu}w] \exp\left\{ i \int dt \left[ \hat{\mu}\Big( \sum_{k\neq j\in\mathcal{T}} w_{kj} - mgN_t \Big) \right]_{t,\tau} \right\}$$
$$\times \exp\left\{ i \int dt \left[ \hat{m}\Big( mN_t - \sum_{j\in\mathcal{T}} x_j \Big) + \hat{g}\Big( gN_t - \sum_{k\in\mathcal{T}} \hat{f}_k \Big) \right]_{t,\tau} \right\}.$$

Using this identity to perform the sum over $\{\mathbf{J}(\tau)\}$ in Eq. (0.6$a$), we can write the moment generating functional $Z[\boldsymbol{\psi}, \boldsymbol{\Psi}]$ as follows:

$$Z[\boldsymbol{\psi}, \boldsymbol{\Psi}] = \int D[x\hat{x}f\hat{f}h\hat{h}] \exp\Big\{ \hat{\mathcal{L}}_0 + \sum_{\tau=0}^{T_{\max}-1} [B(\tau) + D(\tau)] \Big\} \tag{0.15}$$

$$\hat{\mathcal{L}}_0 \equiv \mathcal{L}_0 + \sum_{k\neq j\in\mathcal{T}} \ln 2\cosh(\Psi_{kj}(0) - i w_{kj}(0)) \tag{0.16}$$

where $B(\tau)$ and $D(\tau)$ shall be obtained by summing over the respective parts $J_{kj}^{(tt)}(\tau)$ and $\Delta J_{kj}(\tau)$ of $J_{kj}^{(e)}$ given in Equation (0.10). The sum over $\Big\{ J_{kj}^{(tt)}(\tau) \Big\}$ is straightforward to be carried out and yields

$$B(\tau) = \ln\left\{ \int D[\hat{\mu}wmg] \prod_{k\neq j\in\mathcal{T}} 2\cosh\left( \frac{\Psi_{kj}(\tau+1) + \beta h_{kj}^{(tt)}(\tau) - i w_{kj}(\tau+1)}{N_t} \right) \right\}$$
$$+ \ln\left\{ \exp\Big( i\hat{\mu}(\tau+1) \int dt \Big[ g\Big( mN_t - \sum_{k\in\mathcal{T}} x_k \Big) + m\Big( gN_t - \sum_{k\in\mathcal{T}} \hat{f}_k \Big) \Big]_{t,\tau+1} \Big) \right\}$$

To compute the sum over $\Big\{\Delta J_{kj}(\tau)\Big\}$, for $\Delta J_{kj}(\tau)=\sum_{l\in\mathcal{O}} J^{(to)}_{kl}(\tau)J^{(ot)}_{lj}(\tau)$ we introduce:

$$\begin{aligned}\exp\left(\sum_{k\neq j\in\mathcal{T}}\tilde{I}_{kj}\right) &:= \exp\left\{-\frac{i}{N_t}\sum_{k\neq j\in\mathcal{T}}\sum_{l\notin\mathcal{T}}\int dt\Big[\hat{f}_k x_j J^{(to)}_{kl}(\tau)J^{(ot)}_{lj}(\tau)\Big]_{t,\tau}\right\}\\ &= \exp\left\{-i\sum_{l\notin\mathcal{T}}\int dt\left[\sum_{k\in\mathcal{T}}\frac{J^{(to)}_{kl}(\tau)}{\sqrt{N_t}}\hat{f}_k\right]_{t,\tau}\left[\sum_{j\in\mathcal{T}}\frac{J^{(ot)}_{lj}(\tau)}{\sqrt{N_t}}x_j\right]_{t,\tau}\right\}\\ &= \int\prod_{l\notin\mathcal{T}}Dy_lDz_l e^{-i\int dt\, z_l(t,\tau)y_l(t,\tau)}\\ &\times\int dt\left[\delta\Big(z_l-\sum_{k\in\mathcal{T}}\frac{J^{(to)}_{kl}\hat{f}_k}{\sqrt{N_t}}\Big)\delta\Big(y_l-\sum_{j\in\mathcal{T}}\frac{J^{(ot)}_{lj}x_j}{\sqrt{N_t}}\Big)\right]_{t,\tau}\end{aligned}$$

The delta functions in this expression can be represented using the conjugate fields $\hat{y}_l$ and $\hat{z}_l$ as follows

$$\int D[y\hat{y}z\hat{z}]\exp\left\{i\sum_{l\notin\mathcal{T}}\int dt[N_t^{1/2}(z_l\hat{z}_l+y_l\hat{y}_l)-z_ly_l]_{t,\tau}-i\sum_{k\in\mathcal{T},l\notin\mathcal{T}}\int dt\,[J^{(to)}_{kl}\hat{f}_k\hat{z}_l+J^{(ot)}_{lk}x_k\hat{y}_l]_{t,\tau}\right\}$$

Substituting $\hat{y}_l=z_l/\sqrt{N_t}$ and $\hat{z}_l=y_l/\sqrt{N_t}$ that are obtained by varying the exponent wrt $y_l(t;\tau)$ and $z_l(t;\tau)$, respectively and denoting

$$X_{kl}(\tau)\equiv N_t^{-1/2}\Big\{J^{(to)}_{kl}(\tau)[\Psi_{kl}(\tau)+\beta h^{(to)}_{kl}(\tau-1)]+J^{(ot)}_{lk}(\tau)[\Psi_{lk}(\tau)+\beta h^{(ot)}_{lk}(\tau-1)]\Big\},$$

we have

$$D(\tau)=\ln\left\{\sum_{\left\{J^{(to)}_{kl}(\tau),J^{(ot)}_{lk}(\tau)\right\}}\exp\Big(\sum_{k\in\mathcal{T},l\notin\mathcal{T}}X_{kl}(\tau)+\sum_{k\neq j\in\mathcal{T}}\tilde{I}_{kj}\Big)\right\} \tag{0.17a}$$

$$=\ln\left\{\int D[yzq\hat{q}Q\hat{Q}K\hat{K}]\,e^{N_t[i\Phi+\Omega]}\exp\Big(\sum_{l\notin\mathcal{T}}\int dt\, z_l(t;\tau)y_l(t;\tau)\Big)\right\} \tag{0.17b}$$

where we introduced

$$\begin{aligned}\Phi = &-\frac{1}{N_t}\sum_{k\in\mathcal{T}}\int dtdt'\,\hat{q}(t,t',\tau)x_k(t,\tau)x_k(t',\tau)\\ &-\frac{1}{N_t}\sum_{k\in\mathcal{T}}\int dtdt'\,\Big[\hat{Q}(t,t',\tau)\hat{f}_k(t,\tau)\hat{f}_k(t',\tau)+\hat{K}(t,t',\tau)\hat{f}_k(t',\tau)x_k(t,\tau)\Big]\\ &+\int dtdt'\left\{q(t,t',\tau)\left[\hat{q}(t,t',\tau)+\frac{i}{2N_t}\sum_{l\notin\mathcal{T}}z_l(t;\tau)z_l(t',\tau)\right]\right\}\\ &+\int dtdt'\left\{Q(t,t',\tau)\left[\hat{Q}(t,t',\tau)+\frac{i}{2N_t}\sum_{l\notin\mathcal{T}}y_l(t;\tau)y_l(t',\tau)\right]\right\}\\ &+\int dtdt'\left\{K(t,t',\tau)\left[\hat{K}(t,t',\tau)+\nu\frac{i}{N_t}\sum_{l\notin\mathcal{T}}z_l(t;\tau)y_l(t',\tau)\right]\right\}\end{aligned} \tag{0.18}$$

$$\Omega = \frac{1}{2N_t^2}\sum_{k\in\mathcal{T},l\notin\mathcal{T}}\Big(\Psi_{kl}(\tau+1)+\beta h_{kl}^{(to)}(\tau)+\Psi_{lk}(\tau+1)+\beta h_{lk}^{(ot)}(\tau)\Big)^2$$
$$-\frac{i}{N_t^2}\sum_{k\in\mathcal{T},l\notin\mathcal{T}}\Big[\Psi_{kl}+\beta h_{kl}^{(to)}+\Psi_{lk}+\beta h_{lk}^{(ot)}\Big]\int dt[y_l\hat{f}_k+z_l x_k]_{t,\tau} \qquad (0.19)$$

We can now have write $Z[\boldsymbol{\psi},\boldsymbol{\Psi}]$ in Eq. (0.15) in terms of an action $S$

$$Z[\boldsymbol{\psi},\boldsymbol{\Psi}] = \int D[x\hat{x}f\hat{f}h\hat{h}\hat{\mu}wmgyzq\hat{q}Q\hat{Q}K\hat{K}]e^{S[x,\hat{x},f,\hat{f},h,\hat{h},\hat{\mu},w,m,g,y,z,q,\hat{q},Q,\hat{Q},K,\hat{K}]}$$

where

$$S = W_0 + \sum_{k\neq j\in\mathcal{T}} W_{kj} + \sum_{k\in\mathcal{T},l\notin\mathcal{T}}\tilde{W}_{kl} + i\sum_{k\in\mathcal{T}}\left(\int dt S_k^{(\psi)}(t;\tau) - M_k(\tau)\right) \qquad (0.20)$$

$$W_0 = \int dtdt'\left\{q(t,t',\tau)\left[\hat{q}(t,t',\tau)+\frac{i}{2N_t}\sum_{l\notin\mathcal{T}}z_l(t;\tau)z_l(t',\tau)\right]\right\}$$
$$+\int dtdt'\left\{Q(t,t',\tau)\left[\hat{Q}(t,t',\tau)+\frac{i}{2N_t}\sum_{l\notin\mathcal{T}}y_l(t;\tau)y_l(t',\tau)\right]\right\}$$
$$+\int dtdt'\left\{K(t,t',\tau)\left[\hat{K}(t,t',\tau)+\nu\frac{i}{N_t}\sum_{l\notin\mathcal{T}}z_l(t;\tau)y_l(t',\tau)\right]\right\}$$
$$+iN_t\int dt\,\hat{\mu}(\tau)m(t;\tau)g(t;\tau) \qquad (0.21)$$

$$M_k = \hat{\mu}(\tau)\int dt[g(t,\tau)x_k(t;\tau)+m(t;\tau)\hat{f}_k(t;\tau)]$$
$$+\int dtdt'[\hat{q}(t,t',\tau)x_k(t;\tau)x_k(t',\tau)]$$
$$+\int dtdt'\left[\hat{Q}(t,t',\tau)\hat{f}_k(t;\tau)\hat{f}_k(t',\tau)\ +\hat{K}(t,t',\tau)\hat{f}_k(t',\tau)x_k(t;\tau)\right] \qquad (0.22)$$

$$W_{kj} = i\hat{h}_{kj}^{(tt)}(h_{kj}^{(tt)}-\tilde{h}_{kj}) - \ln\left[2\cosh\left(\beta h_{kj}^{(tt)}(\tau)/N_t\right)\right]$$
$$+\ln\left[2\cosh\left(\frac{\Psi_{kj}(\tau+1)+\beta h_{kj}^{(tt)}(\tau)-iw_{kj}(\tau)}{N_t}\right)\right] \qquad (0.23)$$

$$\tilde{W}_{kl} = \frac{(\Psi_{kl}+\beta h_{kl}^{(to)}+\Psi_{lk}+\beta h_{lk}^{(ot)})^2}{2N_t} + i\hat{h}_{kl}^{(to)}(h_{kl}^{(to)}-\tilde{h}_{kl})$$
$$+i\hat{h}_{lk}^{(ot)}(h_{lk}^{(ot)}-\tilde{h}_{lk}) - \ln\left[2\cosh\left(\frac{\beta h_{kl}^{(to)}}{\sqrt{N_t}}\right)\right] - \ln\left[2\cosh\left(\frac{\beta h_{lk}^{(ot)}}{\sqrt{N_t}}\right)\right]$$
$$-i\frac{\Psi_{kl}+\beta h_{kl}^{(to)}+\Psi_{lk}+\beta h_{lk}^{(ot)}}{N_t}\int dt[y_l\hat{f}_k+z_l x_k]_{t,\tau} \qquad (0.24)$$

Setting $\mathbf{\Psi} = \mathbf{0}$, saddle-point conditions wrt $h_{kl}^{(to)}$, $\hat{h}_{kl}^{(to)}$, $h_{lk}^{(ot)}$, $\hat{h}_{lk}^{(ot)}$, $h_{kj}^{(tt)}$ and $\hat{h}_{kj}^{(tt)}$ yield

$$\hat{\lambda}(\tau+1) \equiv \sqrt{N_t}\Big\langle J_{kl}^{(to)}(\tau+1)\Big\rangle := \sqrt{N_t}\left(\lim_{\Psi\to 0}\frac{\partial Z}{\partial \Psi_{kl}}\right) = \tanh\left(\frac{\beta}{\sqrt{N_t}}h_{kl}^{(to)}(\tau)\right)$$

$$\hat{\lambda}(\tau+1) \equiv \sqrt{N_t}\Big\langle J_{lk}^{(ot)}(\tau+1)\Big\rangle := \sqrt{N_t}\left(\lim_{\Psi\to 0}\frac{\partial Z}{\partial \Psi_{lk}}\right) = \tanh\left(\frac{\beta}{\sqrt{N_t}}h_{lk}^{(ot)}(\tau)\right)$$

$$\hat{\mu}(\tau+1) \equiv N_t\Big\langle J_{kj}^{(tt)}(\tau+1)\Big\rangle := N_t\left(\lim_{\Psi\to 0}\frac{\partial Z}{\partial \Psi_{kj}}\right) = \tanh\left(\frac{\beta}{N_t}h_{kj}^{(tt)}(\tau)\right)$$

Next we introduce a measure for the effective dynamics of a single unit $x$ that, for an observable $O = O(x, \hat{x}, f, \hat{f}, h_*, \hat{h}_*, w_*, \hat{\mu}_*)$, can be defined as

$$\Big\langle O(x, \hat{x}, f, \hat{f}, h_*, \hat{h}_*, w_*, \hat{\mu}_*)\Big\rangle_* = \frac{\displaystyle\int DyDzD[x\hat{x}f\hat{f}]Oe^{S_*}}{\displaystyle\int DyDzD[x\hat{x}f\hat{f}]e^{S_*}} \tag{0.26}$$

where $S_*$ denotes the value of the action $S$ evaluated at the saddle point $\mathcal{M}_* := (m_*, g_*, q_*, \hat{q}_*, Q_*, \hat{Q}_*, K_*, \hat{K}_*)$ for $\boldsymbol{\psi} = \mathbf{0}$.

The saddle-point conditions are:

$$\begin{cases} \dfrac{\partial S}{\partial g(t;\tau)} = 0\,, \quad \to m_*(t,\tau) = \langle x(t;\tau)\rangle_* \\[2ex] \dfrac{\partial S}{\partial q(t,t',\tau)} = 0\,, \quad \to \hat{q}_*(t,t',\tau) = -\dfrac{i\alpha}{2}\langle z(t,\tau)z(t,\tau')\rangle_* \\[2ex] \dfrac{\partial S}{\partial Q(t,t',\tau)} = 0\,, \quad \to \hat{Q}_*(t,t',\tau) = -\dfrac{i\alpha}{2}\langle y(t;\tau)y(t',\tau)\rangle_* \\[2ex] \dfrac{\partial S}{\partial K(t,t',\tau)} = 0\,, \quad \to \hat{K}_*(t,t',\tau) = -i\nu\alpha\langle z(t;\tau)y(t',\tau)\rangle_* \end{cases} \tag{0.27}$$

$$\begin{cases} \dfrac{\partial S}{\partial m(t;\tau)} = 0\,, \quad \to g_*(t;\tau) = \langle \hat{f}(t;\tau)\rangle_* \\[2ex] \dfrac{\partial S}{\partial \hat{q}(t,t',\tau)} = 0\,, \quad \to q_*(t,t',\tau) = \langle x(t;\tau)x(t',\tau)\rangle_* \\[2ex] \dfrac{\partial S}{\partial \hat{Q}(t,t',\tau)} = 0\,, \quad \to Q_*(t,t',\tau) = \langle \hat{f}(t;\tau)\hat{f}(t',\tau)\rangle_* \\[2ex] \dfrac{\partial S}{\partial \hat{K}(t,t',\tau)} = 0\,, \quad \to K_*(t,t',\tau) = \langle x(t,\tau)\hat{f}(t',\tau)\rangle_* \end{cases}, \tag{0.28}$$

Additionally, from the normalisation $Z[\boldsymbol{\psi} = \mathbf{0}, \mathbf{\Psi} = \mathbf{0}] = 1$, we have $\langle \hat{f}(t;\tau)\rangle_* = 0$ and $\langle \hat{f}(t;\tau)\hat{f}(t',\tau)\rangle_* = 0$. These imply that

$$g_*(t;\tau) = Q_*(t,t',\tau) = 0$$

The final step consists of integrating out $y$ and $z$ to find $\hat{q}_*, \hat{Q}_*, \hat{K}_*$ as follows

$$\begin{aligned}
\hat{q}_*(t,t',\tau) &= -\frac{i\alpha}{2}\left\langle \left(\sum_{\ell'\in\mathcal{T}} J^{(to)}_{\ell'}(\tau)\hat{f}_{\ell'}(t;\tau)\right)\left(\sum_{\ell\in\mathcal{T}} J^{(to)}_{\ell}(\tau)\hat{f}_{\ell}(t',\tau)\right)\right\rangle_* \\
&= -\frac{i\alpha}{2}\sum_{\ell\in\mathcal{T}}\left\langle [J^{(to)}_{\ell}(\tau)]^2\hat{f}_{\ell}(t;\tau)\hat{f}_{\ell}(t',\tau)\right\rangle_* \\
&= 0 \qquad (0.29)\\
\hat{Q}_*(t,t',\tau) &= -\frac{i\alpha}{2}\left\langle \left(\sum_{\ell'\in\mathcal{T}} J^{(ot)}_{\ell'}(\tau)x_{\ell'}(t;\tau)\right)\left(\sum_{\ell\in\mathcal{T}} J^{(ot)}_{\ell}(\tau)x_{\ell}(t',\tau)\right)\right\rangle_* \\
&= -\frac{i\alpha}{2}\sum_{\ell}\left\langle [J^{(ot)}_{\ell}(\tau)]^2 x_{\ell}(t;\tau)x_{\ell}(t',\tau)\right\rangle_* \\
&= -\frac{i\alpha}{2}C(t,t',\tau) \qquad (0.30)\\
\hat{K}_*(t,t',\tau) &= -i\nu\alpha\left\langle \left(\sum_{\ell'\in\mathcal{T}} J^{(to)}_{\ell'}(\tau)\hat{f}_{\ell'}(t,\tau)\right)\left(\sum_{\ell\in\mathcal{T}} J^{(ot)}_{\ell}(\tau)x_{\ell}(t',\tau)\right)\right\rangle_* \\
&= -i\nu\alpha\sum_{\ell}\left\langle J^{(to)}_{\ell}(\tau)J^{(ot)}_{\ell}(\tau)\right\rangle_* K(t',t,\tau) \\
&= -i\nu\alpha\left(iG(t',t,\tau)\right) \qquad (0.31)
\end{aligned}$$

where $K(t',t,\tau) = \left\langle \hat{f}_{\ell}(t,\tau)x_{\ell}(t',\tau)\right\rangle_*$ and $G(t',t,\tau) := \partial\langle x(t',\tau)\rangle_*/\partial\theta(t,\tau)$.

The effective single-unit dynamics is generated by the following path probability:

$$\begin{aligned}
\mathcal{P}(\{x\},\{f\}) = \int D\hat{x}D\hat{f}\prod_{\tau=0}^{T_{\max}}\exp\left\{i\int dt\Big[\hat{x}(\partial_t+1)x + i\frac{\sigma^2}{2}\hat{x}^2 - \hat{x}\,F(f)\Big]\right\} \qquad (0.32)\\
\times\exp\left\{i\int dt\hat{f}(f - m_*\hat{\mu}(\tau) - \theta)\right\} \\
\times\exp\left\{-\alpha\int dtdt'\Big[\frac{C(t,t',\tau)}{2}\hat{f}(t;\tau)\hat{f}(t',\tau) + i\nu G(t',t,\tau)x(t;\tau)\hat{f}(t',\tau)\Big]\right\}\Bigg\}
\end{aligned}$$

We can further perform integration over the auxiliary fields $\{\hat{f}\}$, using a Gaussian noise $\eta(t;\tau)$ with correlator $\langle\eta(t;\tau)\eta(t',\tau)\rangle = \alpha C(t,t',\tau)$, to rewrite this path probability as

$$\begin{aligned}
\mathcal{P}(\{x\},\{f\}) = \int D\hat{f}\exp\Big\{-\frac{\alpha}{2}\int dtdt'\,C(t,t',\tau)\hat{f}(t;\tau)\hat{f}(t',\tau) + i\int dt\hat{f}(t;\tau)\eta(t;\tau)\Big\}\Big] \\
\times\int D\xi P(\xi)D\eta\prod_{\tau=0}^{T_{\max}}\Big\{\delta\Big[(\partial_t+1)x - F(f) - \xi\Big] \\
\times\delta\left(f - m_*\hat{\mu}(\tau) - \alpha\nu\int dt'G(t,t',\tau)x(t',\tau) - \eta - \theta\right)\Big\} \qquad (0.33)
\end{aligned}$$

Finally,

$$\mathcal{P}(\{x\},\{f\}) = \int D\xi P(\xi)D\eta P(\eta)\prod_{\tau=0}^{T_{\max}}\delta[(\partial_t+1)x - F(f) - \xi]$$

$$\times \delta\Big(f - m_*\hat{\mu}(\tau) - \alpha\nu \int dt' G(t,t',\tau)x(t') - \eta - \theta\Big) \qquad (0.34)$$

The above equation results in the equivalent SDE form of the $x$'s dynamics in Equation (9). We note that the single-unit effective measure now can also be defined as

$$\langle O(\{x\})\rangle_* = \int Dx \left[\int D\xi P(\xi) D\eta P(\eta) \mathcal{P}(\{x\}|\{\eta\},\{\xi\})\right] O(\{x\}) \qquad (0.35)$$

$$= \int D\xi P(\xi) D\eta P(\eta) O(\{x\}|\{\eta\},\{\xi\}) \qquad (0.36)$$

where the conditional average $O(\{x\}|\{\eta\},\{\xi\})$ is taken wrt the conditional probability

$$\mathcal{P}(\{x\}|\{\eta\},\{\xi\}) := \prod_{\tau=0}^{T_{\max}} \prod_t \Big\langle \delta(\partial_t x(t,\tau) + x(t,\tau) - F(f(t;\tau)) - \xi(t;\tau)) \Big\rangle_* \Big| f = \kappa$$

$$\kappa = m_*(t,\tau)\hat{\mu}(\tau) + \alpha\nu \int dt' G(t,t',\tau) x(t',\tau) + \eta(t,\tau) + \theta(t,\tau)$$

*B. Derivation of Equation (29) for the outlier in the spectrum of the target-gene effective dynamics' Jacobian around the zero fixed point*

In this section we consider only the coupling matrix $\mathbf{J}$ that is achieved at the infinite generation $\tau \to \infty$. So all the $\tau$-dependence can be dropped out in considering the $x$'s dynamics as well as in the notation of $\mathbf{J}$. We have $\hat{\mu}(\tau) := \langle J_{ij}^{(tt)}(\tau)\rangle \to \mu/N_t$ and $\hat{\lambda}(\tau) := \langle J_{ik}^{(to)}\rangle = \langle J_{ki}^{(ot)}\rangle \to \lambda/\sqrt{N_t}$ as $\tau \to \infty$. Denoting by $\underline{\mathbf{J}}^{(tt)}$ the set of effective interactions $J_{ij}^{(e)}$ as defined in Eq. (0.10), we can consider $J_{ij}^{(e)}$ as as Gaussian random variable whose mean is $\mu/N_t$ and variance is $\lambda^4/N_t$ with a covariance between $J_{ij}^{(e)}$ and $J_{ji}^{(e)}$ quantified by a symmetry parameter $\Gamma = \nu^2 \in [0,1]$:

$$J_{ij}^{(e)} = \mathcal{N}\left(\frac{\mu}{N_t}, \frac{\lambda^4}{N_t}\right), \qquad \text{and } [J_{ij}^{(e)}, J_{ji}^{(e)}]_{\underline{\mathbf{J}}} = \frac{\Gamma\lambda^4}{N_t} \qquad (0.37)$$

where we have used $[\,\cdot\,]_{\underline{\mathbf{J}}}$ to denote the average taken with respect to the ensemble of random realisations of $\underline{\mathbf{J}}^{(tt)}$ and $[a,b]_{\underline{\mathbf{J}}}$ to denote the covariance of $a$ and $b$. In order to compute the outlier of spectrum of $\underline{\mathbf{J}}^{(tt)}$, we follow the approach described in [43], which shows that the outlier of $\underline{\mathbf{J}}^{(tt)}$, for $\mu \neq 0$, needs to satisfy [105, 106]

$$R(1+\omega_{\text{outlier}}) = \frac{1}{\mu}, \qquad R(\omega) := \frac{1}{N_t}\Big[\sum_{i,j} R_{ij}(\omega)\Big]_{\underline{\mathbf{J}}} \qquad (0.38)$$

where $R_{ij}$ are the entries of the resolvent $\mathbf{R}(\omega) = (\omega\mathbf{I}_{N_t} - \mathbf{z})^{-1}$, $z_{ij} = J_{ij}^{(e)} - \mu/N_t$ and $\mathbf{I}_{N_t}$ is the $N_t$-by-$N_t$ identity matrix. Using the Neumann series for $\mathbf{R}$, we have

$$R(\omega) = \frac{1}{N_t}\Big[\sum_{i,j}(\omega\delta_{ij} - z_{ij})^{-1}\Big]_{\underline{\mathbf{J}}} = \frac{1}{N_t}\left[\sum_{i,j}\left(\frac{\delta_{ij}}{\omega} + \frac{z_{ij}}{\omega^2} + \sum_k \frac{z_{ik}z_{kj}}{\omega^3} + \cdots\right)\right]_{\underline{\mathbf{J}}} \qquad (0.39)$$

We introduce the moment generating functional $\underline{Z}[\boldsymbol{\phi}]$ for the dynamics of Eq. (0.11):

$$\underline{Z}[\boldsymbol{\phi}] = \int D[x\hat{x}f\hat{f}] \exp\Big\{i\sum_{k\in\mathcal{T}}\int dt \tilde{S}_k^{(0)}(t) - i\sum_{k,j\in\mathcal{T}}\int dt[\phi_{kj}(t)z_{kj}(t) + J_{kj}^{(e)}x_j(t)\hat{f}_k(t)]\Big\}$$

where $\tilde{S}_k^{(0)}(t) = \lim_{\tau\to\infty} S_k^{(0)}(t;\tau)$ with $S_k^{(0)}(t;\tau)$ given by Eq. (0.2$b$). When taking the average $\Big[\underline{Z}[\boldsymbol{\phi}]\Big]_{\underline{\mathbf{J}}}$, one needs to compute the average of the second exponent

$$
\begin{aligned}
H[\boldsymbol{\phi}] &= \left[\exp\left\{-i\sum_{k,j\in\mathcal{T}}\int dt[\phi_{kj}(t)z_{kj}(t) + J_{kj}^{(e)}x_j(t)\hat{f}_k(t)]\right\}\right]_{\underline{\mathbf{J}}} \\
&\propto \exp\left\{-\frac{\lambda^4}{2N_t}\sum_{i,j}\left[\int dt(\phi_{ij} + x_j\hat{f}_i)\right]^2\right\} \\
&\times \exp\left\{-\frac{\Gamma\lambda^4}{2N_t}\sum_{i,j}\int dtdt'\Big[\phi_{ij}(t) + x_j(t)\hat{f}_i(t)\Big]\cdot\Big[\phi_{ji}(t') + x_i(t')\hat{f}_j(t')\Big]\right\}
\end{aligned}
$$

Denoting

$$
V_{ij} := \frac{\delta(\ln H)}{\delta\phi_{ij}(t)} = -\frac{\lambda^4}{N_t}\left\{\int dt'\Big[\phi_{ij}(t') + x_j\hat{f}_i(t')\Big] + \Gamma\int dt'\Big[\phi_{ji}(t') + x_i(t')\hat{f}_j(t')\Big]\right\} \tag{0.40}
$$

we can express the moments of $z_{ij}$ via the derivatives of $H[\boldsymbol{\phi}]$ averaged over the dynamical realisations of $\mathbf{x}$ (this kind of averages is denoted by $\langle\cdot\rangle$) as, for example,

$$
[z_{ij}]_{\underline{\mathbf{J}}} = i\frac{\delta Z}{\delta\phi_{ij}}\Big|_{\boldsymbol{\phi}=0} = i\left\langle\frac{\delta\ln H}{\delta\phi_{ij}}\right\rangle\Big|_{\boldsymbol{\phi}=0} = i\langle V_{ij}\rangle|_{\boldsymbol{\phi}=0}
$$

This allows one to rewrite Eq. (0.39) as

$$
\begin{aligned}
R(\omega) = \frac{1}{N_t}\sum_{i,j}\Bigg[&\frac{\delta_{ij}}{\omega} + \frac{i}{\omega^2}\,\langle V_{ij}\rangle - \frac{1}{\omega^3}\sum_k\left\langle V_{ik}V_{kj} + \frac{\delta V_{ik}}{\delta\phi_{kj}}\right\rangle \\
&- \frac{i}{\omega^4}\sum_{k,l}\left\langle V_{ik}V_{kl}V_{lj} + V_{ik}\frac{\delta V_{kl}}{\delta\phi_{lj}} + \frac{\delta V_{ik}}{\delta\phi_{kl}}V_{lj}\right\rangle + \cdots\Bigg]_{\boldsymbol{\phi}=0}
\end{aligned}
$$

We remark the two properties of $V_{ik}$ (terms that are of order $O(N^{-1})$ are neglected), namely

$$
\begin{cases}
\dfrac{1}{N_t}\displaystyle\sum_{i,j,k}\left\langle\frac{\delta V_{ik}}{\delta\phi_{kj}}\right\rangle_{\boldsymbol{\phi}=0} = -\frac{\Gamma\lambda^4}{N_t^2}\sum_{i,j,k}\delta_{ij} = -\Gamma\lambda^4 \\[2ex]
\dfrac{1}{N_t}\displaystyle\sum_{i,j,k,l}\left\langle V_{ik}\frac{\delta V_{kl}}{\delta\phi_{lj}}\right\rangle_{\boldsymbol{\phi}=0} = -\frac{\Gamma\lambda^4}{N_t}\sum_{i,j}\langle V_{ij}\rangle|_{\boldsymbol{\phi}=0}
\end{cases} \tag{0.42}
$$

After some calculations following the procedure detailed in [43], we find

$$
R(\omega) = u(\omega) + v(\omega) \tag{0.43}
$$

where

$$
u(\omega) = \frac{1}{\omega} - \frac{1}{N\omega^3}\sum_{i,j,k}\left\langle\frac{\delta V_{ik}}{\delta\phi_{kj}}\right\rangle_{\boldsymbol{\phi}=0} + \frac{1}{N\omega^5}\sum_{i,j,k,l,m}\left\langle\frac{\delta V_{ik}}{\delta\phi_{kl}}\frac{\delta V_{lm}}{\delta\phi_{mj}} + \frac{\delta V_{ik}}{\delta\phi_{mj}}\frac{\delta V_{lm}}{\delta\phi_{kl}}\right\rangle_{\boldsymbol{\phi}=0} + \cdots
$$

$$
v(\omega) = \frac{iu^2}{N}\sum_{i,j}\langle V_{ij}\rangle|_{\boldsymbol{\phi}=0} - \frac{u^3}{N}\sum_{i,j,k}\langle V_{ik}V_{kj}\rangle|_{\boldsymbol{\phi}=0} - \frac{iu^4}{N}\sum_{i,j,k,l}\langle V_{ik}V_{kl}V_{lj}\rangle|_{\boldsymbol{\phi}=0} + \cdots
$$

The function $u(\omega)$ can be obtained as a solution to the equation [43]

$$u(\omega) = \frac{\omega - \sqrt{\omega^2 - 4\Gamma\lambda^4}}{2\Gamma\lambda^4} \tag{0.44}$$

while for $v(\omega)$ we have

$$v(\omega) = \frac{u^2\Big[(1+\Gamma)m\hat{f}_* + \Gamma u(Q_* m^2 + q_*\hat{f}_*^2 - 2\chi m\hat{f}_*)\Big]}{1-(1+\Gamma)\chi u + \Gamma u^2(\chi^2 - q_* Q_*)} \tag{0.45}$$

Since $\hat{f}_* := \sum_k \hat{f}_k/N = 0$ and $Q_* = 0$, $v(\omega) = 0$. Therefore,

$$R(1+\omega_{\text{outlier}}) = \frac{\omega_{\text{outlier}} + 1 - \sqrt{(\omega_{\text{outlier}}+1)^2 - 4\Gamma\lambda^4}}{2\Gamma\lambda^4} = \frac{1}{\mu} \tag{0.46}$$

The solution to this equation, for $\Gamma = 1$ (i.e. fully symmetric interactions), reads

$$\omega_{\text{outlier}} = \frac{\lambda^4 + \mu^2}{\mu} - 1 \tag{0.47}$$

This equation shows when $\mu = 1$, $\omega_{\text{outlier}} = \lambda^4$. Since the Jacobian for the dynamics of target genes around the zero fixed point $\mathbf{x} = \mathbf{0}$ is $\mathcal{J}_{ij} = -\delta_{ij} + \underline{J}_{ij}^{(tt)}$. The largest eigenvalue of $\mathcal{J}$ is hence $\lambda^4 - 1$. In the case of fully symmetric interactions, the zero fixed point thus becomes marginally stable in the robust phase where both $\mu, \lambda \to 1$. As a final remark, while it is possible to extend the above computation of the outlier to the more general case of non-zero fixed point, we leave it for future work.

### *C. Derivations of Equation (18) and Equation (28)*

From the Appendix A, the evolution of the averages of the intergroup couplings can be described by the following equations:

$$\hat{\lambda}(\tau) := \frac{1}{N_t N_o}\sum_{i\in\mathcal{T}, j\in\mathcal{O}} J_{ij}^{(to)}(\tau)\,, \qquad \hat{\lambda}(\tau+1) = \tanh(\beta h^{(to)}(\tau)) \tag{0.48a}$$

$$\tilde{\lambda}(\tau) := \frac{1}{N_t N_o}\sum_{i\in\mathcal{T}, j\in\mathcal{O}} J_{ji}^{(ot)}(\tau)\,, \qquad \tilde{\lambda}(\tau+1) = \tanh(\beta h^{(ot)}(\tau)) \tag{0.48b}$$

For $k \in \mathcal{T}$ and $j \in \mathcal{T}$, provided that all $J_{ki}^{(tt)}(\tau)$, $i \neq j$, are fixed, then $x_k(\tau)$ changes to $x_k(\tau) + \Delta x_k^{(j)}$ as $J_{kj}^{(tt)}(\tau) \to J_{kj}^{(tt)}(\tau) + \Delta J_{kj}$. To the first order in $\Delta J_{kj}$, using Eq. (0.12) with $F = (\cdot)$ we have

$$\Delta x_k^{(j)} \simeq x_k\Big(J_{kj}^{(tt)}(\tau) + \Delta J_{kj}\Big) - x_k\Big(J_{kj}^{(tt)}(\tau)\Big) \tag{0.49a}$$

$$= \frac{\partial x_k}{\partial J_{kj}}\Big|_{J_{kj}=J_{kj}^{(tt)}(\tau)} \Delta J_{kj} + O\Big((\Delta J_{kj})^2\Big) \tag{0.49b}$$

$$= (1 - x_k^2(\tau))\, x_j(\tau)\Delta J_{kj} + O\Big((\Delta J_{kj})^2\Big) \tag{0.49c}$$

Let $\delta\Psi_k^2$ denote the total change of the fitness $\Psi^2(\tau)$ given in Eq. (15), we have then

$$\delta\Psi_k^2 = \frac{2\Psi(\tau)}{N_t}\Delta x_k^{(j)}\,, \qquad \text{if } x_k(\tau) \to x_k(\tau) + \Delta x_k^{(j)} \tag{0.50}$$

Plugging Eqs. (0.49$c$) and (0.13) into Eq. (0.50), we obtain

$$\delta\Psi_k^2 \simeq \frac{2\Psi(\tau)}{N_t} x_j(\tau)(1-x_k^2(\tau)) \left( \sum_{\ell\in\mathcal{O}} J_{k\ell}^{(to)}(\tau) J_{\ell j}^{(ot)}(\tau) \right)$$

Following similar arguments in specifying $h_{kj}^{(tt)}$ from the fitness $\Psi^2(\tau)$ in Eq. (17), here we can also consider $\delta\Psi_k^2$ as the average of some feedback field $h_{k\ell}^{(to)}$ over all the pairs $(\ell, j)$ such that $\ell \in \mathcal{O}$ and $j \in \mathcal{T}$, i.e. $\delta\Psi^2(\tau) \propto (N_t N_o)^{-1} \sum_{(\ell,j)} h_{j\ell}^{(to)}(\tau))$. This means $h_{\ell,j}^{(to)}$ can be defined as:

$$h_{k\ell}^{(to)}(\tau) = \left(1 - \frac{x_k^2(\tau)}{\hat{q}(\tau)}\right) \frac{\Psi(\tau)}{\hat{q}(\tau)} \Big( J_{k\ell}^{(to)}(\tau) J_{\ell j}^{(ot)}(\tau) x_j(\tau) \Big) \,. \tag{0.51}$$

where we have used $x_k(\tau) \to x_k(\tau)/\sqrt{\hat{q}(\tau)}$ to make use of the relative fitness. Substituting this form of $h_{k\ell}^{(to)}(\tau)$ into Eq. (0.48$a$) and then taking the limit $t, \tau \to \infty$, we arrive at Eq. (28).

*D. Linear stability of the fixed point in Equation (26) in the absence of noise*

Here we follow the approach of [29], where we perturb the fixed point $x_*$ by some small Gaussian white noise and check how the system responds to it. To this end, we linearize the ADMFT second Eq. (9) around $x_*$ and perform a Fourier analysis as follows. For $\sigma = 0$, with $z = \mathcal{N}(0,1)$, Eq. (26) then becomes

$$x_*(z) = f_0(x_*) := \tanh\Big(\mu m_\infty + J_0\sqrt{q}z + \chi\nu\alpha x_*\Big) \tag{0.52}$$

where $m_\infty = \langle x_* \rangle_z$ and $q = \langle x_*^2 \rangle_z$. For $\mu = 0$, $x_* = 0$ is always a fixed point. Perturbations around any fixed point $x_*$ can be described by $x(t) = x_* + \varepsilon x_1(t)$ and $\eta(t) = J_0\sqrt{q}z + \varepsilon z_1(t)$. We have $\langle z_1(t) z_1(t') \rangle = \alpha \langle x_1(t) x_1(t') \rangle$ from the self-consitency relation Eq. (8$d$). Adding a term $\varepsilon\phi(t)$ into the effective process Eq. (9) for $\sigma = 0$, where $\phi(t)$ is a white noise of unit variance, the Fourier components $\tilde{x}_1(\omega)$ of $x_1(t)$ satisfy

$$\tilde{x}_1 = \frac{[1 - f_0^2(x_*)]\tilde{\phi}(\omega)}{i\omega + 1 - \nu\alpha[1 - f_0^2(x_*)]\tilde{G}(\omega)} \,, \tag{0.53}$$

Solving this equation for the square modulus of $\tilde{x}_1(\omega)$, i.e. $\tilde{R}_0(\omega) = |\tilde{x}_1(\omega)|^2$ gives us

$$\tilde{R}_0^{-1}(\omega) = \left| \frac{(i\omega + 1) - \nu\alpha\tilde{G}(\omega)[1 - f_0^2(x_*)]}{1 - f_0^2(x_*)} \right|^2 - \alpha \tag{0.54}$$

The stability of the steady-state solution in Eq. (0.52) is only determined by $\omega = 0$. In this case, $\tilde{G}(\omega = 0) = \chi$ and the above equation simplifies to

$$[\tilde{R}_0(\omega = 0)]^{-1} = \frac{\Big[1 - \nu\alpha\chi(1 - f_0^2(x_*))\Big]^2}{[1 - f_0^2(x_*)]^2} - \alpha \tag{0.55}$$

If $\tilde{R}_0(\omega = 0)$ diverges, perturbations do not decay to zero, signalling an instability of the fixed point. This means the fixed point $x_*$ loses its stability at $\Sigma = 0$, while remains stable at $\Sigma > 0$, for

$$\Sigma(x_*) = \Big[1 - \nu\alpha\chi(1 - f_0^2(x_*))\Big]^2 - \alpha[1 - f_0^2(x_*)]^2 \tag{0.56}$$

For the paramagnetic fixed point with $x_* = q = 0$, so that $f_0(x_*) = 0$, this condition simplifies to

$$\Sigma(x_* = 0) = [1 - \nu\alpha\chi]^2 - \alpha \tag{0.57}$$

For $\nu = 0$ this reduces to the well-known result $\alpha_c = 1$ for static random networks [27], while for $\nu \neq 0$ we can proceed with the definition of $\chi$

$$\chi = \frac{1}{J_0\sqrt{q}} \left\langle \frac{\partial x_*(z)}{\partial z} \right\rangle_z = \left\langle \frac{1 - f_0^2(x_*)}{1 - \nu\alpha\chi(1 - f_0^2(x_*))} \right\rangle_z \tag{0.58}$$

that, around the critical line where $x_* \simeq 0$ and hence $m_\infty$ and $q$ are small, satisfies the following quadratic equation:

$$\chi(1 - \chi\nu\alpha) = 1 \tag{0.59}$$

The *physical* solution to this equation reads

$$\chi_* = \frac{1 - \sqrt{1 - 4\nu\alpha}}{2\nu\alpha} \tag{0.60}$$

Substituting this $\chi_*$ into Eq. (0.57), we arrive at

$$\Sigma(x_* = 0) = \frac{1}{4}\Big[1 + \sqrt{1 - 4\nu\alpha}\Big]^2 - \alpha \tag{0.61}$$

It is easy to check that this reduces to $\alpha_c(1 + \nu) = 1$ and hence $\alpha_c = 0.5$ for $\nu = 1$.

*E. Generalisation of Eqs. (12), (14) and (17) to environments whose optimal phenotypes are given by* quenched *binary patterns*

Suppose in a given environment $a$, the optimal phenotype is $\mathbf{S}^{(a)}$ with $S_i^{(a)} \in \{-1, 1\}$. The definition of fitness Eq. (11) then can be generalised in this case as follows:

$$m_a(\tau) = \frac{1}{N_t} \sum_{i \in \mathcal{T}} w_i(\tau)\,, \quad w_i(\tau) := S_i^{(a)} x_i(\tau) \tag{0.62}$$

For $S_i^{(a)} = 1$, $\forall\, i$, we recover $\Psi(\tau)$ from $m_a(\tau)$. Next since $F(\cdot) = \tanh$ is an odd-symmetric function, by multiplying both sides of Eq. (12), we arrive at

$$\left(\frac{\partial}{\partial t} + 1\right) w_k(t;\tau) = \tanh\Big(\sum_j S_k^{(a)} S_j^{(a)} J_{kj}(\tau) w_j(t;\tau)\Big) + S_k^{(a)} \xi_k(t;\tau) \tag{0.63}$$

Now introducing $\tilde{J}_{kj}(\tau) = S_k^{(a)} S_j^{(a)} J_{kj}(\tau)$ and $\zeta(t;\tau)$ such that $\langle \zeta_k(t,\tau)\zeta_j(t',\tau')\rangle = \sigma^2 S_k^{(a)} S_j^{(a)} \delta_{kj}\delta_{\tau\tau'}\delta(t - t')$, we have the Eq. (0.63) becomes exactly the same as Eq. (12) but for the new variable $w_k(t;\tau)$. Finally, it is rather straighforwards to see that in this case the learning rule for $\tilde{J}_{kj}$ can be similarly obtained by replacing $x_k(\tau)$ by $w_k(\tau)$ and $\hat{q}(\tau)$ by $\hat{q}_a(\tau) := N_t^{-2} \int d^{N_t} w P(w_{i\in\mathcal{T}} | \tilde{J}(\tau))(\sum_{i\in\mathcal{T}} w_i)^2$ in Eq. (17).

## References


[1] Daniel E. Koshland, Albert Goldbeter, and Jeffry B. Stock. Amplification and adaptation in regulatory and sensory systems. *Science*, 217(4556):220–225, 1982.

[2] Shay Stern, Tali Dror, Elad Stolovicki, Naama Brenner, and Erez Braun. Genome-wide transcriptional plasticity underlies cellular adaptation to novel challenge. *Molecular Systems Biology*, 3(1):106, 2007.

[3] Masayo Inoue and Kunihiko Kaneko. Cooperative adaptive responses in gene regulatory networks with many degrees of freedom. *PLOS Computational Biology*, 9(4):1–12, 04 2013.

[4] Herbert Levine, Mohit Kumar Jolly, Prakash Kulkarni, and Vidyanand Nanjundiah. *Phenotypic switching: implications in biology and medicine*. Academic Press, 2020.

[5] CH Waddington. *The Strategy of the Genes*. George Allen & Unwin, 1957.

[6] Chunhe Li and Jin Wang. Quantifying waddington landscapes and paths of non-adiabatic cell fate decisions for differentiation, reprogramming and transdifferentiation. *Journal of The Royal Society Interface*, 10(89):20130787, 2013.

[7] Sui Huang. The molecular and mathematical basis of waddington's epigenetic landscape: A framework for post-darwinian biology? *BioEssays*, 34(2):149–157, 2012.

[8] Tadashi Miyamoto, Chikara Furusawa, and Kunihiko Kaneko. Pluripotency, differentiation, and reprogramming: A gene expression dynamics model with epigenetic feedback regulation. *PLOS Computational Biology*, 11(8):1–25, 08 2015.

[9] Yuuki Matsushita and Kunihiko Kaneko. Homeorhesis in waddington's landscape by epigenetic feedback regulation. *Phys. Rev. Res.*, 2:023083, Apr 2020.

[10] Wulfram Gerstner, Werner M Kistler, Richard Naud, and Liam Paninski. *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press, 2014.

[11] Benoit Duchet, Christian Bick, and Aine Byrne. Mean-Field Approximations With Adaptive Coupling for Networks With Spike-Timing-Dependent Plasticity. *Neural Computation*, 35(9):1481–1528, 08 2023.

[12] Manuel Beiran and Srdjan Ostojic. Contrasting the effects of adaptation and synaptic filtering on the timescales of dynamics in recurrent networks. *PLOS Computational Biology*, 15(3):1–33, 03 2019.

[13] Samuel P. Muscinelli, Wulfram Gerstner, and Tilo Schwalger. How single neuron properties shape chaotic dynamics and signal transmission in random neural networks. *PLOS Computational Biology*, 15(6):1–35, 06 2019.

[14] Lukas Herron, Pablo Sartori, and BingKan Xue. Robust retrieval of dynamic sequences through interaction modulation. *PRX Life*, 1:023012, Dec 2023.

[15] David M. Post and Eric P. Palkovacs. Eco-evolutionary feedbacks in community and ecosystem ecology: interactions between the ecological theatre and the evolutionary play. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1523):1629–1640, 2009.

[16] Jacob Moran and Mikhail Tikhonov. Defining coarse-grainability in a model of structured microbial ecosystems. *Phys. Rev. X*, 12:021038, May 2022.

[17] Jules Fraboul, Giulio Biroli, and Silvia De Monte. Artificial selection of communities drives the emergence of structured interactions. *Journal of Theoretical Biology*, 571:111557, 2023.

[18] Thilo Gross and Bernd Blasius. Adaptive coevolutionary networks: a review. *Journal of The Royal Society Interface*, 5:259–271, 2008.

[19] Rico Berner, Thilo Gross, Christian Kuehn, Jürgen Kurths, and Serhiy Yanchuk. Adaptive dynamical networks. *Physics Reports*, 1031:1–59, 2023.

[20] Christian Kuehn. *Multiple time scale dynamics*, volume 191. Springer, 2015.

[21] John A Hertz, Yasser Roudi, and Peter Sollich. Path integral methods for the dynamics of stochastic and disordered systems. *Journal of Physics A: Mathematical and Theoretical*, 50(3):033001, dec 2016.

[22] P. C. Martin, E. D. Siggia, and H. A. Rose. Statistical dynamics of classical systems. *Phys. Rev.*

*A*, 8:423–437, Jul 1973.

[23] C. De Dominicis and L. Peliti. Field-theory renormalization and critical dynamics above $T_c$: Helium, antiferromagnets, and liquid-gas systems. *Phys. Rev. B*, 18:353–376, Jul 1978.

[24] A.C.C. Coolen. Statistical mechanics of recurrent neural networks II — Dynamics. In F. Moss and S. Gielen, editors, *Neuro-Informatics and Neural Modelling*, volume 4 of *Handbook of Biological Physics*, pages 619–684. North-Holland, 2001.

[25] Carson C Chow and Michael A Buice. Path integral methods for stochastic differential equations. *J. Math. Neurosci.*, 5(1):8, mar 2015.

[26] H. Sompolinsky and Annette Zippelius. Dynamic theory of the spin-glass phase. *Phys. Rev. Lett.*, 47:359–362, Aug 1981.

[27] H. Sompolinsky, A. Crisanti, and H. J. Sommers. Chaos in random neural networks. *Phys. Rev. Lett.*, 61:259–262, Jul 1988.

[28] A. Crisanti and H. Sompolinsky. Dynamics of spin systems with randomly asymmetric bonds: Langevin dynamics and a spherical model. *Phys. Rev. A*, 36:4922—4939, Nov 1987.

[29] Manfred Opper and Sigurd Diederich. Phase transition and 1/f noise in a game dynamical model. *Phys. Rev. Lett.*, 69:1616–1619, Sep 1992.

[30] L. Molgedey, J. Schuchhardt, and H. G. Schuster. Suppressing chaos in neural networks by noise. *Phys. Rev. Lett.*, 69:3717–3719, Dec 1992.

[31] J P L Hatchett and A C C Coolen. Asymmetrically extremely dilute neural networks with langevin dynamics and unconventional results. *Journal of Physics A: Mathematical and General*, 37(29):7199, jul 2004.

[32] Takashi Ichinomiya. Path-integral approach to dynamics in a sparse random network. *Phys. Rev. E*, 72:016109, Jul 2005.

[33] Tobias Galla. Dynamics of random replicators with Hebbian interactions. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(11):P11005, nov 2005.

[34] J. C. Stiller and G. Radons. Dynamics of nonlinear oscillators with random interactions. *Phys. Rev. E*, 58:1789–1799, Aug 1998.

[35] Jonathan Kadmon and Haim Sompolinsky. Transition to chaos in random neuronal networks. *Phys. Rev. X*, 5:041030, Nov 2015.

[36] Francesca Mastrogiuseppe and Srdjan Ostojic. Intrinsically-generated fluctuating activity in excitatory-inhibitory networks. *PLOS Computational Biology*, 13(4):1–40, 04 2017.

[37] Daniel Martí, Nicolas Brunel, and Srdjan Ostojic. Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks. *Phys. Rev. E*, 97:062314, Jun 2018.

[38] A. Crisanti and H. Sompolinsky. Path integral approach to random neural networks. *Phys. Rev. E*, 98:062120, Dec 2018.

[39] Jannis Schuecker, Sven Goedeke, and Moritz Helias. Optimal sequence memory in driven random networks. *Phys. Rev. X*, 8:041029, Nov 2018.

[40] Elisabeth Agoritsas, Thibaud Maimbourg, and Francesco Zamponi. Out-of-equilibrium dynamical equations of infinite-dimensional particle systems i. the isotropic case. *Journal of Physics A: Mathematical and Theoretical*, 52(14):144002, mar 2019.

[41] Ada Altieri, Felix Roy, Chiara Cammarota, and Giulio Biroli. Properties of equilibria and glassy phases of the random lotka-volterra model with demographic noise. *Phys. Rev. Lett.*, 126:258301, Jun 2021.

[42] Eric De Giuli and Camille Scalliet. Dynamical mean-field theory: from ecosystems to reaction networks. *Journal of Physics A: Mathematical and Theoretical*, 55(47):474002, nov 2022.

[43] Joseph W. Baron, Thomas Jun Jewell, Christopher Ryder, and Tobias Galla. Breakdown of random-matrix universality in persistent Lotka-Volterra communities. *Phys. Rev. Lett.*, 130:137401, Mar 2023.

[44] Thibaut Arnoulx de Pirey and Guy Bunin. Aging by near-extinctions in many-variable interacting populations. *Phys. Rev. Lett.*, 130:098401, Feb 2023.

[45] Miguel Aguilera, Masanao Igarashi, and Hideaki Shimazaki. Nonequilibrium thermodynamics of the asymmetric Sherrington-Kirkpatrick model. *Nature Communications*, 14:3685, Jun 2023.

[46] Agnish Kumar Behera, Madan Rao, Srikanth Sastry, and Suriyanarayanan Vaikuntanathan. Enhanced associative memory, classification, and learning with active dynamics. *Phys. Rev. X*, 13:041043, Dec 2023.

[47] David G. Clark, L. F. Abbott, and Ashok Litwin-Kumar. Dimension of activity in random neural networks. *Phys. Rev. Lett.*, 131:118401, Sep 2023.

[48] Axel Prüser, Sebastian Rosmej, and Andreas Engel. Nature of the volcano transition in the fully disordered kuramoto model. *Phys. Rev. Lett.*, 132:187201, Apr 2024.

[49] Sergei Shmakov and Peter B. Littlewood. Coalescence of limit cycles in the presence of noise. *Phys. Rev. E*, 109:024220, Feb 2024.

[50] Carles Martorell, Rubén Calvo, Alessia Annibale, and Miguel A. Muñoz. Dynamically selected steady states and criticality in non-reciprocal networks. *Chaos, Solitons & Fractals*, 182:114809, 2024.

[51] Jerome Garnier-Brun, Michael Benzaquen, and Jean-Philippe Bouchaud. Unlearnable games and "satisficing" decisions: A simple model for a complex world, 2024.

[52] A. C. C. Coolen, R. W. Penney, and D. Sherrington. Coupled dynamics of fast spins and slow interactions: An alternative perspective on replicas. *Phys. Rev. B*, 48:16116–16118, Dec 1993.

[53] D E Feldman and V S Dotsenko. Partially annealed neural networks. *Journal of Physics A: Mathematical and General*, 27(13):4401, jul 1994.

[54] V S Dotsenko and D E Feldman. Statistical mechanics of training in neural networks. *Journal of Physics A: Mathematical and General*, 27(21):L821, nov 1994.

[55] T. Uezu, K. Abe, S. Miyoshi, and M. Okada. Statistical mechanical study of partial annealing of a neural network model. *Journal of Physics A: Mathematical and Theoretical*, 43(2):025004, dec 2009.

[56] B Wemmenhove, N S Skantzos, and A C C Coolen. Slowly evolving connectivity in recurrent neural networks: I. the extreme dilution regime. *Journal of Physics A: Mathematical and General*, 37(31):7653, jul 2004.

[57] B Wemmenhove and N S Skantzos. Slowly evolving random graphs ii: adaptive geometry in finite-connectivity hopfield models. *Journal of Physics A: Mathematical and General*, 37(32):7843, jul 2004.

[58] S Rabello, A C C Coolen, C J Pérez-Vicente, and F Fraternali. A solvable model of the genesis of amino-acid sequences via coupled dynamics of folding and slow-genetic variation. *Journal of Physics A: Mathematical and Theoretical*, 41(28):285004, jun 2008.

[59] Fábio C Poderoso and José F Fontanari. Model ecosystem with variable interspecies interactions. *Journal of Physics A: Mathematical and Theoretical*, 40(30):8723, jul 2007.

[60] Diego Alberici, Nicolas Macris, and Emanuele Mingione. On the convergence to the non-equilibrium steady state of a langevin dynamics with widely separated time scales and different temperatures. *Annales Henri Poincaré*, 2024.

[61] A. E. Allahverdyan and D. B. Saakian. Spin-glass model with partially annealed asymmetric bonds. *Phys. Rev. E*, 58:R5201–R5204, Nov 1998.

[62] Michael B. Elowitz, Arnold J. Levine, Eric D. Siggia, and Peter S. Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.

[63] Mads Kaern, Timothy C Elston, William J Blake, and James J Collins. Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics*, 6(6):451–464, 2005.

[64] Johan Paulsson. Models of stochastic gene expression. *Physics of Life Reviews*, 2(2):157–175, 2005.

[65] Avigdor Eldar and Michael B. Elowitz. Functional roles for noise in genetic circuits. *Nature*, 467:167–173, 2010.

[66] Marie-Anne Félix and Michalis Barkoulas. Pervasive robustness in biological systems. *Nature Reviews Genetics*, 16:483–496, Mar 2015.

[67] Chikara Furusawa, Takao Suzuki, Akiko Kashiwagi, Tetsuya Yomo, and Kunihiko Kaneko. Ubiquity of log-normal distributions in intra-cellular reaction dynamics. *Biophysics*, 1:25–31, 2005.

[68] Tuan Minh Pham and Kunihiko Kaneko. Double-replica theory for evolution of genotype-phenotype interrelationship. *Phys. Rev. Res.*, 5:023049, Apr 2023.

[69] Paul C Bressloff. Stochastic switching in biology: from genotype to phenotype. *Journal of Physics A: Mathematical and Theoretical*, 50(13):133001, feb 2017.

[70] Daniel Nichol, Mark Robertson-Tessi, Alexander R. A. Anderson, and Peter Jeavons. Model genotype–phenotype mappings and the algorithmic structure of evolution. *Journal of The Royal Society Interface*, 16(160):20190332, 2019.

[71] Pengyao Jiang, Martin Kreitman, and John Reinitz. The effect of mutational robustness on the evolvability of multicellular organisms and eukaryotic cells. *Journal of Evolutionary Biology*, 36(6):906–924, 2023.

[72] Paula Garcia-Galindo, Sebastian E. Ahnert, and Nora S. Martin. The non-deterministic genotype–phenotype map of rna secondary structure. *Journal of The Royal Society Interface*, 20(205):20230132, 2023.

[73] Susanna Manrubia, José A. Cuesta, Jacobo Aguirre, Sebastian E. Ahnert, Lee Altenberg, Alejandro V. Cano, Pablo Catalán, Ramon Diaz-Uriarte, Santiago F. Elena, Juan Antonio García-Martín, Paulien Hogeweg, Bhavin S. Khatri, Joachim Krug, Ard A. Louis, Nora S. Martin, Joshua L. Payne, Matthew J. Tarnowski, and Marcel Weiß. From genotypes to organisms: State-of-the-art and perspectives of a cornerstone in evolutionary dynamics. *Physics of Life Reviews*, 38:55–106, 2021.

[74] Kunihiko Kaneko. Evolution of robustness to noise and mutation in gene expression dynamics. *PLOS ONE*, 2(5):1–8, 05 2007.

[75] Aviv Bergman and Mark L. Siegal. Evolutionary capacitance as a general feature of complex gene networks. *Nature*, 424:549–552, Jul 2003.

[76] Stefano Ciliberti, Olivier C Martin, and Andreas Wagner. Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLOS Computational Biology*, 3(2):1–10, 02 2007.

[77] Anton Crombach and Paulien Hogeweg. Evolution of evolvability in gene regulatory networks. *PLOS Computational Biology*, 4(7):1–13, 07 2008.

[78] Tadamune Kaneko and Macoto Kikuchi. Evolution enhances mutational robustness and suppresses the emergence of a new phenotype: A new computational approach for studying evolution. *PLOS Computational Biology*, 18(1):1–20, 01 2022.

[79] Masayo Inoue and Kunihiko Kaneko. Entangled gene regulatory networks with cooperative expression endow robust adaptive responses to unforeseen environmental changes. *Phys. Rev. Res.*, 3:033183, Aug 2021.

[80] Peijie Zhou and Tiejun Li. Construction of the landscape for multi-stable systems: Potential landscape, quasi-potential, A-type integral and beyond. *The Journal of Chemical Physics*, 144(9):094109, 03 2016.

[81] Anissa Guillemin and Michael P. H. Stumpf. Non-equilibrium statistical physics, transitory epigenetic landscapes, and cell fate decision dynamics. *Mathematical Biosciences and Engineering*, 17(6):7916–7930, 2020.

[82] Tanguy Cabana and Jonathan Touboul. Large deviations, dynamics and phase transitions in large stochastic and disordered neural networks. *Journal of Statistical Physics*, 153:211–269, Oct 2013.

[83] H. Eissfeller and M. Opper. New method for studying the dynamics of disordered spin systems without finite-size effects. *Phys. Rev. Lett.*, 68:2094–2097, Mar 1992.

[84] F. Roy, G. Biroli, G Bunin, and C Cammarota. Numerical implementation of dynamical mean field theory for disordered systems: application to the Lotka–Volterra model of ecosystems. *Journal of Physics A: Mathematical and Theoretical*, 52(48):484001, nov 2019.

[85] Wenxuan Zou and Haiping Huang. Introduction to dynamical mean-field theory of randomly connected neural networks with bidirectionally correlated couplings. *SciPost Phys. Lect. Notes*, page 79, 2024.

[86] Jordan Breffle, Subhadra Mokashe, Siwei Qiu, and Paul Miller. Multistability in neural systems with random cross-connections. *Biological Cybernetics*, Dec 2023.

[87] David Sussillo and L.F. Abbott. Generating coherent patterns of activity from chaotic neural networks. *Neuron*, 63(4):544–557, 2009.

[88] David Saad and Alexander Mozeika. Emergence of equilibriumlike domains within nonequilibrium ising spin systems. *Phys. Rev. E*, 87:032131, Mar 2013.

[89] Ronald A Fisher. *The genetical theory of natural selection*. Oxford, 1930.

[90] DO Hebb. *The organization of behavior. A neuropsychological theory*. John Wiley, 1949.

[91] Oskar H. Schnaack, Luca Peliti, and Armita Nourmohammad. Learning and organization of memory for evolving patterns. *Phys. Rev. X*, 12:021063, Jun 2022.

[92] Chikara Furusawa and Kunihiko Kaneko. Formation of dominant mode by evolution in biological systems. *Phys. Rev. E*, 97:042410, Apr 2018.

[93] Takuya U. Sato and Kunihiko Kaneko. Evolutionary dimension reduction in phenotypic space. *Phys. Rev. Res.*, 2:013197, Feb 2020.

[94] Takuya U. Sato, Chikara Furusawa, and Kunihiko Kaneko. Prediction of cross-fitness for adaptive evolution to different environmental conditions: Consequence of phenotypic dimensional reduction. *Phys. Rev. Res.*, 5:043222, Dec 2023.

[95] Katsuhiko Sato, Yoichiro Ito, Tetsuya Yomo, and Kunihiko Kaneko. On the relation between fluctuation and response in biological systems. *Proceedings of the National Academy of Sciences*, 100(24):14086–14090, 2003.

[96] Kunihiko Kaneko and Chikara Furusawa. An evolutionary relationship between genetic variation and phenotypic fluctuation. *Journal of Theoretical Biology*, 240(1):78–86, 2006.

[97] Ayaka Sakata and Kunihiko Kaneko. Dimensional reduction in evolving spin-glass model: Correlation of phenotypic responses to environmental and mutational changes. *Phys. Rev. Lett.*, 124:218101, May 2020.

[98] Qian-Yuan Tang and Kunihiko Kaneko. Dynamics-evolution correspondence in protein structures. *Phys. Rev. Lett.*, 127:098103, Aug 2021.

[99] Christian. R. Landry, Bernardo Lemos, Scott. A. Rifkin, W. J. Dickinson, and Daniel. L. Hartl. Genetic properties influencing the evolvability of gene expression. *Science*, 317(5834):118–121, 2007.

[100] Rafael Silva-Rocha and Víctor de Lorenzo. Noise and robustness in prokaryotic regulatory networks. *Annual Review of Microbiology*, 64(1):257–275, 2010.

[101] Yui Uchida, Shuji Shigenobu, Hiroyuki Takeda, Chikara Furusawa, and Naoki Irie. Potential contribution of intrinsic developmental stability toward body plan conservation. *BMC Biology*, 20:82, April 2022.

[102] Erez Dekel and Uri Alon. Optimality and evolutionary tuning of the expression level of a protein. *Nature*, 436:588–592, 07 2005.

[103] David G. Clark and L. F. Abbott. Theory of coupled neuronal-synaptic dynamics. *Phys. Rev. X*, 14:021001, Apr 2024.

[104] Ulises Pereira-Obilinovic, Johnatan Aljadeff, and Nicolas Brunel. Forgetting leads to chaos in attractor networks. *Phys. Rev. X*, 13:011009, Jan 2023.

[105] Florent Benaych-Georges and Raj Rao Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 227(1):494–521, 2011.

[106] Sean O'Rourke and David Renfrew. Low rank perturbations of large elliptic random matrices. *Electron. J. Probab.*, 19:1–65, 2014.