Supplementary Text for
Universal structural requirements for maximal robust perfect adaptation in biomolecular networks

Ankit Gupta and Mustafa Khammash

ETH Zürich, Department of Biosystems Science and Engineering,
Mattenstrasse 26, 4058 Basel, Switzerland
Supplementary Text

S1 Preliminaries 3
S1.1 Reaction Networks .......................................................... 3
S1.2 The maxRPA property .......................................................... 6
S1.3 RPA and Integrators ........................................................... 7
S1.4 Encoding of the set-point .................................................... 8

S2 Main Results 9
S2.1 Characterisation of deterministic maxRPA networks ................. 9
S2.2 Characterisation of stochastic maxRPA networks ..................... 14
S2.3 A couple of simple lemmas ................................................. 20
S2.4 Generic decomposition of stochastic maxRPA networks .......... 21

S3 maxRPA networks with minimal IMs 24
S3.1 Homothetic networks with singleton IMs ............................... 24
   S3.1.1 Zero-order degradation networks .................................. 24
   S3.1.2 Autocatalytic networks .............................................. 25
S3.2 Antithetic maxRPA networks with dyadic IMs ....................... 25

S4 Automated discovery of optimal maxRPA controllers 27
S1 Preliminaries

In this section we discuss reaction networks in both deterministic and stochastic settings. Then we mathematically define our criterion for maxRPA and discuss how this property is connected to the existence of integrators [1].

S1.1 Reaction Networks

Consider a reaction network with \( N \) biomolecular species \( \mathbf{X}_1, \ldots, \mathbf{X}_N \) that interact among each other through \( K \) reactions of the form (1) where \( \nu_{ik} \) (resp. \( \nu'_{ik} \)) are stoichiometric constants in \( \mathbb{N}_0 \) (i.e. the set of non-negative integers) denoting the number of molecules of species \( \mathbf{X}_i \) that are consumed (resp. produced) by reaction \( k \). Hence the change in the vector of species molecular counts, brought about by a firing of reaction \( k \), is given by the stoichiometric vector \( \zeta_k = (\nu'_{ik} - \nu_{ik}, \ldots, \nu'_{Nk} - \nu_{Nk}) \in \mathbb{Z}^N \) (where \( \mathbb{Z} \) is the set of all integers).

Deterministic model: In the classical deterministic model of a reaction network, the state \( x = (x_1, \ldots, x_N) \in \mathbb{R}_+^N \) (where \( \mathbb{R}_+ \) is the set of non-negative reals) denotes the vector of species concentrations. For each reaction \( k \), the state \( x \) is mapped to the flux for reaction \( k \) according to a propensity function \( \lambda_k : \mathbb{R}_+^N \rightarrow \mathbb{R}_+ \). Then the reaction dynamics \( (x(t))_{t \geq 0} \) is described by the following system of ODEs

\[
\frac{dx}{dt} = D(x) := \sum_{k=1}^{K} \lambda_k(x) \zeta_k. \tag{S1}
\]

For our results, we shall be assuming that a solution of this system of ODEs exists uniquely for some open set \( \mathcal{E} \subset \mathbb{R}_+^N \) which is time-invariant under the dynamics, and there is a globally attracting fixed point \( \bar{x} \), satisfying

\[
D(\bar{x}) = 0. \tag{S2}
\]

If the propensity function for reaction \( k \) has mass-action form, then it means that \( \lambda_k \) is given by

\[
\lambda_k(x) = \theta_k \prod_{i=1}^{N} x_i^{\nu_{ik}} \tag{S3}
\]

where \( \theta_k \) is a positive reaction rate constant and \( \nu_{ik} \)-s are as in (1).

Stochastic model: In the standard stochastic model of a reaction network, the state \( x = (x_1, \ldots, x_N) \in \mathbb{N}_0^N \) denotes the vector of species copy-numbers or molecular counts. The propensity function \( \lambda_k : \mathbb{N}_0^N \rightarrow \mathbb{R}_+ \) maps the state \( x \) to the rate of firing \( \lambda_k(x) \) of reaction \( k \), and when this reaction fires it displaces the state to \( (x + \zeta_k) \) where \( \zeta_k \) is the...
stoichiometry vector as before. It is natural to impose the following restriction on each propensity function $\lambda_k$

$$\lambda_k(x) > 0 \iff x \geq \nu_k \text{ (componentwise)}, \quad (S4)$$

that essentially means that each reaction $k$ has a positive rate of firing if and only if the requisite number of reactant molecules (given by $\nu_{ik}$ for species $X_i$) are present. This assumption is biologically reasonable and it ensures that the non-negative integer orthant $\mathbb{N}_0^N$ is invariant for the stochastic dynamics. This assumption is trivially satisfied for the propensity function $\lambda_k$ if it has the mass-action [2] form

$$\lambda_k(x) = \theta_k \prod_{i=1}^N \frac{x_i(x_i - 1) \ldots (x_i - \nu_{ki} + 1)}{\nu_{ki}!}, \quad (S5)$$

where $\theta_k$ is a positive reaction rate constant as before.

In the stochastic model, the reaction dynamics $(X(t))_{t \geq 0}$ is described by a continuous-time Markov chain (CTMC) [2]. Letting $\mathcal{E} \subset \mathbb{N}_0^N$ to be the set of all accessible states for the CTMC. Then this CTMC can be characterised by its generator [3] defined as

$$A^*f(x) = \sum_{k=1}^K \lambda_k(x)(f(x + \zeta_k) - f(x)), \quad (S6)$$

where $f$ is a bounded real-valued function on $\mathcal{E}$. Under certain special conditions, the domain of the generator can be taken to be the set of all polynomially growing functions on the state-space $\mathcal{E}$ [4]. Letting $p_t$ to be probability distribution of $X(t)$ i.e.

$$p_t(x) = \mathbb{P}(X(t) = x), \quad x \in \mathcal{E},$$

the well-known Chemical Master Equation (CME) [2] expresses the time-evolution of $p_t$ as

$$\frac{dp_t}{dt} = A^*p_t, \quad (S7)$$

where $A^*$ is the adjoint of linear operator $A$ given by

$$A^*\xi(x) = \sum_{k=1}^K (\lambda_k(x - \zeta_k)\xi(x - \zeta_k) - \lambda_k(x)\xi(x)).$$

The CTMC $(X(t))_{t \geq 0}$ is said to be ergodic if regardless of the initial distribution, as $t \to \infty$, the probability distribution $p_t$ converges to a stationary distribution $\pi$ in the $\ell_1$ norm i.e.

$$\lim_{t \to \infty} \|p_t - \pi\|_{\ell_1} := \lim_{t \to \infty} \sum_{x \in \mathcal{E}} |p_t(x) - \pi(x)| = 0.$$

If this convergence is exponential i.e. there exists a constant $\rho > 0$ and another constant $C > 0$ (depending on $p_0$), such that for any $t > 0$

$$\|p(t) - \pi\|_{\ell_1} \leq Ce^{-\rho t}, \quad (S8)$$

4
then the CTMC is called *exponentially ergodic* [5]. The stationary distribution $\pi$ is a globally attracting fixed point for the CME (S7), over the space of probability distributions over $\mathcal{E}$, and it satisfies

$$A^*\pi = 0. \quad (S9)$$

In order to verify ergodicity, a two-step approach is needed. First one needs to check if the state-space $\mathcal{E}$, which is typically infinite, is *irreducible* and then construct a suitable Foster-Lyapunov function over the state-space $\mathcal{E}$ [5]. In order for the state-space $\mathcal{E}$ to be irreducible, for any two states $x, y \in \mathcal{E}$ there must exist a positive-probability sequence of reactions $k_1, \ldots, k_n$ that take state $x$ to $y$. In light of condition (S4) this is equivalent to having

$$z_{j-1} \geq \nu_{k_j} \quad \text{for each} \quad j = 1, \ldots, n$$

where $z_0 = x$, $z_n = y$ and

$$z_j = z_0 + \sum_{l=1}^{j} \zeta_{k_l} \quad \text{for} \quad j = 1, \ldots, n.$$

A Foster-Lyapunov function $V : \mathcal{E} \to [1, \infty)$ on the state-space $\mathcal{E}$ essentially shows that the CTMC has an attractive tendency toward a compact set [5]. To verify exponential ergodicity this function needs to be *norm-like* (i.e. all sub-level sets must be compact) and for some $C_1, C_2 > 0$

$$AV(x) \leq C_1 - C_2 V(x) \quad \text{for all} \quad x \in \mathcal{E}, \quad (S10)$$

where $A$ is the generator of the CTMC (see Theorem 7.1 in [5]). For stochastic reaction networks, computational frameworks for systematically checking state-space irreducibility and constructing Foster-Lyapunov functions are provided in [6] and [4]. In fact, for many biological networks, a linear Foster-Lyapunov

$$V(x) = 1 + \langle v, x \rangle \quad (S11)$$

can be constructed, where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^N$ and $v \in \mathbb{R}^N_+$ is a vector found by linear programming [4].

For any real-valued function $f(x)$ over the state-space $\mathcal{E}$ satisfying

$$\|f\|_V := \sup_{x \in \mathcal{E}} \frac{|f(x)|}{V(x)} < \infty \quad (S12)$$

we have the ergodic convergence (see Theorem 6.1 in [5])

$$\lim_{t \to \infty} \mathbb{E}(f(X(t))) = \mathbb{E}_\pi(f) := \sum_{x \in \mathcal{E}} f(x)\pi(x), \quad (S13)$$

where $\mathbb{E}$ denotes the expectation operator and $\mathbb{E}_\pi$ denotes expectation under the stationary distribution. Moreover (S9) implies that

$$\mathbb{E}_\pi(Af) = 0. \quad (S14)$$
S1.2 The maxRPA property

Suppose that the reaction structure, given by reactions (1), is fixed. Henceforth we designate the first species $X_1$ as the output species. The property of RPA pertains to robustly driving the steady-state abundance level of the output species to some set-point $\theta^*$, despite perturbations or disturbances to the network. The steady-state of the network, in both deterministic and stochastic settings, is determined by the propensity functions, which may depend on parameters such as reaction rate constants for mass-action reactions or other parameters, like temperature, which influence the kinetics. It is natural to assume that the set-point is encoded as a function of these parameters, and we suppose that the first $m$ reactions in the network have mass-action kinetics and their rate constants $\theta_1, \ldots, \theta_m$ encode the set-point $\theta^*$ via some smooth function $\phi$, i.e.

$$\phi(\theta_1, \ldots, \theta_m) = \theta^*.$$

Noting that this set-point will remain invariant to time-rescaling, one can derive a first-order partial differential equation (PDE) for $\phi$ which shows that $m$ must be at least two (see Section S1.4). As we are concerned with maxRPA networks, we would like our set-point to be determined by the least number of parameters. Hence we set $m = 2$ and in this case our PDE shows that $\phi$ is essentially only a function of the ratio of the two parameters $\theta_1$ and $\theta_2$. Denoting this function as $\phi_{\text{out}}$, the set-point becomes

$$\theta^* = \phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right) \quad \text{with} \quad \phi'_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right) \neq 0.$$  \hfill (S15)

The non-zero derivative condition is added to rule out the trivial case where the set-point encoding function does not depend on its argument. Our results will show that this set-point encoding function cannot be arbitrary for maxRPA systems and must have a specific form depending on the stoichiometry of the first two reactions.

We shall allow disturbances to enter the system through perturbations of the propensity functions of all the reactions except the first two. These $(K - 2)$ reactions need not follow mass-action kinetics and we call the vector of their propensity functions $\lambda = (\lambda_3, \ldots, \lambda_K)$ as the uncertain propensity map which belongs to some pre-defined set $\Lambda_u$ of uncertain propensity maps. In our setup, constant-in-time disturbances affect the system by perturbing the uncertain propensity map $\lambda$ to another map in the uncertain set $\Lambda_u$.

Given this framework, we now define the maxRPA property in a mathematically precise way. Let $\theta = (\theta_1, \theta_2)$ be the vector of rate constants for the first two reactions, which we assume belongs to some feasible open set $\Theta \subset \mathbb{R}^2_+$, and let $f : \mathbb{R}^N \to \mathbb{R}$ be the projection map, that maps the state vector to the state of the output species

$$f(x_1, \ldots, x_N) = x_1.$$  \hfill (S16)

For any fixed $(\theta, \lambda) \in \Theta \times \Lambda_u$, let $(x_{\theta,\lambda}(t))_{t\geq 0}$ be the deterministic dynamics with fixed point $x_{\theta,\lambda}$. Then we say that the network satisfies the maxRPA property in the deterministic setting if and only if

$$\lim_{t \to \infty} f(x_{\theta,\lambda}(t)) = f(\bar{x}_{\theta,\lambda}) = \phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right),$$  \hfill (S17)

with $\phi'_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right) \neq 0$. 

(6)
for any \((\theta, \lambda) \in \Theta \times \Lambda_u\). Similarly let \((X_{\theta,\lambda}(t))_{t \geq 0}\) be the stochastic dynamics with stationary distribution \(\pi_{\theta,\lambda}\). Moreover let \(E\) (resp. \(E_\pi\)) denote the expectation operator under the natural probability distribution (resp. the stationary probability distribution) for CTMC.

Then we say that the network satisfies the maxRPA property in the \textit{stochastic setting} if and only if

\[
\lim_{t \to \infty} E(f(X_{\theta,\lambda}(t))) = E_{\pi_{\theta,\lambda}}(f) = \phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right),
\]

(S18)

for any \((\theta, \lambda) \in \Theta \times \Lambda_u\).

\section*{S1.3 RPA and Integrators}

The \textit{internal model principle} (IMP) \cite{7, 8} from control theory implies that for a system to reject constant-in-time disturbances, it must contain a subsystem, called the \textit{internal model} (IM), that is able to generate such disturbances. For this to hold, the subsystem must be an “integrator” as constant disturbances are generated by the ODE

\[
\dot{z} = 0.
\]

Additionally, IMP stipulates that the internal model IM is only affected by the system output, and it receives no other direct inputs - neither from the rest of the system, nor from the disturbance itself (see Figure 2(A)). Intuitively, IM can be viewed as providing an estimate of the external disturbances, based \textit{only} on the deviation of the output from its set-point.

Consider the deterministic setting and suppose there is a real-valued function \(F(x)\) such that if we let \(z_{\theta,\lambda}(t) = F(x_{\theta,\lambda}(t))\) then we have

\[
\dot{z}_{\theta,\lambda}(t) = \phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right) - f(x_{\theta,\lambda}(t)).
\]

(S19)

In this case \(z_{\theta,\lambda}(t)\) is an \textit{integrator} whose time-derivative measures the deviation of the output from the set-point. The function \(F(x)\) is encoded by the IM and it determines the change-of-coordinates that produces the integrator. Observe that for an integrator to achieve its objective of measuring output-deviations, we can allow for more flexibility in the r.h.s. of (S19). For example, if \(\Psi(x)\) is any nonlinear function of the state-vector \(x = (x_1, \ldots, x_N)\) which is monotonic in the first coordinate \(x_1\), then a more general version of the integrator equation (S19) is given by

\[
\dot{z}_{\theta,\lambda}(t) = \Psi\left(\phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right), y_{\theta,\lambda}(t)\right) - \Psi(x_{\theta,\lambda}(t)),
\]

(S20)

where \(y_{\theta,\lambda}(t) = (x_{\theta,\lambda,2}(t), \ldots, x_{\theta,\lambda,N}(t))\) denotes the last \((N-1)\) components of \(x_{\theta,\lambda}(t)\). This integrator is able to sense the deviation of the output only if the dynamics of \(y_{\theta,\lambda}(t)\) stays away those values \(y\) which make the mapping \(x_1 \mapsto \Psi(x_1, y)\) identically zero. Typically \(\Psi(x_1, y)\) is monomial of variables in \(y\) and in this case the integrator is often called \textit{constrained} as the dynamics of \(y_{\theta,\lambda}(t)\) is constrained to stay away from the boundary of \(\mathbb{R}^{N-1}_+\) \cite{9}.
We now turn to the stochastic setting where the output exhibiting maxRPA is the expectation of the copy-number of species $X_1$ (see (S18)). In this case, for a real-valued function $F(x)$ to specify an integrator we must have that for $z_{\theta,\lambda}(t) = \mathbb{E}(F(X_{\theta,\lambda}(t)))$

$$\dot{z}_{\theta,\lambda}(t) = \phi_{\text{out}} \left( \frac{\theta_1}{\theta_2} \right) - \mathbb{E}(f(X_{\theta,\lambda}(t))). \quad (S21)$$

Under the assumption of ergodicity, such a function $F(x)$ exists uniquely (up to addition by a constant) and it can be found by solving the Poisson equation (see Chapter 8 in [10] and [11]). One key difference from the deterministic setting is that we cannot generalise the integrator equation through the introduction of an arbitrary nonlinear function $\Psi$ (see (S20)), as the expectation operator would not generally commute with the nonlinearity introduced by $\Psi$. As our results will indicate, this constraint in the form of the integrator equation manifests itself in structural constraints for the maxRPA networks.

We end this section with the simple observation that existence of an integrator is sufficient to guarantee maxRPA in both deterministic and stochastic setting, as the r.h.s. of the integrator equations need to be zero at steady-state.

### 1.4 Encoding of the set-point

Recall from Section S1.2 that we suppose that the first $m$ reactions have mass-action kinetics and their rate constants (denoted by $\theta_1,\ldots,\theta_m$) encode the set-point $\theta^*$ via some smooth function $\phi$, i.e.

$$\phi(\theta_1,\ldots,\theta_m) = \theta^*.$$  

If we multiply all the network propensity functions by a positive scalar $r$, then the steady-state will not change (in both deterministic and stochastic settings) as we are essentially just recalibrating time. Hence

$$\phi(r\theta_1,\ldots,r\theta_m) = \phi(\theta_1,\ldots,\theta_m).$$

Differentiating this relation w.r.t. $r$ we obtain the following first-order partial differential equation (PDE)

$$\theta_1 \frac{\partial \phi}{\partial \theta_1} + \cdots + \theta_m \frac{\partial \phi}{\partial \theta_m} = 0.$$  

Note that if $m = 1$ then the only possible solution is that $\phi$ is a constant function. This is not feasible as this would imply that the set-point does not depend on any of the propensity functions! Hence we look for a solution with $m = 2$ of the PDE

$$\theta_1 \frac{\partial \phi}{\partial \theta_1} + \theta_2 \frac{\partial \phi}{\partial \theta_2} = 0.$$  

The characteristic curves for this PDE, with initial values $(\theta_1,\theta_2)$ are $\theta_1(s) = \theta_1 e^s$ and $\theta_2(s) = \theta_2 e^s$. Along these curves the function $\phi$ is constant, i.e.

$$\phi(\theta_1 e^s,\theta_2 e^s) = \phi(\theta_1,\theta_2).$$
Notice that this also proves that for any \( \theta_1, \theta_2 > 0 \)
\[
\phi(\theta_1, \theta_2) = \phi\left(\frac{\theta_1}{\theta_2}, 1\right).
\]
Hence the set-point encoding function \( \phi \) is essentially a function of only one variable, which is the ratio of \( \theta_1 \) and \( \theta_2 \). This function is denoted by \( \phi_{\text{out}} \) in the main text.

### S2 Main Results

In this section we state and prove the main results in the paper. These results provide the necessary and sufficient linear-algebraic conditions for a system to achieve maxRPA, in both deterministic and stochastic settings. It must be noted that these structural conditions are different between the two settings, which may seem contradictory to the fact that the stochastic model converges to the deterministic model under certain volume scaling limits [12]. However this apparent contradiction does not exist because the convergence only holds over finite time-intervals and hence the steady-state behavior (like RPA) can be quite divergent between the stochastic and deterministic scenarios.

#### S2.1 Characterisation of deterministic maxRPA networks

We now characterise maxRPA networks in the deterministic setting and explicitly identify the form of the set-point encoding function (S15) as well as an integrator that demonstrates the maxRPA property. Recall that in our setting, the first two reactions have mass-action propensities with rate constants \( \theta = (\theta_1, \theta_2) \in \Theta \). Hence these propensity functions can be expressed as \( \lambda_k(x) = \theta_k m_k(x) \) for \( k = 1, 2 \), where \( m_k \) denotes the mass-action monomial given by (3). Disturbances enter the reaction system via uncertainties or perturbations in the propensities of the other \( K - 2 \) reactions. These \( K - 2 \) propensities constitute the uncertain propensity map \( \lambda \) which belongs to the uncertain set \( \Lambda_u \). The deterministic reaction dynamics \( (x_{\theta,\lambda}(t))_{t \geq 0} \) satisfies the following system of ODEs

\[
\frac{dx}{dt} = D_{\theta,\lambda}(x) := \sum_{k=1}^{2} \theta_k m_k(x) \zeta_k + \sum_{k=3}^{K} \lambda_k(x) \zeta_k.
\]  

(S22)

We shall make the following assumptions.

**Assumption S2.1** The full row-rank condition (6) holds and for any fixed \( (\theta, \lambda) \in \Theta \times \Lambda_u \) we have the following:

(A) **Stability:** The ODE system (S22) is well-defined for initial values in some open set \( \mathcal{E} \subset \mathbb{R}_+^n \) which is time-invariant under the dynamics and there is a globally attracting fixed point \( \bar{x}_{\theta,\lambda} \in \mathcal{E} \).

(B) **Local richness of the uncertain set \( \Lambda_u \):** There exist \( R \) functions \( \phi_1, \ldots, \phi_R : \mathcal{E} \to \mathbb{R}_{+}^{K-2} \) such that for some \( \epsilon > 0 \)

\[
\lambda_\gamma := \lambda + \sum_{i=1}^{R} \gamma_i \phi_i \in \Lambda_u
\]  

(S23)
for all $\gamma = (\gamma_1, \ldots, \gamma_R) \in (-\epsilon, \epsilon)^R$ and if we define a $(K - 2) \times R$ matrix as

$$\Phi(x) = [\phi_1(x), \ldots, \phi_R(x)]$$  \hspace{1cm} (S24)

then this matrix has full row-rank at $x = \pi_{\theta, \lambda}$.

Part (B) of this assumption implies that around any uncertain propensity map $\lambda$, the uncertain set $\Lambda_u$ is \textit{locally rich} in the sense that it is possible to solely perturb the propensity for each disturbance inducing reaction $k \in \{3, \ldots, K\}$ by translating the propensity map $\lambda$ by a linear combination of functions $\phi_1, \ldots, \phi_R$. Typically we would expect that each propensity gets independently disturbed, i.e. $R = K - 2$ and $\phi_k(x) = (0, \ldots, \hat{\phi}_k(x), 0, \ldots, 0)$. In this case, the rank condition is trivially satisfied and the condition in part (B) reduces to saying that for each $k = 1, \ldots, K - 2$, the propensity $\lambda_{k+2}(x)$ can be perturbed to

$$\lambda_{k+2}(x) + \gamma \hat{\phi}_k(x) \quad \text{for all} \quad \gamma \in (-\epsilon, \epsilon).$$

In the special case of mass-action kinetics, i.e.

$$\lambda_{k+2}(x) = \theta_{k+2} m_{k+2}(x),$$

if we have

$$\hat{\phi}_k(x) = m_{k+2}(x),$$

then this condition implies that the disturbance perturbs the rate constant $\theta_k$ to a value in $(\theta_k - \epsilon, \theta_k + \epsilon)$. While this covers many scenarios of interest, the general condition, as stated in Assumption S2.1(B), is satisfied by many other forms of disturbances, like those arising from dependent sources like metabolic burden, cross-talk, temperature and other extrinsic factors etc.

Observe that if some component of the uncertain propensity map $\lambda$ is the zero function, then one cannot perturb it negatively, and hence (S23) will not hold for all $\gamma \in (-\epsilon, \epsilon)^R$. However our characterisation result will remain valid as its proof will continue to hold upon replacing the usual derivatives w.r.t. components of $\gamma$ with appropriate one-sided derivatives. A zero component in the uncertain propensity map $\lambda$ will naturally arise when one wants to check if adding a disturbance inducing reaction would preserve or violate the maxRPA property.

**Theorem S2.2 (Characterisation of deterministic maxRPA networks)** 
Suppose that the network reaction dynamics is given by (S22) and Assumption S2.1 holds. Then this network exhibits maxRPA for the output species $X_1$ if and only if the following two stoichiometric conditions are satisfied:

(A) Reactions 1 and 2 have as reactants, strictly unequal number of molecules of the output species $X_1$ but equal number of molecules of all other network species, i.e. (7) holds.

(B) There exists a pair $(q, \kappa)$, with $\kappa > 0$, satisfying the linear system (8).
Moreover, if the network exhibits maxRPA, the set-point encoding function is given by (9) and an integrator is given by the function

\[ F(x) := \frac{1}{\theta_2} \sum_{i=1}^{N} q_i \int \frac{1}{m_1(x)} dx_i = \frac{1}{\theta_2} \sum_{i=1}^{N} q_i v_i(x_i) \]  

(S25)

where

\[ v_i(x) = \int \frac{1}{x^{\nu_{i1}}} dx = \begin{cases} 
\log x & \text{if } \nu_{i1} = 1 \\
\frac{1}{1-\nu_{i1}} x^{1-\nu_{i1}} & \text{otherwise} 
\end{cases} \]

Remark S2.3 Observe that the gradient of integrator (S25) is given by

\[ \nabla_x F(x) = \frac{1}{\theta_2 m_1(x)} q \]

which shows that if \( q \) satisfies the linear system (8) then for \( z_{\theta,\lambda}(t) = F(x_{\theta,\lambda}(t)) \) we have

\[ \dot{z}_{\theta,\lambda}(t) = \kappa \frac{\theta_1}{\theta_2} - \frac{m_2(x_{\theta,\lambda}(t))}{m_1(x_{\theta,\lambda}(t))} = \kappa \frac{\theta_1}{\theta_2} - (f(x_{\theta,\lambda}(t)))^\alpha, \]  

(S26)

where \( \alpha := \nu_{12} - \nu_{11} \) and \( f(x) = x_1 \). The last relation holds because \( m_2(x) = x_1^\alpha m_1(x) \) due to condition in part (A). Instead of (S25) we can also define a simpler linear integrator

\[ F_c(x) = \frac{1}{\theta_2} \sum_{i=1}^{N} q_i x_i. \]

In this case for \( z_{\theta,\lambda}(t) = F_c(x_{\theta,\lambda}(t)) \) we have

\[ \dot{z}_{\theta,\lambda}(t) = m_1(x_{\theta,\lambda}(t)) \left( \kappa \frac{\theta_1}{\theta_2} - (f(x_{\theta,\lambda}(t)))^\alpha \right). \]  

(S27)

Both (S26) and (S27) are examples of the general integrator given by (S20).

Proof. Fix a \( (\theta, \lambda) \in \Theta \times \Lambda_u \) and for any \( \gamma = (\gamma_1, \ldots, \gamma_R) \in (-\epsilon, \epsilon)^R \) let

\[ \lambda_{\gamma} = \lambda + \sum_{i=1}^{R} \gamma_i \phi_i \]

where \( \phi_1, \ldots, \phi_R \) are as in Assumption S2.1(B). Due to Assumption S2.1(A) there is a fixed point \( \bar{x}_{\theta,\lambda_{\gamma}} \) satisfying

\[ D_{\theta,\lambda_{\gamma}}(\bar{x}_{\theta,\lambda_{\gamma}}) = 0. \]  

(S28)

Differentiating this jointly w.r.t. parameters \( (\theta, \gamma) \) gives us

\[ \mathcal{J}_x D_{\theta,\lambda_{\gamma}}(\bar{x}_{\theta,\lambda_{\gamma}}) \mathcal{J}_{\theta,\gamma} \bar{x}_{\theta,\lambda_{\gamma}} + \mathcal{J}_{\theta,\gamma} D_{\theta,\lambda_{\gamma}}(\bar{x}_{\theta,\lambda_{\gamma}}) = 0 \]  

(S29)
where $\mathcal{J}_\alpha$ denotes the Jacobian w.r.t. $\alpha$. Note that $\lambda_0 = \lambda$, the fixed point $\mathbf{x}_{\theta,\lambda}$ is globally attracting, and the matrix $\mathcal{J}_x D_{\theta,\lambda}(\mathbf{x}_{\theta,\lambda})$ is invertible. Setting $\gamma = 0$ and multiplying relation (S29) by the inverse of this Jacobian matrix on the left we obtain

$$\mathcal{J}_{\theta,\gamma} \mathbf{x}_{\theta,\lambda} \bigg|_{\gamma = 0} = - \left[ \mathcal{J}_x D_{\theta,\lambda}(\mathbf{x}_{\theta,\lambda}) \right]^{-1} \mathcal{J}_{\theta,\gamma} D_{\theta,\lambda} \mathbf{x}_{\theta,\lambda} \bigg|_{\gamma = 0}. \quad (S30)$$

Observe that since the stoichiometric vectors $\zeta_1, \ldots, \zeta_K$ form the columns of the stoichiometry matrix $S$ (see (5)) we can express the vector field $D_{\theta,\lambda}(x)$ (see (S22)) as

$$D_{\theta,\lambda}(x) = \sum_{k=3}^K \lambda_k(x) \zeta_k + S \Psi(x) \begin{bmatrix} \theta_1 \\ \theta_2 \\ \gamma \end{bmatrix},$$

where $\Psi(x)$ is the $K \times (R + 2)$ dimensional matrix given by

$$\Psi(x) = \begin{bmatrix} m_1(x) & 0 & 0 \\ 0 & m_2(x) & 0 \\ 0 & 0 & \Phi(x) \end{bmatrix},$$

and $\Phi(x)$ is the matrix given by (S24). Since the map $(\theta, \gamma) \mapsto D_{\theta,\lambda}(x)$ is affine we can compute its Jacobian at $x = \mathbf{x}_{\theta,\lambda}$ as

$$\mathcal{J}_{\theta,\gamma} D_{\theta,\lambda} \mathbf{x}_{\theta,\lambda} \bigg|_{\gamma = 0} = S \Psi(\mathbf{x}_{\theta,\lambda}). \quad (S31)$$

Let $f$ be the projection map (S16) that maps the state vector to the state of the output species. For (S17) to hold we must have

$$[\nabla_{\theta,\gamma} f(\mathbf{x}_{\theta,\lambda})]^T = \phi'_\text{out} \left( \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \right) \begin{bmatrix} 1 \\ \theta_2 \\ -\theta_1 \theta_2 \end{bmatrix}, \quad (S32)$$

Let $e_1$ denote the vector in $\mathbb{R}^N$ with the first component as 1 and the rest as zeros. Substituting the value of

$$\mathcal{J}_{\theta,\gamma} D_{\theta,\lambda} \mathbf{x}_{\theta,\lambda} \bigg|_{\gamma = 0}$$

from (S31) into (S30) we obtain

$$[\nabla_{\theta,\gamma} f(\mathbf{x}_{\theta,\lambda})]^T = e_1^T \mathcal{J}_{\theta,\gamma} \mathbf{x}_{\theta,\lambda} \bigg|_{\gamma = 0} = -e_1^T [\mathcal{J}_x D_{\theta,\lambda}(\mathbf{x}_{\theta,\lambda})]^{-1} S \Psi(\mathbf{x}_{\theta,\lambda}). \quad (S33)$$

Letting

$$q_{\theta,\lambda}^T = -e_1^T \left[ \mathcal{J}_x D_{\theta,\lambda}(\mathbf{x}_{\theta,\lambda}) \right]^{-1} \quad (S34)$$

and equating (S32) and (S33) we get

$$q_{\theta,\lambda}^T \zeta_1 = \frac{\phi'_\text{out} \left( \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \right)}{\theta_2 m_1(\mathbf{x}_{\theta,\lambda})}, \quad q_{\theta,\lambda}^T \zeta_2 = -\frac{\theta_1 \phi'_\text{out} \left( \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \right)}{\theta_2^2 m_2(\mathbf{x}_{\theta,\lambda})} \quad (S35)$$

12
and $q^T_{\theta, \lambda} \hat{S} \Phi(\pi_{\theta, \lambda}) = 0$, \hspace{1cm} (S36)

where $\hat{S} = [\zeta_3, \ldots, \zeta_K]$ contains the last $(K - 2)$ columns of the stoichiometry matrix $S$. Due to Assumption S2.1(B), matrix $\Phi(\pi_{\theta, \lambda})$ has full row rank and therefore (S36) can be simplified to

$$q^T_{\theta, \lambda} \zeta_k = 0 \quad \text{for all} \quad k = 3, \ldots, K.$$ \hspace{1cm} (S37)

Let $c$ be a vector satisfying (S63). Using (S35) and (S37) we obtain

$$0 = q^T_{\theta, \lambda} S c = \phi'_\text{out} \left( \frac{\theta_1}{\theta_2} \right) \left[ \frac{c_1}{\theta_2 m_1(\pi_{\theta, \lambda})} - \frac{\theta_1 c_2}{\theta_2^2 m_2(\pi_{\theta, \lambda})} \right],$$

and since $\phi'_\text{out} \left( \frac{\theta_1}{\theta_2} \right) \neq 0$ (see (S15)), this shows that

$$\frac{c_1}{c_2} = \frac{\theta_1 m_1(\pi_{\theta, \lambda})}{\theta_2 m_2(\pi_{\theta, \lambda})} = \phi^{-1}_\text{out}(x_{\theta, \lambda, 1}) \prod_{i=1}^{N} x_{\theta, \lambda, i}^{(\nu_{i1} - \nu_{i2})} \hspace{1cm} (S38)$$

where the last equality follows from (S17). Here $\phi^{-1}_\text{out}(\cdot)$ denotes the inverse of the set-point encoding function and $x_{\theta, \lambda, i}$ is the $i$-th component of the steady-state vector $x_{\theta, \lambda}$. Observe that (S38) must hold for all $(\theta, \lambda) \in \Theta \times \Lambda_u$ with the same constant $c_1/c_2$.

From (S30) and (S31) we note that

$$J_{\theta, \lambda} := J_{\theta, \gamma} x_{\theta, \lambda, \gamma} \bigg|_{\gamma = 0} = - [J_x D_{\theta, \lambda}(x_{\theta, \lambda})]^{-1} S \Psi(x_{\theta, \lambda}).$$

Due to (6) and Assumption S2.1(B), matrices $S$ and $\Psi(x_{\theta, \lambda})$ have full row-ranks, and since this property does not change upon multiplication by a nonsingular matrix on the left, we can conclude that matrix $J_{\theta, \lambda}$ also has full row-rank

$$\text{Rank}(J_{\theta, \lambda}) = N. \hspace{1cm} (S39)$$

Applying the constant rank theorem (see Theorem 11.1 and Proposition 11.7 in [13]) we can conclude that the map $(\theta, \gamma) \mapsto x_{\theta, \lambda, \gamma}$ is locally open and surjective, i.e. it maps an open set $O \subset \Theta \times (-\epsilon, \epsilon)^K$ around $(\theta, 0)$ to an open set $P \subset \mathcal{E}$ around $x_{\theta, \lambda}$. As relation (S38) holds for each $x \in P$, we must have that $\nu_{i2} = \nu_{i1}$ for each $i = 2, \ldots, N$ and

$$\phi^{-1}_\text{out}(x) = \kappa^{-1} x^\alpha$$

for $\alpha = \nu_{i2} - \nu_{i1}$ and $\kappa = \frac{c_2}{c_1} > 0$. Setting

$$q = \frac{\theta_2 m_1(\pi_{\theta, \lambda}) q_{\theta, \lambda}}{\phi'_\text{out} \left( \frac{\theta_1}{\theta_2} \right)} \hspace{1cm} (S40)$$

we see that the pair $(q, \kappa)$ must satisfy the linear system (8). As no other such pair can exist under our assumptions (see Lemma S2.8) $(q, \kappa)$ does not depend on the choice of $(\theta, \lambda) \in \Theta \times \Lambda_u$. This proves the “only if” part of the result.

The “if” part of the result is trivial as the stability of the deterministic dynamics along with the presence of an integrator (S25) (see also Remark S2.3) shows that the network exhibits maxRPA. This completes the proof of this theorem. \hspace{1cm} \square

13
S2.2 Characterisation of stochastic maxRPA networks

In the stochastic setting, the mass-action propensities for the first two reactions, with rate-constants \( \theta = (\theta_1, \theta_2) \in \Theta \), can be written as \( \lambda_k(x) = \theta_k m_k(x) \) for \( k = 1, 2 \), where \( m_k \) denotes the combinatorial factor

\[
m_k(x) = \prod_{i=1}^{N} \frac{x_i(x_i - 1) \ldots (x_i - \nu_{ki} + 1)}{\nu_{ki}!}.
\] (S41)

As in the deterministic scenario, the system is affected by disturbances that perturb the propensity map \( \lambda \) which belongs to the uncertain set \( \Lambda \). The stochastic reaction dynamics is represented by a CTMC \((X_{\theta,\lambda}(t))_{t \geq 0}\) with state-space \( \mathcal{E} \subseteq \mathbb{N}_0^N \) and generator \( \mathbb{A}_{\theta,\lambda} \) given by

\[
\mathbb{A}_{\theta,\lambda} f(x) = 2 \sum_{k=1}^{2} \theta_k m_k(x) \Delta_k f(x) + \sum_{k=3}^{K} \lambda_k(x) \Delta_k f(x),
\] (S42)

where \( f \) is a real-valued function on the state-space \( \mathcal{E} \) and \( \Delta_k \) is the difference operator

\[
\Delta_k f(x) = f(x + \zeta_k) - f(x).
\]

As discussed in Section S1.3, for maxRPA to hold in the stochastic setting, there exists a real-valued function \( F_{\theta,\lambda} \) on the state-space \( \mathcal{E} = \mathbb{N}_0^N \) that specifies the integrator \( z_{\theta,\lambda}(t) = \mathbb{E}(F(X_{\theta,\lambda}(t))) \) which satisfies (S21). Due to Dynkin’s Theorem [3] we can express the time-derivative of the integrator as

\[
\dot{z}_{\theta,\lambda}(t) = \mathbb{E}(\mathbb{A}_{\theta,\lambda} F_{\theta,\lambda}(X_{\theta,\lambda}(t))).
\]

Substituting this in (S21) and observing that the resulting relation needs to hold for any initial state \( X_{\theta,\lambda}(0) \in \mathcal{E} \) we can conclude that \( F_{\theta,\lambda} \) must satisfy the Poisson equation (see Chapter 8 in [10])

\[
\mathbb{A}_{\theta,\lambda} F_{\theta,\lambda}(x) = \phi_{\text{out}} \left( \frac{\theta_1}{\theta_2} \right) - x_1 \quad \text{for any} \quad x \in \mathcal{E}.
\] (S43)

Such a function exists uniquely under our ergodicity assumption [11] and corresponding to this function \( F_{\theta,\lambda} \) we define a bounded function \( G_{\theta,\lambda,k-2}(x) = (G_{\theta,\lambda,1}(x), \ldots, G_{\theta,\lambda,K-2}(x)) \) by

\[
G_{\theta,\lambda,k-2}(x) = \text{Sign} (\Delta_k F_{\theta,\lambda}(x)) \quad \text{for} \quad k = 3, \ldots, K,
\] (S44)

where the sign function is given by

\[
\text{Sign}(x) = \begin{cases} 
-1 & \text{if } x < 0 \\
1 & \text{if } x \geq 0
\end{cases}
\]

In other words, \( G_{\theta,\lambda,k-2}(x) \) captures the signs of the actions of the difference operator \( \Delta_k \) on the function \( F_{\theta,\lambda} \) for the disturbance inducing reactions \( k = 3, \ldots, K \). We now state our assumptions in the stochastic setting.
Assumption S2.4 For any \((\theta, \lambda) \in \Theta \times \Lambda_u\) we have the following:

(A) **Stability:** The state-space \(\mathcal{E} = \mathbb{N}_0^N\) is irreducible for the CTMC dynamics \((X_{\theta, \lambda}(t))_{t \geq 0}\) and there exists a Foster-Lyapunov function \(V_{\theta, \lambda}\) satisfying (S10) with \(\mathbb{A} = \mathbb{A}_{\theta, \lambda}\). Hence this CTMC is exponentially ergodic with a unique stationary distribution \(\pi_{\theta, \lambda}\). The projection map \(f\) (see (S16)) satisfies \(\|f\|_{V_{\theta, \lambda}} < \infty\) where the norm \(\|\cdot\|_{V_{\theta, \lambda}}\) is defined in (S12).

(B) **Local richness of the uncertain set \(\Lambda_u\):** Let \(F_{\theta, \lambda}\) be the solution of the Poisson equation (S43) with \(\mathbb{A} = \mathbb{A}_{\theta, \lambda}\) and let \(G_{\theta, \lambda}(x)\) be the function given by (S48). Then for any \(\epsilon > 0\) there exist \(R\) functions \(\phi_1, \ldots, \phi_R : \mathcal{E} \to \mathbb{R}_+\) such that

\[
\lambda_\gamma := \lambda + \sum_{i=1}^{R} \gamma_i \phi_i \in \Lambda_u \quad \text{for all} \quad \gamma = (\gamma_1, \ldots, \gamma_R) \in (-\epsilon, \epsilon)^R \quad \text{(S45)}
\]

and

\[
\sum_{x \in \mathcal{E}} \left\| G_{\theta, \lambda}(x) - \sum_{i=1}^{R} c_i \phi_i(x) \right\|_{L_1} V_{\theta, \lambda}(x) \pi_{\theta, \lambda}(x) < \epsilon \quad \text{(S46)}
\]

for some constants \(c_1, \ldots, c_R\).

The exponential ergodicity assumption is quite restrictive but it has been found that many biological networks readily satisfy this assumption [4] and this can be shown by constructing a linear Foster-Lyapunov function \(V_{\theta, \lambda}\) (S11). Note that in the case of a linear \(V_{\theta, \lambda}\), the condition \(\|f\|_{V_{\theta, \lambda}} < \infty\) is trivially satisfied. From Theorem 4.5 in [5], we can see that if we define a measure \(\mu\) over \(\mathcal{E}\) as

\[
\mu(x) = V_{\theta, \lambda}(x) \pi_{\theta, \lambda}(x)
\]

then this is a finite measure i.e.

\[
\sum_{x \in \mathcal{E}} \mu(x) < \infty.
\]

Part (B) of Assumption S2.4 essentially says that the set \(\Lambda_u\) of uncertain propensity maps is locally rich enough to allow perturbations by functions whose linear combination can closely approximate the vector-valued function \(G_{\theta, \lambda}(x)\) in the sense of \(L_1\) norm over the measurable space \((\mathcal{E}, \mu)\). Before we prove our main result (Theorem S2.6) we need the following proposition.

**Proposition S2.5** Suppose the conditions of Theorem S2.6 hold, and let \(F_{\theta, \lambda}\) be the solution of the Poisson equation (S43) with \(\mathbb{A} = \mathbb{A}_{\theta, \lambda}\) for some \((\theta, \lambda) \in \Theta \times \Lambda_u\). Then for the maxRPA property to hold, function \(F_{\theta, \lambda}\) must be invariant under translations by the stoichiometric vectors \(\zeta_k\) corresponding to disturbance inducing reactions i.e. for each \(k = 3, \ldots, K\) and each state \(x \in \mathcal{E}\) such that \(x \geq \nu_k\)

\[
F_{\theta, \lambda}(x + \zeta_k) = F_{\theta, \lambda}(x). \quad \text{(S47)}
\]
Proof. Fix a \((\theta, \lambda) \in \Theta \times \Lambda_u\) and an \(\epsilon > 0\). Due to Assumption S2.4(B) there exist \(R\) functions \(\phi_1, \ldots, \phi_R\) and constants \(c_1, \ldots, c_R\) such that (S46) holds. For any \(\gamma = (\gamma_1, \ldots, \gamma_R) \in (-\epsilon, \epsilon)^R\) let

\[ \lambda_\gamma := \lambda + \sum_{i=1}^{R} \gamma_i \phi_i \in \Lambda_u. \]

Define a function \(H_{\theta, \lambda}(x) = (H_{\theta, \lambda, 1}(x), \ldots, H_{\theta, \lambda, K-2}(x))\) by

\[ H_{\theta, \lambda, k-2}(x) = \Delta_k F_{\theta, \lambda}(x) \quad \text{for} \quad k = 3, \ldots, K, \quad (S48) \]

which implies that \(G_{\theta, \lambda}(x) = \text{Sign}(H_{\theta, \lambda}(x))\) where the sign function is being applied coordinate-wise. As \(\|f\|_{V_{\theta, \lambda}} < \infty\) (see Assumption S2.4(A)), Theorem 2.3 in [11] implies that the same holds for the solution of the Poisson equation (S43) \(F_{\theta, \lambda}(x)\). Hence there exists a constant \(C_H > 0\) such that

\[ \sup_{x \in E} \max_{k=1, \ldots, K-2} \frac{|H_{\theta, \lambda, k}(x)|}{V(x)} \leq C_H. \quad (S49) \]

From Theorem 3.3 in [14] it follows that for any \(\gamma_i\) we have

\[ \partial_\gamma \mathbb{E}_{\pi_{\theta, \lambda, \gamma}}(f) \big|_{\gamma = 0} = \sum_{k=3}^{K} \sum_{x \in N_0^N} \partial_\gamma \lambda_k(x) \Delta_k F_{\theta, \lambda}(x) \pi_{\theta, \lambda}(x) \]

\[ = \sum_{k=1}^{K-2} \sum_{x \in N_0^N} \phi_{ik}(x) H_{\theta, \lambda, k}(x) \pi_{\theta, \lambda}(x) \]

\[ = \sum_{x \in N_0^N} \langle H_{\theta, \lambda}(x), \phi_i(x) \rangle \pi_{\theta, \lambda}(x), \quad (S50) \]

where \(\langle \cdot, \cdot \rangle\) denotes the standard inner product in \(\mathbb{R}^{K-2}\), \(f\) is the projection map (S16) and \(\phi_{ik}(x)\) is the \(k\)-th component of \(\phi_i(x)\). For the maxRPA property to hold, the parametric sensitivity defined by the l.h.s. of (S50) must be 0 for each \(\gamma_i\). Multiplying it by \(c_i\) and summing over \(i\) we get

\[ 0 = \sum_{x \in N_0^N} \left( H_{\theta, \lambda}(x), \sum_{i=1}^{R} c_i \phi_i(x) \right) \pi_{\theta, \lambda}(x) \]

\[ = \sum_{x \in N_0^N} \langle H_{\theta, \lambda}(x), G_{\theta, \lambda}(x) \rangle \pi_{\theta, \lambda}(x) - \sum_{x \in N_0^N} \left( \langle H_{\theta, \lambda}(x), G_{\theta, \lambda}(x) - \sum_{i=1}^{R} c_i \phi_i(x) \rangle \right) \pi_{\theta, \lambda}(x). \]

Since

\[ \langle H_{\theta, \lambda}(x), G_{\theta, \lambda}(x) \rangle = \sum_{k=1}^{K-1} |H_{\theta, \lambda, k-2}(x)| = \|H_{\theta, \lambda}(x)\|_{L_1} \]

16
we obtain
\[ \sum_{x \in \mathbb{N}^N} \| H_{\theta,\lambda}(x) \|_{\ell_1, \pi_{\theta,\lambda}(x)} = \sum_{x \in \mathbb{N}^N} \left( H_{\theta,\lambda}(x), G_{\theta,\lambda}(x) - \sum_{i=1}^{R} c_i \phi_i(x) \right) \pi_{\theta,\lambda}(x) \]
\[ \leq C_H \sum_{x \in \mathcal{E}} \left\| G_{\theta,\lambda}(x) - \sum_{i=1}^{R} c_i \phi_i(x) \right\|_{\ell_1} V_{\theta,\lambda}(x) \pi_{\theta,\lambda}(x) \]
\[ \leq C_H \epsilon, \]

where \( C_H \) is the constant in (S49) and the last inequality follows from (S46). As \( \epsilon > 0 \) is arbitrary, letting \( \epsilon \to 0 \) shows that
\[ \sum_{x \in \mathbb{N}^N} \| H_{\theta,\lambda}(x) \|_{\ell_1, \pi_{\theta,\lambda}(x)} = 0. \]

Due to irreducibility of the state-space \( \mathcal{E} = \mathbb{N}^N \) and ergodicity of the CTMC (see Assumption S2.4(A)), we know that \( \pi_{\theta,\lambda}(x) > 0 \) for each \( x \in \mathcal{E} \). This implies that \( H_{\theta,\lambda}(x) = 0 \) for each \( x \in \mathbb{N}^N \) and completes the proof of this proposition. \( \square \)

Theorem S2.6 (Characterisation of stochastic maxRPA networks) Suppose that the network reaction dynamics is given by the CTMC \( (X_{\theta,\lambda}(t))_{t \geq 0} \) and also suppose that Assumption S2.4 holds. Then this network exhibits maxRPA for the output species \( X_1 \) if and only if the following two stoichiometric conditions are satisfied:

(A) Reactions 1 and 2 do not have molecules of species \( X_2, \ldots, X_N \) as reactants, and exactly one of these reactions involves a single molecule of the output species \( X_1 \) as reactant while the other involves none i.e. (12) holds.

(B) There exists a pair \((q, \kappa)\), with \( \kappa > 0 \), satisfying the linear system (8).

Moreover, if the network exhibits maxRPA, the set-point encoding function is given by (13) and an integrator is given by the linear function
\[ F(x) = \frac{1}{\theta_2} \sum_{i=1}^{N} q_i x_i. \]  

(S51)

Remark S2.7 Observe that the stoichiometric conditions imposed by this theorem are stricter than those imposed by Theorem S2.2. This shows that if the underlying assumptions hold, then the set of networks exhibiting maxRPA in the stochastic setting is a strict subset of the set of networks exhibiting maxRPA in the deterministic setting. Also note that in this stochastic case the difference in the number of reactant molecules for the output species \( X_1 \), between reactions 1 and 2 is either 1 or -1. Hence \( \nu_{12} - \nu_{11} \in \{-1, 1\} \) and so the set-point encoding function can also be expressed as
\[ \phi_{\text{out}}(x) = (\kappa x)^{\nu_{12} - \nu_{11}}. \]
Proof. Applying the generator $A_{\theta,\lambda}$ to the function $F_{\theta,\lambda}$ and using the translation invariance shown by Proposition S2.5, we obtain

$$A_{\theta,\lambda}F_{\theta,\lambda}(x) = \theta_1 m_1(x)(F_{\theta,\lambda}(x + \zeta_1) - F_{\theta,\lambda}(x)) + \theta_2 m_2(x)(F_{\theta,\lambda}(x + \zeta_2) - F_{\theta,\lambda}(x)).$$

Since $F_{\theta,\lambda}$ solves the Poisson equation (S43), the r.h.s. of this equation must equate to the r.h.s. of (S43) for all $x \in \mathbb{N}_0^N$. This can only happen if for some distinct $j_1, j_2 \in \{1, 2\}$

$$m_{j_1}(x) \equiv 1 \text{ and } F_{\theta,\lambda}(x + \zeta_{j_1}) - F_{\theta,\lambda}(x) = \frac{1}{\theta_{j_1}} \phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right) \quad (S52)$$

and either

$$m_{j_2}(x) = x \text{ and } F_{\theta,\lambda}(x + \zeta_{j_2}) - F_{\theta,\lambda}(x) = -\frac{1}{\theta_{j_2}} \quad (S53)$$

or

$$m_{j_2}(x) = 1 \text{ and } F_{\theta,\lambda}(x + \zeta_{j_2}) - F_{\theta,\lambda}(x) = -\frac{x_1}{\theta_{j_2}} \quad (S54)$$

We shall later argue that condition (S54) cannot be satisfied at all $x \in \mathbb{N}_0^N$. Hence we proceed under the assumption that (S53) holds. Without loss of generality we can assume that $j_1 = 1$ and $j_2 = 2$ (the other case is symmetric) and $F_{\theta,\lambda}(0) = 0$ where 0 is the $N$-dimensional vector of zeros. Observe that the form of the mass-action monomials $m_1(x)$ and $m_2(x)$ in (S52) and (S53) implies part (A).

Applying the parameter sensitivity formula given by Theorem 3.3 in [14] we can conclude that

$$\partial_{\theta_1}E_{\pi_{\theta,\lambda}}(f) = \sum_{x \in \mathbb{N}_0^N} m_1(x)(F_{\theta,\lambda}(x + \zeta_1) - F_{\theta,\lambda}(x))\pi_{\theta,\lambda}(x) = \frac{1}{\theta_2} \phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right) \quad (S55)$$

and

$$\partial_{\theta_2}E_{\pi_{\theta,\lambda}}(f) = \sum_{x \in \mathbb{N}_0^N} m_2(x)(F_{\theta,\lambda}(x + \zeta_2) - F_{\theta,\lambda}(x))\pi_{\theta,\lambda}(x) = -\frac{\theta_1}{\theta_2^2} \phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right). \quad (S56)$$

However (S52) and (S53) imply that

$$\sum_{x \in \mathbb{N}_0^N} m_1(x)(F_{\theta,\lambda}(x + \zeta_1) - F_{\theta,\lambda}(x))\pi_{\theta,\lambda}(x) = \frac{1}{\theta_1} \phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right) \quad (S57)$$

and

$$\sum_{x \in \mathbb{N}_0^N} m_2(x)(F_{\theta,\lambda}(x + \zeta_2) - F_{\theta,\lambda}(x))\pi_{\theta,\lambda}(x) = -\frac{1}{\theta_2} \phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right). \quad (S58)$$

Equating (S55) and (S57) we get

$$\phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right) = \frac{\theta_1}{\theta_2} \phi_{\text{out}}\left(\frac{\theta_1}{\theta_2}\right).$$

18
which also ensures that (S56) and (S58) are equal. This relation proves that
\[ \phi_{out}(x) = \kappa x \] (S59)
for some \( \kappa > 0 \). We shall soon see this \( \kappa \) is exactly as stated in part (B) of this theorem.

Let \( e_m \) be the \( m \)-th standard basis vector in \( \mathbb{R}^N \) (i.e. its \( m \)-th component is one and the rest are all zeros). As the state-space \( \mathbb{N}_0^N \) is irreducible for the dynamics, for each species \( X_i \) there exists a sequence of reactions \( k_1^{(i)}, \ldots, k_n^{(i)} \) such that there is a positive probability of going from state 0 to state \( e_i \) via these reactions. Due to (S4) we must have
\[ z_j := \sum_{l=1}^{j} \zeta_{k_l^{(i)}} \geq \nu_{k_{j+1}^{(i)}} \quad \text{for each} \quad j = 0, 1, \ldots, n-1 \]
and
\[ z_n := \sum_{l=1}^{n} \zeta_{k_l^{(i)}} = e_i. \]

We can also observe that for any \( z \in \mathbb{N}_0^N \) this same sequence of reactions will also allow the dynamics to go from state \( z \) to state \( (z + e_i) \) with a positive probability (see [6]). Setting \( z_0 = 0 \) we note that
\[ F_{\theta,\lambda}(z + e_i) - F_{\theta,\lambda}(z) = \sum_{j=1}^{n} (F_{\theta,\lambda}(z + z_j) - F_{\theta,\lambda}(z + z_{j-1})) \]
\[ = \sum_{j=1}^{n} \left( F_{\theta,\lambda}(z + z_{j-1} + \zeta_{k_j^{(i)}}) - F_{\theta,\lambda}(z + z_{j-1}) \right). \]

Observe that because of conditions (S47), (S52) and (S53) we obtain
\[ \theta_2(F_{\theta,\lambda}(z + z_{j-1} + \zeta_{k_j^{(i)}}) - F_{\theta,\lambda}(z + z_{j-1})) = \begin{cases} 0 & \text{if } k_j^{(i)} \in \{3, \ldots, K\} \\ \kappa & \text{if } k_j^{(i)} = 1 \\ -1 & \text{if } k_j^{(i)} = 2. \end{cases} \]

Therefore if \( n_i^+ \) and \( n_i^- \) denote the number of instances of reactions 1 and 2 respectively. Then we have
\[ \theta_2(F_{\theta,\lambda}(z + e_i) - F_{\theta,\lambda}(z)) = q_i := \kappa n_i^+ - n_i^- \] (S60)
for any \( z \in \mathbb{N}_0^N \). This allows us to identify an integer \( q_i \) with each species \( X_i \) and it shows that \( F_{\theta,\lambda} \) must be of the form (S51). Proposition S2.5 implies that \( q \) must be orthogonal to the stoichiometric vectors of the disturbance inducing reactions, i.e.
\[ \langle q, \zeta_k \rangle = 0 \quad \text{for} \quad k = 3, \ldots, K. \] (S61)

Moreover since (S59) holds, (S52) and (S53) imply that
\[ \langle q, \zeta_1 \rangle = \kappa \quad \text{and} \quad \langle q, \zeta_2 \rangle = -1. \] (S62)
Taken together, (S2.2) and (S62), prove that the vector $q = (q_1, \ldots, q_N)$ must satisfy the linear system (8).

We now argue that under our assumptions, (S54) cannot hold at each $x \in \mathbb{N}_0^N$. Suppose that this condition indeed holds. Then what changes in the preceding analysis is that now

$$\theta_2(F_{\theta,\lambda}(z + z_{j-1} + \zeta_k^{(i)})) - F_{\theta,\lambda}(z + z_{j-1}) = -f(z + z_{j-1}) \quad \text{if} \quad k^{(i)}_j = 2,$$

where $f$ is the projection map (S16) that maps the state vector to the state of the output species. Note that in this case the l.h.s. of (S60) will depend on the choice of the reaction path, and the intermediate states encountered, in going from state $z$ to state $(z + e_i)$. This violates the fact that the solution $F_{\theta,\lambda}$ of the Poisson equation (S43) is unique up to addition by a constant.

So far we had assumed that $j_1 = 1$ and $j_2 = 2$ in (S52) and (S53). The treatment of the other symmetric case ($j_1 = 2$ and $j_2 = 1$) is similar to above. In this case $\phi_{\text{out}}(x) = (\kappa x)^{-1}$. This proves the “only if” part of the result.

The “if” part of the result follows from ergodicity of the stochastic dynamics along with the presence of an integrator (S51) that satisfies (S21). This completes the proof of this theorem. \qed

### S2.3 A couple of simple lemmas

Recall that $S$ is the $N \times K$ stoichiometric matrix for the reaction network given by (5) and we shall assume that it has full row-rank (6). Next we state and prove a couple of simple lemmas.

**Lemma S2.8** Consider a network satisfying our stability assumption (i.e. Assumption S2.1(A) in the deterministic setting or Assumption S2.4(A) in the stochastic setting). Let $S$ be its stoichiometric matrix satisfying (6). Then there can exist at most one solution $(q, \kappa)$ for which (8) is satisfied and this solution can be found by solving a linear system of the form $Ax = b$.

**Proof.** Let us first assume that we are in the deterministic setting and fix a nominal value of $(\theta^*, \lambda^*) \in \Theta \times \Lambda_u$ for which the steady-state is $x = x_{\theta^*, \lambda^*}$. Let $c = (c_1, \ldots, c_K)$ be the vector of all the propensity functions at this steady-state, i.e.

$$c = (\theta_1^* m_1(x), \theta_2^* m_2(x), \lambda_3^*(x), \ldots, \lambda_K^*(x)).$$

The fixed-point relation (S1) implies that

$$Sc = 0. \quad \text{(S63)}$$

Since the fixed-point $x$ is in the positive orthant and $m_1(x)$ and $m_2(x)$ are mass-action monomials we must have that both $c_1$ and $c_2$ are strictly positive. Now suppose that there are two pairs $(q_1, \kappa_1)$ and $(q_2, \kappa_2)$ satisfying (8). Then we have

$$0 = q_i^T Sc = (\kappa_i, -1, 0, \ldots, 0)c = \kappa_i c_1 - c_2 \quad \text{for} \quad i = 1, 2.$$
Therefore $\kappa_1 = \kappa_2 = c_2/c_1$ and since the matrix $S$ has full row-rank we must also have $q_1 = q_2$. This proves the uniqueness of the solution of the linear system $(8)$ if it exists. Checking this existence is equivalent to checking the existence of a solution to the linear system $Ax = b$ where $x = (q, \kappa)$ is the $(N + 1)$ dimensional column vector of unknowns, $b$ is the $K$ dimensional vector given by $b = (0, -1, 0, \ldots, 0)$ and $A$ is the $K \times (N + 1)$ dimensional matrix obtained by appending $-e_1$ as a column to the transpose of $S$. Here $e_1$ denotes the $N$ dimensional vector with the first component as one and the rest of the components as zeros.

The proof in the stochastic case is similar except that $c$ is the vector of expected propensity functions of all the reactions at the stationary distribution. Relation (S63) can be obtained from (S14) by setting $f(x) = x$, i.e. the identity function of the state vector. □

**Lemma S2.9** Consider a maxRPA network characterised by the solution $(q, \kappa)$ to the linear-algebraic system $(8)$. If (15) holds then $X_1 \notin C_\pm$ and hence the output species $X_1$ belongs to $C_0$ and it is not present in the internal model.

**Proof.** Condition (15) implies the existence of a vector $c = (c_1, \ldots, c_K)$ such that $c_1 = c_2 = 0$ and

$$Sc = e_1$$

where $S$ is the stoichiometry matrix for the network. Multiplying (8) by the $K \times 1$ vector $c$ on the right we see that

$$0 = (\kappa, -1, 0, \ldots, 0)c = q^T Sc = q^T e_1 = q_1.$$ 

Hence $X_1 \notin C_0$ and this completes the proof of this lemma. □

**S2.4 Generic decomposition of stochastic maxRPA networks**

In this section we study the structure of a maxRPA network in the stochastic setting. We restrict ourselves with bimolecular reactions in which all reactions can have at most two reactants i.e. $\sum_i \nu_{ik} \leq 2$ for each reaction $k$ of the form (1). We shall show that in the stochastic setting, a maxRPA network cannot be homothetic but it needs to be antithetic. A key property of such antithetic networks is the existence of a generalised annihilation or sequestration reaction between the internal model (IM) (comprising species in $C_\pm$) and the rest of the network (RoN) (comprising species in $C_0$). This reaction is of the form

$$X_p + X_t \rightarrow C_0^*,$$ 

(S64)

where $X_p \in C_+$, $X_t \in C_-$ and $C_0^*$ denotes any combination of species that belong to the set $C_0$. Note that such a reaction cannot arise in a homothetic network because either $C_+$ or $C_-$ is empty.

**Theorem S2.10 (Generic decomposition of stochastic maxRPA networks)**

Suppose that Assumption S2.4 and condition (15) holds and all reactions are bimolecular. Then for a maxRPA network in the stochastic setting we have the following up to renaming of species $X_2, \ldots, X_N$ and interchanging of reactions 1 and 2:
(A) The maxRPA network must be antithetic.

(B) There is a species \( X_2 \in C_+ \) such that the first reaction is of the form (16).

(C) There is a species \( X_3 \in C_- \), such that the second reaction is of the form (17).

(D) There exists at least one generalised sequestration reaction of the form (14). Moreover reactions of this form are the only possible reactions where the products are species in \( C_0 \) and one of the reactant species is in \( C_\pm \).

Remark S2.11 As mentioned in the main text (see Section 1.4), (16) is the set-point encoding reaction (as its reaction rate \( \theta_1 \) sets the numerator of the set-point) and to (17) as the output sensing reaction (as its reaction rate is \( \theta_2 x_1 \) and \( x_1 \) is the output species copy-number). The encoding and sensing is performed by the internal model (IM) (comprising species in \( C_\pm \)) via production of distinct species \( X_2 \) and \( X_3 \) (the species are renamed if necessary) in the IM. The generic decomposition of stochastic maxRPA networks is depicted in Figure 3.

Proof. We first prove by contradiction that no maxRPA network in the stochastic setting can be homothetic (recall Definition 1.1). Consider a maxRPA network in the stochastic setting characterised by a pair \((q, \kappa)\) that satisfies the linear-algebraic system (8) (see Theorem S2.6(B)). Suppose that \( C_- = \emptyset \) and hence all nonzero components are \( q \) are positive. From Theorem S2.6(A) we know that the second reaction must be of the form \( X_1 \to * \) and its stoichiometry vector \( \zeta_2 \) satisfies

\[
q^T \zeta_2 = -1. \tag{S65}
\]

Note that except for the first coordinate (corresponding to the output species \( X_1 \)) all other components of \( \zeta_2 \) must be non-negative. Since \( q_1 = 0 \) and all other components of \( q \) are non-negative, we see that (S65) cannot hold, resulting in a contradiction. Now we consider the other case that \( C_+ = \emptyset \) and so all nonzero components are \( q \) are negative. Again from Theorem S2.6(A) we know that the first reaction must be of the form \( \emptyset \to * \) and its stoichiometry vector \( \zeta_1 \) satisfies

\[
q^T \zeta_1 = \kappa > 0. \tag{S66}
\]

As all components of \( \zeta_1 \) need to be non-negative and all components of \( q \) are nonpositive, (S66) yields a contradiction. This shows that the maxRPA network cannot be homothetic and proves part (A) of this theorem.

We now prove part (B) with the assumption that the maxRPA network is antithetic and hence \( q \) has both positive and negative components. As all components of \( \zeta_1 \) are non-negative, condition (S66) implies that there exists some species, let us call it \( X_2 \), with \( q_2 > 0 \) and \( \zeta_{21} > 0 \).

\[1\]

Hence \( X_2 \in C_+ \) and the first reaction is of the form (16). This completes the proof of part (B). The proof of part (C) is similar. Lemma S2.9 shows that \( q_1 = 0 \), and so from (S65) we can conclude that there exists some species, let us call it \( X_3 \), with \( q_3 < 0 \) and \( \zeta_{32} > 0 \). Hence \( X_3 \in C_- \) and the second reaction is of the form (17). This completes the proof of part (C).

\[1\] Here \( \zeta_{21} \) denotes the second component of the first stoichiometric vector \( \zeta_1 \)
We now prove part (D). Let \( I_0 \) be the set of addresses of species in \( C_0 \) respectively, i.e.
\[
I_0 = \{ i = 1, \ldots, N : X_i \in C_0 \}.
\]

Define \( S_0 \subset \mathbb{N}_0^N \) to be the set given by
\[
S_0 = \{ x = (x_1, \ldots, x_N) \in \mathbb{N}_0^N : x_i = 0 \text{ for each } i \notin I_0 \}
\]
and so if the state is in \( S_0 \) then all the species in the internal model (i.e. in the set \( C_\pm \)) have zero copy-numbers. Pick any species \( X_m \in C_\pm \). Since the state-space \( \mathcal{E} = \mathbb{N}_0^N \) is irreducible, there exists a sequence of reactions \( k_1, \ldots, k_n \) such that there is a positive probability of going from state \( e_m \) to state \( 0 \) via these reactions, thereby eliminating exactly one molecule of species \( X_m \). For this to hold we must have
\[
z_j := e_m + \sum_{l=1}^{j} \zeta_{k_l} \geq \nu_{k_{j+1}} \text{ for each } j = 0, 1, \ldots, n - 1
\]
and
\[
z_n := e_m + \sum_{l=1}^{n} \zeta_{k_l} = 0.
\]
Note that \( z_n \in S_0 \) and \( z_0 = e_m \notin S_0 \). Let \( k_* \) be the first reaction in this sequence which originates from a state outside \( S_0 \) and leads to a state within \( S_0 \), i.e. \( k_* = k_l \) where
\[
l = \min\{ j = 1, \ldots, n : z_{j-1} \notin S_0 \text{ and } z_j \in S_0 \}.
\]
Observe that \( k_* \) needs to eliminate molecules of some species \( X_p \) in \( C_\pm \) and all its products must be species in \( C_0 \). Hence this reaction \( k_* \) cannot be the first reaction (16) or the second reaction (17). Since all reactions are bimolecular this reaction \( k_* \) must necessarily be of the form
\[
X_p \longrightarrow C_0^*
\]
or of the form
\[
X_p + X_\ell \longrightarrow C_0^*.
\]
As \( q \) satisfies (8), reaction \( k_* \) being of the form (S67) would imply that \( q_p = 0 \) which is a contradiction since \( X_p \) is in \( C_\pm \). Therefore reaction \( k_* \) must be of the form (S68) and in this case condition (8) implies that \( q_p + q_\ell = 0 \) or
\[
q_p = -q_\ell.
\]
Clearly for this to happen one of the species \( X_p \) or \( X_\ell \) belongs to \( C_+ \) and the other belongs to \( C_- \), thereby proving the existence of a generalised antithetic reaction (14). This analysis also shows that all reactions where the products are species in \( C_0 \) and one of the reactants is a species in \( C_\pm \), must necessarily be of the form (14). This completes the proof of this theorem.
\[\square\]
Remark S2.12  Notice that the proof of part (D) of Theorem S2.10 only hinges on the assumption that the state-space $\mathbb{N}_0^N$ is irreducible for the reaction network (see Assumption S2.4(A)). This irreducibility assumption typically holds when the stoichiometry matrix satisfies the full-rank condition (6) (see [6] for more details). Hence for most bimolecular deterministic antithetic maxRPA networks satisfying (15), the assertion of part (D) of this theorem continues to hold, and in particular such networks would have a generalised sequestration reaction of the form (14). If the deterministic maxRPA network is homothetic, this argument cannot be applied as for these networks the state-space $\mathbb{N}_0^N$ will not be irreducible for the corresponding stochastic network.

S3  maxRPA networks with minimal IMs

S3.1  Homothetic networks with singleton IMs

Consider a deterministic maxRPA network whose IM consists of only the species $X_2$, i.e. $\mathcal{C}_\pm = \{X_2\}$. By definition, this network is homothetic and we shall assume that (15) holds. Let $(q, \kappa)$ be the maxRPA characterising pair that satisfies (8). Then only the second component of $q$ (i.e. $q_2$) is non-zero while the rest are all zeros. From (8) we can conclude that the copy-numbers of $X_2$ are not altered by the disturbance inducing reactions (i.e. $\zeta_{2k} = 0$ for $k = 3, \ldots, K$) and for the first two reactions we have

$$n_1 := \zeta_{21} = \frac{\kappa}{q_2} \quad \text{and} \quad n_2 := -\zeta_{22} = \frac{1}{q_2}.$$  

Recall that the non-negative integer $\nu_{ik}$ denotes the number of molecules of species $X_i$ that are consumed as reactants by the $k$-th reaction. Due to Theorem S2.2 the first two reactions, with mass-action kinetics, must have the form

$$n_1 := \zeta_{21} = \frac{\kappa}{q_2} \quad \text{and} \quad n_2 := -\zeta_{22} = \frac{1}{q_2}.$$

(S69)

\[ \nu_1X_1 + \nu_2X_2 + \sum_{i=3}^{N} \nu_{1i}X_i \xrightarrow{\theta_1} (\nu_{21} + n_1)X_2 + \mathcal{C}_0^* \]

and

\[ (\nu_1 + \alpha)X_1 + \nu_2X_2 + \sum_{i=3}^{N} \nu_{1i}X_i \xrightarrow{\theta_2} (\nu_{21} - n_2)X_2 + \mathcal{C}_0^*, \]

(S70)

where $\alpha$ is some non-zero integer satisfying $\alpha \geq -\nu_{11}$ and $\mathcal{C}_0^*$ denotes any combination of species in $\mathcal{C}_0 = \{X_1, X_3, \ldots, X_N\}$. Recall from Remark S2.11, that reaction 1 is the set-point encoding reaction, while reaction 2 is the output sensing reaction. Note that since $\kappa > 0$, both $n_1$ and $n_2$ must have the same sign as $q_2$ and the output set-point encoding is given by

$$\theta^* = \phi_{\text{out}} \left( \frac{\theta_1}{\theta_2} \right) = \sqrt{\frac{\theta_1 n_1}{\theta_2 n_2}}.$$  

S3.1.1  Zero-order degradation networks

Consider the case $\nu_{21} = 0$ in which the lone IM species $X_2$ is not a reactant in either the first reaction (S69) or the second reaction (S70). This implies that exactly one of these
two reactions causes negative production or loss of $X_2$ molecules, violating the reaction form given by (1). The rate of this reaction is not proportional to the abundance of $X_2$ molecules which can lead to negative abundance levels for this species. Such a reaction is called zero-order degradation and it can be approximately realised when the degradation is assisted by an enzyme which operates at saturation [15].

Depending on whether $q_2$ is positive (i.e. $C_+ = \{X_2\}$) or negative (i.e. $C_- = \{X_2\}$), two types of maxRPA networks based on zero-order degradation can be constructed. An example of both these network types is shown in Figure 4(A, B). In both the examples, we set $\nu_{i1} = 0$ for all $i$. The key difference between them is that $q_2$ is positive in one case (panel A) while it is negative in the other (panel B). In the former, the zero-order degradation reaction acts like a sensing reaction while in the latter, it acts like a set-point encoding reaction.

S3.1.2 Autocatalytic networks

Now consider the case $\nu_{21} > 0$ in which the lone IM species $X_2$ is a reactant in both the first reaction (S69) and the second reaction (S70). Hence, irrespective of the sign of $q_2$, $X_2$ stimulates its own production and therefore we call such networks autocatalytic.

In Figure 4(C, D) we present two autocatalytic maxRPA networks that resemble the networks in Figure 4(A, B), except that non-physical zero-order degradation reactions are replaced by autocatalytic reactions. In both the examples we set $\nu_{i1} = 0$ for all $i$ except $i = 2$. As before, the difference between them is that $q_2$ is positive in one case (panel C) while it is negative in the other (panel D). In the former, we choose $\nu_{21} = n_2$ and so reaction 1 (set-point encoding) is autocatalytic, while in the latter we choose $\nu_{21} = -n_1$ and so reaction 2 (output sensing) is autocatalytic.

In the special case of $\nu_{21} = -n_1 = n_2 = 1$, $C^*_0 = X_1$ in (S69) and $C^*_0 = X_0 = \emptyset$ in (S70), these two reactions encode the simplest network exhibiting the property of Absolute Concentration Robustness (ACR), whereby one of the species ($X_1$ in our case) has the same value at any equilibrium value. For more details we refer the readers to the celebrated paper of Shinar and Feinberg [16] which present sufficient structural conditions for networks to exhibit ACR. In a recent paper the existence of a linear constrained integrator for a wide class of ACR networks is shown [17].

S3.2 Antithetic maxRPA networks with dyadic IMs

Consider an antithetic maxRPA network in the deterministic setting whose IM consists of two species $X_2$ and $X_3$, i.e. $C_\pm = \{X_2, X_3\}$. Let $(q, \kappa)$ be the maxRPA characterising pair that satisfies (8). Without loss of generality we can assume that $q_2 > 0$ (i.e. $C_+ = \{X_2\}$) and $q_3 < 0$ (i.e. $C_- = \{X_3\}$). All the other components of the $q$ vector are zero. In this section we shall assume that all reactions are bimolecular i.e. $\sum_i \nu_{ik} \leq 2$ for each reaction $k$ of the form (1).

From (8) we can conclude that for each disturbance inducing reaction $k = 3, \ldots, K$ we must have

$$q_2 \zeta_{2k} = -q_3 \zeta_{3k},$$  \hspace{1cm} (S71)
and so $\zeta_{2k}$ and $\zeta_{3k}$ have the same sign. Hence each disturbance inducing reaction either simultaneously degrades or produces both species $X_2$ and $X_3$. Suppose that one such reaction $k$ causes degradation of both the species, then as this reaction is bimolecular it must be of the generalised sequestration form

$$X_2 + X_3 \longrightarrow C_0^*.$$  \hspace{1cm} (S72)

Since $\zeta_{2k} = \zeta_{3k} = -1$, due to (S71) we have $q_3 = -q_2$. To preserve dynamical stability, such a degradation causing reaction must indeed exist provided one of species $X_2$ or $X_3$ is only produced by the first two reactions, and not degraded.

Let us examine the form of the first two reactions more closely under the restrictions imposed by Theorem S2.2. Without loss of generality we can assume that $\alpha > 0$, as the other case is symmetric. As all reactions are bimolecular, there must exist a species $X_m$ such that the first two reactions, with mass-action kinetics, have the form

$$X_m \xrightarrow{\theta_1} \nu'_{21} X_2 + \nu'_{31} X_3 + C_0^*$$ \hspace{1cm} (S73)

and

$$\alpha X_1 + X_m \xrightarrow{\theta_2} \nu'_{22} X_2 + \nu'_{32} X_3 + C_0^*.$$  \hspace{1cm} (S74)

Here $X_m$ can be any species, including the null species $X_0 = \emptyset$ for $m = 0$. Note that for the second reaction to be bimolecular, $\alpha$ can be 2 only when $m = 0$, and for other values of $m$, $\alpha$ should be 1. We now separate our analysis into two cases, depending on the value of $m$.

**Case 1**: $m \notin \{2, 3\}$: In this case $q_m = 0$ and (8) implies that

$$\nu'_{21} q_2 + \nu'_{31} q_3 = \kappa > 0 \quad \text{and} \quad \nu'_{22} q_2 + \nu'_{32} q_3 = -1.$$

Since $q_2 > 0$ and $q_3 < 0$ we must have that $\nu'_{21} > 0$ and $\nu'_{32} > 0$ which says that the first reaction must produce species $X_2$ while the second reaction must produce species $X_3$. As these two reactions are not degrading any $X_2$ or $X_3$, there must exist a disturbance inducing reaction of the form given by (S72), as we explained earlier, which in turn implies that $q_3 = -q_2$. We illustrate this case with an example shown in Figure 5 where we set $q_2 = -q_3 = \nu'_{21} = \nu'_{32} = \alpha = 1$ and $\nu'_{31} = \nu'_{22} = 0$. Note that $\kappa$ must be 1 for this example and so the set-point is exactly the ratio $\theta_1/\theta_2$. If $m = 0$ this antithetic network is also maxRPA in the stochastic setting (see Theorem S2.6).

**Case 2**: $m \in \{2, 3\}$: Without loss of generality we assume that $m = 2$ as the other case ($m = 3$) is similar. Then (8) implies that

$$(\nu'_{21} - 1) q_2 + \nu'_{31} q_3 = \kappa > 0 \quad \text{and} \quad (\nu'_{22} - 1) q_2 + \nu'_{32} q_3 = -1.$$

As $q_2 > 0$ and $q_3 < 0$, the first reaction must produce at least two molecules of $X_2$ (i.e. $\nu'_{21} \geq 2$). If either $\nu'_{31}$ or $\nu'_{32}$ is positive then species $X_3$ is being produced by these reactions but not being degraded, and so $q_2 = -q_3$ and the two IM species must degrade mutually via a generalised annihilation reaction (S72), giving rise to a similar antithetic structure as in the previous case. On the other hand if $\nu'_{31} = \nu'_{32} = 0$, then species $X_3$ is not getting created by the first two reactions, and we must necessarily have $\nu'_{22} = 0$, implying that reaction 2
causes degradation of $X_2$ molecules, without the need of additional annihilation reactions. In this scenario, species $X_3$ merely acts as a bystander and does not participate in creating the integral action that generates robustness. This integral action is being created only by species $X_2$ as in the case of homothetic autocatalytic networks with a singleton internal model (see Section S3.1).

### S4 Automated discovery of optimal maxRPA controllers

In this section we provide more details on how the analysis in Section 3 of the main paper was conducted. For ease of explanation, we shall restrict ourselves to the gene-expression system considered in Section 3. However this approach for identifying optimal maxRPA controllers can be applied to any network of interest which is to be controlled.

Our aim is to find two-species maxRPA controllers for controlling the gene-expression system which consists of two species – the mRNA $X_4$ and the output protein $X_1$. The two controller species are $X_2$ and $X_3$, and we assume that $X_2$ acts like a transcription factor that ‘actuates’ the gene-expression system by catalysing the production of mRNA

$$X_2 \rightarrow X_2 + X_4.$$  \hfill (S75)

Furthermore, the gene-expression network consists of the following three reactions:

- **Protein translation**
  $$X_4 \rightarrow X_4 + X_1$$ \hfill (S76)

- **mRNA degradation**
  $$X_4 \rightarrow \emptyset$$ \hfill (S77)

- **Protein degradation**
  $$X_1 \rightarrow \emptyset$$ \hfill (S78)

With the actuation reaction fixed as (S75) and the gene-expression reactions fixed as (S76), (S77) and (S78), our goal is to find maxRPA controllers that will consist of additional reactions involving the controller species $X_2$ and $X_3$ as well as the output species $X_1$. Note that controller reactions are not allowed to involve the mRNA $X_4$.

Observe that each reaction (1) can be expressed as $\nu \rightarrow \nu'$, where $\nu$ (resp. $\nu'$) is a non-negative integer vector whose components denote the number of molecules of each species consumed (resp. produced) by the reaction. We only allow doubly bimolecular reactions (i.e. $\sum_i \nu_i \leq 2$ and $\sum_i \nu'_i \leq 2$). So the controller can only involve four species – viz. the two controller species ($X_2$ and $X_3$), the output species ($X_1$) and the null species ($\emptyset$) as reactants or products. We construct all possible bimolecular reactant (or product) combinations by generating all possible combinations of two species out of the four species, and this gives us
the set $\mathcal{B}$ of size $|\mathcal{B}| = 10$. The set of all possible bimolecular reactions $\mathcal{B}_2$ can be constructed by selecting all possible ordered pairs $(\nu, \nu')$ (without replacement) from the set $\mathcal{B}$. Therefore there are $|\mathcal{B}_2| = 10 \times 9 = 90$ bimolecular reactions, and each maxRPA controller would be a subset of $\mathcal{B}_2$. Observe that for any $(\nu, \nu') \in \mathcal{B}_2$ the corresponding stoichiometric vector is simply $\zeta = \nu' - \nu$.

Rather than going through all possible subsets of $\mathcal{B}_2$ (which are $2^{90} \geq 10^{27}$ in number) and checking each for maxRPA, we exploit our characterisation results to systematically parse through the set of maxRPA designs. For this we define the set of possible bimolecular first-two reactions as

$$\mathcal{F} = \{(\nu_1, \nu'_1), (\nu_2, \nu'_2) : \nu_1 = e_i, \nu_2 = e_1 + e_i \text{ for some } i = 0, 1, 2, 3 \text{ and } \nu'_1, \nu'_2 \in \mathcal{B}\},$$

where $e_i$ is the vector of all zeros with 1 at the $i$-th coordinate, and $e_0$ is just the vector of all zeros. Hence the conditions on $\nu_1$ and $\nu_2$ imply that the first reaction (set-point encoding) is of the form $\mathbf{X}_i \rightarrow \ast$ and the second reaction (sensing) is of the form $\mathbf{X}_1 + \mathbf{X}_i \rightarrow \ast$, where $\mathbf{X}_i$ is either a controller species, or the output species or the null species $\emptyset$ (for $i = 0$). This mimics the conditions on the first-two reactions imposed by Theorem S2.2 for deterministic maxRPA networks with the additional constraint that the first two reactions are bimolecular and they do not involve $\mathbf{X}_4$. In the special case of $i = 0$, the equivalent restriction imposed by Theorem S2.6 is also satisfied for stochastic maxRPA networks. We then construct the set of possible charge vectors $q$ that satisfy the linear system (8). Note that each such vector must be of the form $q = (0, q_1, q_2, 0)$ where $q_1 \neq 0$ and $q_2 \neq 0$. Allowing (8) to be scaled by some positive constant if necessary, we can fix $q_2$ to be either $+1$ (i.e. species $\mathbf{X}_2$ has a positive charge) or $-1$ (i.e. species $\mathbf{X}_2$ has a negative charge). In both the scenarios we let $q_3$ be a rational number whose numerator and denominator is any pair chosen from $I_k = \{-(k+1), -k, \ldots, -1, 1, \ldots, k+1\}$ where $k$ is a positive integer which we fix as 2 in our case. Hence the set of admissible $q$-vectors is

$$\mathcal{Q} = \left\{ q = (0, q_2, q_3, 0) : q_2 = \pm 1 \text{ and } q_3 = \frac{p_1}{p_2} \text{ for any } p_1, p_2 \in I_k \right\}.$$

The rationale behind this choice for the set of $q$-vectors is that $q$ satisfies a linear system of the form $Ax = b$, where $A$ and $b$ have integer entries (see Lemma S2.8), and hence the components of $q$ must be rational numbers. We can expand the set $\mathcal{Q}$ of possible charge vectors by increasing $k$, but for our example it does not lead to an increase in the number of maxRPA networks.

For each $q \in \mathcal{Q}$, the set of compatible first-two reactions is given by

$$\mathcal{F}_q = \{((\nu_1, \nu'_1), (\nu_2, \nu'_2)) \in \mathcal{F} : \langle q, \nu'_1 - \nu_1 \rangle > 0 \text{ and } \langle q, \nu'_2 - \nu_2 \rangle < 0\},$$

which basically says that the stoichiometric vectors for the first two reactions satisfy the first two equations in (8) up to positive scaling. Here $\langle \cdot, \cdot \rangle$ denotes the standard inner product on the Euclidean space. The set of compatible bimolecular controller reactions for this $q$ is given by

$$\mathcal{C}_q = \{(q, \nu') \in \mathcal{B}_2 : \langle q, \nu' - \nu \rangle = 0 \text{ and } \text{supp}(\nu) \cup \text{supp}(\nu') \subset \{1, 2, 3\}\},$$

28
where \( \text{supp}(\nu) = \{i : \nu_i > 0\} \) denotes the support of vector \( \nu \). Observe that this is the set of all possible disturbance inducing reactions whose stoichiometric vectors are orthogonal to \( q \) as mandated by the last \((K-2)\) equations in (8). The condition on the support ensures that the controller does not involve the mRNA \( X_4 \).

The set of all plausible maxRPA controllers corresponding to charge vector \( q \) is given by the product space \( \mathcal{F}_q \times \mathcal{P}(\mathcal{C}_q) \), where \( \mathcal{P}(\mathcal{C}_q) \) is the power-set of \( \mathcal{C}_q \) consisting of all possible subsets of \( \mathcal{C}_q \). By iterating over elements in this product space we can explore the whole space of maxRPA controllers and identify the best designs. We found that out of more than \( 10^{27} \) possible controllers, the number of maxRPA controllers was less than 10000. A large majority of them (\( \geq 95\%) \) were homothetic, while the rest were antithetic. The antithetic designs that enabled stochastic maxRPA were only around 50 in number.

The full Python code for carrying out this analysis can be downloaded from the GitHub repository: [https://github.com/ankitgupta83/maxRPA.git](https://github.com/ankitgupta83/maxRPA.git). This implementation numerically verifies the stability of the closed-loop network, along with other conditions such as full row-rank of the stoichiometric matrix and strict positivity of the components of the fixed point.

Observe that in Figure 9(B) corresponding to a deterministic (but not stochastic) antithetic maxRPA controller, there is a bias in the output mean w.r.t. the set-point. To understand the source of this error, let us consider the linear integrator (S51) for stochastic maxRPA. If \((X_{\theta,\lambda}(t))_{t \geq 0}\) is the stochastic dynamics with generator \( A_{\theta,\lambda} \) (see Section S2.2), then for the integrator \( z_{\theta,\lambda}(t) = \mathbb{E}(F(X_{\theta,\lambda}(t))) \) its time-derivative is given by

\[
\dot{z}_{\theta,\lambda}(t) = \mathbb{E}(A_{\theta,\lambda}F_{\theta,\lambda}(X_{\theta,\lambda}(t))) = \frac{\theta_1}{\theta_2}\mathbb{E}(m_1(X_{\theta,\lambda}(t))) - \mathbb{E}(m_2(X_{\theta,\lambda}(t))),
\]

where for \( k = 1, 2 \), \( m_k(x) \) is the combinatorial factor (S41) for the mass-action kinetics for the first two reactions. Here we have used the fact that \( q \) satisfies the linear system (8).

Note that at steady-state, we would have \( \lim_{t \to \infty} \dot{z}_{\theta,\lambda}(t) = 0 \) and this implies that

\[
\lim_{t \to \infty} \frac{\mathbb{E}(m_2(X_{\theta,\lambda}(t)))}{\mathbb{E}(m_1(X_{\theta,\lambda}(t)))} = \frac{\theta_1}{\theta_2} \tag{S79}
\]

In the case of deterministic dynamics, the expectation operator \( \mathbb{E} \) is absent and so we can cancel out the common factors between \( m_1(x) \) and \( m_2(x) \) (see (7)) and this would prove that the output reaches the stipulated set-point. However due to the presence of the expectation operator, and lack of independence between the species’ copy-numbers, we cannot cancel the common factors and this manifests in output bias. For example, if the first reaction is of the form \( X_2 \to * \) and the second reaction is of the form \( X_1 + X_2 \to * \) then \( m_1(x) = x_2 \) and \( m_2(x) = x_1x_2 \). So if \( X_1(t) \) and \( X_2(t) \) are the first two components of \((X_{\theta,\lambda}(t))_{t \geq 0}\), then (S79) becomes

\[
\lim_{t \to \infty} \frac{\mathbb{E}(X_1(t)X_2(t))}{\mathbb{E}(X_2(t))} = \frac{\theta_1}{\theta_2},
\]

which does not imply that

\[
\lim_{t \to \infty} \mathbb{E}(X_1(t)) = \frac{\theta_1}{\theta_2}, \tag{S80}
\]
unless $X_1(t)$ and $X_2(t)$ are independent (which is generally not the case). Of course if condition (12) is satisfied then the first reaction is zero-order $\emptyset \rightarrow *$ (so $m_1(x) = 1$) and the second reaction is first-order $X_1 \rightarrow *$ (so $m_2(x) = x_1$) then hence (S80) holds.

**Supplementary References**

[1] Mustafa H Khammash. Perfect adaptation in biology. *Cell Systems*, 12(6):509–521, 2021. S1

[2] D.A. Anderson and T.G. Kurtz. Continuous time Markov chain models for chemical reaction networks. In H. Koeppl, G. Setti, M. di Bernardo, and D. Densmore, editors, *Design and Analysis of Biomolecular Circuits*. Springer-Verlag, 2011. S1.1, S1.1, S1.1

[3] S. N. Ethier and T. G. Kurtz. *Markov processes : Characterization and Convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. S1.1, S2.2

[4] Ankit Gupta, Corentin Briat, and Mustafa Khammash. A scalable computational framework for establishing long-term behavior of stochastic reaction networks. *PLoS Comput Biol*, 10(6):e1003669, 06 2014. S1.1, S1.1, S1.1, S2.2

[5] Sean P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–548, 1993. S1.1, S1.1, S1.1, S2.2

[6] Ankit Gupta and Mustafa Khammash. Computational identification of irreducible state-spaces for stochastic reaction networks. *SIAM Journal on Applied Dynamical Systems*, 17(2):1213–1266, 2018. S1.1, S2.2, S2.12

[7] Bruce A Francis and William M Wonham. The internal model principle for linear multivariable regulators. *Applied mathematics and optimization*, 2(2):170–194, 1975. S1.3

[8] Eduardo D Sontag. Adaptation and regulation with signal detection implies internal model. *Systems & control letters*, 50(2):119–126, 2003. S1.3

[9] Corentin Briat. A biology-inspired approach to the positive integral control of positive systems: The antithetic, exponential, and logistic integral controllers. *SIAM Journal on Applied Dynamical Systems*, 19(1):619–664, 2020. S1.3

[10] Sean Meyn. *Control techniques for complex networks*. Cambridge University Press, Cambridge, 2008. S1.3, S2.2

[11] Peter W Glynn and Sean P Meyn. A liapounov bound for solutions of the poisson equation. *The Annals of Probability*, pages 916–931, 1996. S1.3, S2.2, S2.2

[12] Thomas G. Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *The Journal of Chemical Physics*, 57(7):2976–2978, 1972. S2
[13] Loring W. Tu. *The Rank of a Smooth Map*, pages 105–117. Springer New York, New York, NY, 2008. S2.1

[14] Ankit Gupta, Muruhan Rathinam, and Mustafa Khammash. Estimation of parameter sensitivities for stochastic reaction networks using tau-leap simulations. *SIAM Journal on Numerical Analysis*, 56(2):1134–1167, 2018. S2.2, S2.2

[15] Xiao Yu Ni, Tormod Drengstig, and Peter Ruoff. The control of the controller: molecular mechanisms for robust perfect adaptation and temperature compensation. *Biophysical journal*, 97(5):1244–1253, 2009. S3.1.1

[16] Guy Shinar and Martin Feinberg. Structural sources of robustness in biochemical reaction networks. *Science*, 327(5971):1389–1391, 2010. S3.1.2

[17] Daniele Cappelletti, Ankit Gupta, and Mustafa Khammash. A hidden integral structure endows absolute concentration robust systems with resilience to dynamical concentration disturbances. *Journal of the Royal Society Interface*, 17(171):20200437, 2020. S3.1.2