Noise Control for DNA Computing

Tomislav Plesa¹, Konstantinos C. Zygalakis², David F. Anderson³, Radek Erban¹

Abstract

Synthetic biology is a growing interdisciplinary field of science and engineering that aims to design biochemical systems that behave in a desired manner. With the advancement of strand-displacement DNA computing, a large class of abstract biochemical networks may be physically realized using DNA molecules. Methods for systematic design of the abstract systems with prescribed behaviors have been predominantly developed at the (less-detailed) deterministic level. However, stochastic effects, neglected at the deterministic level, are increasingly found to play an important role in biochemistry. In such circumstances, methods for controlling the intrinsic noise in the system are necessary for a successful network design at the (more-detailed) stochastic level. To bridge the gap, the noise-control algorithm for designing biochemical networks is developed in this paper. The algorithm structurally modifies any given reaction network under mass-action kinetics, in such a way that (i) controllable state-dependent noise is introduced into the stochastic dynamics, while (ii) the deterministic dynamics are preserved. The capabilities of the algorithm are demonstrated on a production-decay reaction system, and on an exotic system displaying bistability. For the production-decay system, it is shown that the algorithm may be used to redesign the network to control the stochastic switching, and achieve noise-induced oscillations.

1. Introduction

Synthetic biology is an interdisciplinary field of science and engineering that aims to construct biochemical systems with prescribed behaviors [1, 2]. At the theoretical level, the synthetic systems may significantly enhance our understanding of biology. At the practical level, they may have broad applications, e.g. in medicine [3, 4, 5, 6, 7, 8], industry [9, 10], and nanotechnology [11, 12]. The systems may also be of interest to NASA for optimizing extraterrestrial explorations [13]. A proof-of-concept for synthetic biology is a synthetic oscillator called the repressilator, which was implemented in vivo [14]. The experimental advances since the repressilator range from isolated synthetic biochemical networks, to microorganisms containing partially, or even fully, synthetic DNA molecules (synthetic life) [15, 16, 17, 18]. Examples include microorganisms containing a synthetic bistable switch [19], and a cell-density controlling quorum sensor [20], microorganisms producing antimalarial drugs [7, 8], and synthetic systems designed for tumor detection, diagnosis and adaptive drug-response [4, 5].

The construction of biochemical networks in synthetic biology may be broken down into two steps: firstly, an abstract system is constructed, displaying prescribed properties, and taking the form of a chemical reaction network [21, 22, 23]. Secondly, the abstract network is mapped to a suitable physical network, which may then be integrated into a desired environment (e.g. a test-tube, or a living cell) [24].

In the first step of network construction, the goal is to obtain an abstract network with desired dynamics. In this paper, we consider two dynamical models of reaction networks under mass-action kinetics [25, 23]: the deterministic model, and the stochastic model (see Methods for more details). The deterministic model takes the form of the reaction rate equations, which are ordinary-differential equations governing the time-evolution of the species concentrations [23, 25]. The stochastic model takes the form of a Markov chain, which may be simulated using the Gillespie stochastic simulation algorithm [26]. The Gillespie algorithm generates noisy copy-number time-series, with the copy-number distribution matching that obtained from the underlying chemical master equation [23, 25, 26, 27]. The stochastic model is more-detailed, taking into an account the discreteness of the species counts, and the stochastic nature of the dynamics, which may be particularly important in biochemistry, where reaction networks may contain low-abundance species [31, 32, 14, 19, 22, 28, 29, 30]. On the other hand, the deterministic model is less-detailed, and more appropriate when the species are in high-abundance, and the discreteness and stochasticity are negligible [33].

In the second step of network construction, the goal is to engineer a physical network whose dynamics match well the dynamics of a given abstract network, over a suitable time-interval. Engineering an appropriate physical network may proceed indirectly, by altering a preexisting physical network, or directly, by engineering a network, which involves a given set of physical species, from scratch. The advantage of the...
Algorithm 1 which maps an input reaction network to output was implemented in vitro [34].

We now wish to structurally modify network (1) in such a way that the deterministic model from (2) is preserved, while preserving the desired deterministic dynamics. The algorithm may also be used to adjust the intrinsic noise to favorably interact with environment-induced effects (e.g. extrinsic noise).

The rest of the paper is organized as follows: we introduce Algorithm 1 by applying it to the test network (1), which at the deterministic level displays a globally attracting equilibrium point. We show that the algorithm can favorably modify the stationary probability distribution underlying (1) at arbitrary points of the state-space, without influencing the deterministic dynamics. For example, it is shown that the algorithm may be used to redesign (1) to achieve noise-induced multimodality (multistability). We then apply Algorithm 1 to the exotic network (11), which at the deterministic level displays a bistability involving an equilibrium point and a limit cycle. The algorithm is used to redesign (11) to increase the stochastic switching between the two attractors, and to achieve noise-induced oscillations.

2. A One-species Regular System

Consider the one-species production-decay reaction network \( \mathcal{R}(s) \), given by (1). 

\[
\mathcal{R}(s) : \quad \emptyset \xrightarrow{k_1} s, \quad \frac{ds}{dt} = k_1 - k_2 s, \quad s \xrightarrow{k_2} \emptyset, \quad (1) \quad \hat{x}(0) = \hat{x}_0. \quad (2)
\]

Species \( s \) from network (1) reacts according to the two reactions with rate coefficients \( k_1, k_2 \in \mathbb{R}_\geq \), where \( \mathbb{R}_\geq \) is the set of nonnegative real numbers, and \( \emptyset \) is the zero-species (denoting species which are not of interest). In this paper, we assume reaction networks are under mass-action kinetics, with the reactions taking place in unit-volume reactors. Let us denote the concentration of species \( s \) from (1) at time \( t \in \mathbb{R}_\geq \) by \( \hat{x} = \hat{x}(t) \in \mathbb{R}_\geq \). The initial value problem for the deterministic model (also called the drift) for network (1) is given by system (2), with \( \hat{x}_0 \geq 0 \) (see also Methods). Since the deterministic model (2) has a globally attracting equilibrium point, given by \( k_1/ k_2 \), network (1) is said to be regular [23].

Let us denote the copy-number of species \( s \) from (1) at time \( t \geq 0 \) by \( \hat{X}(t) \in \mathbb{N}_0 \), where \( \mathbb{N}_0 \) is the set of integers. Under the stochastic model, \( \hat{X}(t) \) is modelled as a continuous-time, discrete-space Markov chain (see also Methods), which can be generated by using the Gillespie stochastic simulation algorithm [26]. Given \( \hat{X}(t) \), there will be a mean interevent time until one of the reactions from (1) fires. The mean interevent time is given by \( 1/\alpha(X(t)) \), and when the event takes place, the probability that the i-th reaction from (1) fires is equal to \( \alpha_i(X(t))/\alpha(X(t)) \), for \( i \in \{1,2\} \). Here, \( \alpha_1 = k_1 \), and \( \alpha_2(x) = k_2 x \), are the so-called propensity functions of the first, and second, reactions from (1), respectively. Function \( \alpha(x) = k_1 + k_2 x \) is the total propensity function of network (1), i.e. the sum of propensity functions of all the underlying reactions.

We now wish to structurally modify network (1) in such a way that the deterministic model from (2) is preserved,
while an arbitrary nonnegative function, defined on a bounded discrete domain, is added to the total propensity function of (1). The latter requirement implies that the interevent time would be controllably decreased in a state-dependent manner. Equivalently, the two requirements imply that a controllable state-dependent noise would be introduced into the stochastic dynamics. We have designed a three-step algorithm, given as Algorithm 1, which achieves such goals for arbitrary reaction networks under mass-action kinetics. Let us describe properties of the algorithm by applying it on network (1).

Firstly, we wish to introduce an additional species $\bar{s}$ into network (1), in such a way that species $s$ and $\bar{s}$ satisfy a pairwise stoichiometric conservation law. Secondly, we require that the enlarged network has the same deterministic model as network (1), despite the added species $\bar{s}$, which may be achieved by adding another auxiliary species. More precisely, let us consider network $\mathcal{R}^1(s, \bar{s}) \cup \mathcal{R}^2(\bar{s})$, given by:

$$\mathcal{R}^1(s, \bar{s}) : \begin{cases} \bar{s} + l^1 \xrightarrow{k_1} s + l^1, \\
 s \xrightarrow{k_2} \bar{s}, \\
 \mathcal{R}^2(\bar{s}) : \emptyset \xrightarrow{1/\mu} l^1, \\
 \bar{s} + l^1 \xrightarrow{1/\mu} \bar{s}. 
\end{cases}$$

(3)

Species $s, \bar{s}, l^1$ from (3) react according to the four reactions with rate coefficients $k_1, k_2, 1/\mu \in \mathbb{R}_>$. Network $\mathcal{R}^1 = \mathcal{R}^1(s, \bar{s})$, given in (3), is obtained from network $\mathcal{R} = \mathcal{R}(s)$, given by (1), in the following way: since the first reaction in $\mathcal{R}$ increases copy-number of $s$ by one, $\bar{s}$ and $l^1$ are added to the reactants of the reaction, and $l^1$ is added to the products, leading to the first reaction in $\mathcal{R}^1$. Since the second reaction in $\mathcal{R}$ decreases copy-number of $s$ by one, $\bar{s}$ is added to the products, leading to the second reaction in $\mathcal{R}^1$. This ensures that the desired conservation law holds. The superscript in $l^1$ indicates that species $l^1$ is involved as a catalyst in a reaction of $\mathcal{R}^1$ in which $s$ is increased by one. The subscript in $\mathcal{R}^2 = \mathcal{R}^2(\bar{s})$ indicates that the network describes production and decay of $l^1$.

The initial value problem for the deterministic model of (3) is given by

$$\begin{align*}
\frac{dx}{dt} &= k_1(c-x)y - k_2x, \\
\frac{dy}{dt} &= \frac{1}{\mu} (1-(c-x)y), \\
\alpha(0) &= x_0, \\
\beta(y) &= y_0,
\end{align*}$$

(4)

where $x = x(t) \in [0, c] \cap \mathbb{R}_>$, and $y = y(t) \in \mathbb{R}_>$, are the concentrations of species $s$, and $l^1$, from (3), respectively, with $x_0, y_0, c \in \mathbb{R}_>$. We have used the kinetic conservation law $\bar{x}(t) = c - x(t)$, where $\bar{x}(t)$ is the concentration of species $\bar{s}$, and $c < \infty$ is a time-independent conservation constant. Note that the conservation law truncates $x$-state-space. Let us now describe relationships between systems (2) and (4), starting with the weak statement: for $c > k_1/k_2$, and for any $\mu \geq 0$, solutions of (2) and (4) are the same in the long-time limit $t \to \infty$. More precisely, the $x$-component of the equilibrium point of (4) is identical to the equilibrium point of (2), and both are stable. In Supplementary Information (SI) Text, we justify the strong statement: for sufficiently large $c$, and for $\mu \ll 1$, solutions of (2) and (4), with the same initial conditions, are approximately the same at each time $t \geq 0$. For these reasons, we call $\mathcal{R}^1$ a drift-corrector network.

2.1 Zero-Drift Network $\mathcal{R}^3_{1,1}$

Having completed the first two steps, let us focus on the third (and final) step, in which we introduce arbitrary noise into the stochastic model of (3), without influencing the deterministic model (4). Let us start our consideration by embedding into (3) network $\mathcal{R}^3_{1,1} = \mathcal{R}^3_{1,1}(s, \bar{s})$, which is given by

$$\begin{align*}
\mathcal{R}^3_{1,1}(s, \bar{s}) : & \quad s + \bar{s} \xrightarrow{k_1} 2s, \\
& \quad s + \bar{s} \xrightarrow{k_1} 2\bar{s}.
\end{align*}$$

(5)

The subscript in $\mathcal{R}^3_{1,1}$ indicates that the underlying reactions have one molecule of $s$, and one of $\bar{s}$, as reactants. The two reactions in (5) preserve the conservation law from (3). Furthermore, they fire with the same rates, in the first reaction leading to a unit-production, while the second to a unit-decay, of species $s$. Consequently, embedding $\mathcal{R}^3_{1,1}$ into (3) does not affect the underlying deterministic model (4), and we call $\mathcal{R}^3_{1,1}$ a zero-drift network. However, $\mathcal{R}^3_{1,1}$ does affect the underlying stochastic model [36, 37, 38, 23]. To illustrate this, let us consider network $\mathcal{R}^3_{1,1}$ in isolation: the reactions from (5) fire when $X(t) \in (0, C)$, but not when $X(t) \in \{0, C\}$, so that $\mathcal{R}^3_{1,1}$ in isolation fires until $X(t)$ takes one of the extreme values $\{0, C\}$. Here, $X(t) \in \mathbb{N}_0$, and $C \in \mathbb{N}$, $C < \infty$, are the copy-number of species $s$ appearing in (3) and (5) at time $t \geq 0$, and the conservation constant, respectively. Let us note that a possible biologically-relevant realization of network (5), aside from DNA strand-displacement mechanism, is a dimer version of the bifunctional histidine kinase/phosphatase reported in [39].

In SI Text, we derive equation (SI7) which describes the effective behavior of the Markov chain $X(t)$ from network $\mathcal{R}^1 \cup \mathcal{R}^2 \cup \mathcal{R}^3_{1,1}$ in the limit $\mu \to 0$, and it follows that the effective total propensity function of the network, denoted $\alpha(x)$, satisfies

$$\alpha(x) \approx \alpha(x) + 2K_{1,1}\beta_{1,1}(x), \quad \mu \to 0, \quad \alpha(x) = k_1 + k_2x.$$  

(6)

(7)

Function $\alpha(x)$ has the form of the total propensity of network (1), and $K_{1,1}\beta_{1,1}(x)$ is the propensity function of reactions in (5), with the scaled factors given by

$$K_{1,1} = \left(\frac{C}{2}\right)^2 k_{1,1}, \quad \beta_{1,1}(x) = \left(\frac{C}{2}\right)^{-2} x(C-x).$$

(8)
Function $\beta_{1,1}(x)$ is displayed in Figure 1(a), where one can notice its parabolic shape, arising from the underlying conservation law $X(t) + \tilde{X}(t) = C$, which holds for all $t \geq 0$, where $\tilde{X}(t) \in \mathbb{N}_0$ is the copy-number of $\tilde{S}$ at time $t \geq 0$. Comparing (6) and (7), it follows that, as $\mu \to 0$, the mean interevent time for $X(t)$, from network $\mathcal{R}_{1,1}^3 \cup \mathcal{R}_{2,1}^3 \cup \mathcal{R}_{1,1}^3$, is lower than that of $\tilde{X}(t)$, from network (1), in the regions of the common state-space where $\beta_{1,1}(x) \neq 0$, i.e. for $x \in (0, C)$. Coefficient $K_{1,1}$ controls how much the interevent time is reduced. Equivalently, $\beta_{1,1}(x)$, and $K_{1,1}$, determine the support, and magnitude, respectively, of the state-dependent intrinsic noise which network (5) introduces into the dynamics of network (3).

To study this further, in SI Text we derive the following two equations (given as (S19), and (S113), respectively)

$$\lim_{k_{1,1} \to 0} p(x) \approx \begin{cases} \frac{1}{\pi} \left( \frac{k_1}{k_2} \right)^x \exp \left(-\frac{k_1}{k_2}x\right), & \text{if } x \in [0, C], \\ 0, & \text{otherwise}, \end{cases} \tag{9}$$

$$\lim_{k_{1,1} \to \infty} p(x) \approx \begin{cases} 1 - \frac{k_1}{k_2}, & \text{if } x = 0, \\ \frac{k_1}{k_2}, & \text{if } x = C, \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

where $p(x)$ is the stationary probability mass function (PMF) corresponding to network $\mathcal{R}_{1,1}^3 \cup \mathcal{R}_{2,1}^3 \cup \mathcal{R}_{1,1}^3$ in the limit $\mu \to 0$, i.e. the probability that there are $x$ molecules of species $S$ as $\mu \to 0$ in the long-time limit $t \to \infty$. Let us interpret analytical results (9) and (10), and compare them with the numerically obtained counterparts. In Figure 1(b), we display numerically obtained stationary $x$-marginal PMFs for different values of $K_{1,1}$, with the rest of the (dimensionless) parameters fixed to $k_1 = 2.5$, $k_2 = 0.5$, $\mu = 10^{-3}$, and $C = 15$. It can be seen that, for $K_{1,1} = 0$, i.e. when the zero-drift network $\mathcal{R}_{1,1}^3$ does not fire, the PMF matches that of network (1), i.e. it is a Poissonian, as predicted by (9). Let us note that the matching of the PMFs of networks (1) and $\mathcal{R}_{1,1}^3 \cup \mathcal{R}_{2,1}^3 \cup \mathcal{R}_{1,1}^3$ relies on choosing sufficiently large rate coefficients $1/\mu$ in the drift-corrector network $\mathcal{R}_{1,1}^3$. When $K_{1,1} = 5$, the PMF appears closer to a uniform distribution, than does the PMF when $K_{1,1} = 0$. Finally, for the larger value $K_{1,1} = 10^5$, i.e. when zero-drift network $\mathcal{R}_{1,1}^3$ fires much faster than network $\mathcal{R}_{1}^3$, the PMF redistributes across the domain, accumulating at the boundary, and becoming bimodal. This is in qualitative agreement with (6), and in quantitative agreement with (10), which predicts $p(0) \approx 0.7$ and $p(15) \approx 0.3$. In Figure 1(c), a representative sample path is shown, obtained by applying the Gillespie algorithm on network $\mathcal{R}_{1}^3 \cup \mathcal{R}_{2}^3 \cup \mathcal{R}_{1,1}^3$, when $K_{1,1} = 10^5$. Also shown is a trajectory obtained by numerically solving the deterministic model (4). Consistent with Figure 1(b), the sample path switches between the boundary of the state-space, with a bias towards the left boundary point $x = 0$. This is in contrast to the deterministic trajectories, which are globally attracted to the equilibrium point $x = 5$.

### 2.2 General Zero-Drift Networks $\mathcal{R}_{n,\bar{n}}^3$

Zero-drift network $\mathcal{R}_{n,\bar{n}}^3(x, \tilde{x})$, given by (5), involves a single molecule of $S$ and $\tilde{S}$ as reactants, and adds the noise at $x \in [1, C - 1]$, i.e. in the interior of the state-space. Similar networks may be used to add the noise at any point in the state-space, without influencing the deterministic dynamics. In particular, in (19) and (20), we present general zero-drift networks $\mathcal{R}_{n,\bar{n}}^3(x, \tilde{x})$, which involve $n$ molecules of $S$, and $\bar{n}$ of $\tilde{S}$, as reactants, and add the noise at $x \in [\bar{n}, C - \bar{n}]$, where $n, \bar{n} \in \mathbb{N}_0$, and $(n + \bar{n}) \leq C$ (see also SI Text). Embedding a union of such networks, $\cup_{(n,\bar{n})} \mathcal{R}_{n,\bar{n}}^3(x, \tilde{x})$, into (3), we arrive at the result similar to (6), with $K_{1,1} \beta_{1,1}(x)$ replaced by the linear combination $\sum_{(n,\bar{n})} K_{n,\bar{n}} \beta_{n,\bar{n}}(x)$. The scaled rate coefficient $K_{n,\bar{n}}$, and function $\beta_{n,\bar{n}}(x)$, are given as (S14), and (S15), respectively, in SI Text, where we also justify that an arbitrary nonnegative function, with compact support, may be approximated by a suitable sum $\sum_{(n,\bar{n})} K_{n,\bar{n}} \beta_{n,\bar{n}}(x)$. To illustrate general zero-drift networks, let us start with embedding into network (3) zero-drift network $\mathcal{R}_{n,\bar{n}}^3(x, \tilde{x})$, satisfying (19) with $n = 5$ and $\bar{n} = 10$. In Figure 1(d), we show propensities of $\mathcal{R}_{n,\bar{n}}^3(x, \tilde{x})$ for $x = 5$. In (e), we show the numerically approximated stationary $x$-marginal PMFs underlying network $\mathcal{R}_{1}^3 \cup \mathcal{R}_{2}^3 \cup \mathcal{R}_{5,10}^3$ for different values of $K_{3,10}$, with the rest of the parameters as in Figure 1(b).

One can notice that, under the action of network $\mathcal{R}_{3,10}^3$, the PMF is gradually decreased to nearly zero at $x = 5$ (the deterministic equilibrium), and becomes bimodal, with the two noise-induced maxima at $x = 4$ and $x = 6$. In (f), we show a corresponding representative sample path.

In general, noise-induced multimodality may be achieved by a suitable combination of zero-drift networks. For example, let us synthetize noise such that the stationary PMF is trimodal, and nearly zero everywhere, except at $x = \{1, 7, 11\}$. Such a task may always be achieved by a suitable combination of the basis zero-drift networks, i.e. those zero-networks that induce noise only at a single point in the state-space (e.g. subnetwork $\mathcal{R}_{5,10}^3$ with propensity function shown in Figure 1(d), see also SI Text). In the present case, one could construct the thirteen basis zero-drift networks which add large enough noise at $x \in [0, 15] \setminus \{1, 7, 11\}$. Here, for simplicity, we achieve the task with only four zero-drift networks. In Figures 1(g)–(i), we consider network $\mathcal{R}_{1}^3 \cup \mathcal{R}_{2}^3 \cup \mathcal{R}_{5,15}^3 \cup \mathcal{R}_{9,5}^3 \cup \mathcal{R}_{12,0}^3$. We denote $\beta(x) \equiv \beta_{15,15}(x) + \beta_{9,9}(x) + \beta_{8,5}(x) + \beta_{12,0}(x)$, and, for simplicity, take $K = K_{15,15} = K_{9,9} = K_{8,5} = K_{12,0}$. The resultant propensity function $\beta(x)$ is shown in (g), while in (h) it can be seen that the PMF becomes trimodal for sufficiently large $K$, with the maxima at $x = \{1, 7, 11\}$. This is consistent with the corresponding representative sample path shown in blue in panel (i), which display tristability. Let us note that, while the stochastic dynamics display multistability in (c), (f), and (i), the corresponding deterministic dynamics, also shown in the plots, remain monostable.
3. A Two-species Exotic System

Consider the two-species network $\mathcal{R}(s_1, s_2)$, given by

$$\begin{align*}
\mathcal{R}(s_1, s_2) : & \emptyset \xrightarrow{k_1} s_1, & \emptyset \xrightarrow{k_2} s_2, \\
& s_1 \xrightarrow{k_3} 2s_1, & s_2 \xrightarrow{k_4} \emptyset; \\
& 2s_1 \xrightarrow{k_5} 3s_1, & s_1 + s_2 \xrightarrow{k_6} s_1 + 2s_2, \\
& s_1 + s_2 \xrightarrow{k_7} s_2, & 2s_2 \xrightarrow{k_8} 3s_2, \\
& 2s_1 + s_2 \xrightarrow{k_9} s_1 + s_2, & 3s_2 \xrightarrow{k_10} 2s_2, \\
& s_1 + 2s_2 \xrightarrow{k_{11}} 2s_1 + 2s_2,
\end{align*}$$

((11))

where species $s_1$ and $s_2$ react according to the eleven reactions with rate coefficients $k_1, k_2, \ldots, k_{11} \geq 0$. We denote the copy-numbers of species $s_1$ and $s_2$, at time $t$ by $X_1(t)$ and $X_2(t)$, respectively. It was established in [21] that, for particular choices of the rate coefficients, the deterministic model of reaction network (11), given as equation (SI17) in SI Text, exhibits exotic dynamics: it undergoes a homoclinic bifurcation, and displays a bistability involving a limit cycle and an equilibrium point. On the other hand, it is demonstrated in [22] that the stochastic model of (11) is not necessarily sensitive to the deterministic bifurcation, and may effectively behave in a monostable manner. The latter point is demonstrated in Figure 2(c), where we show in red numerically approximated $x_1$-solutions of (SI17), one initiated in the region of attraction of the equilibrium point, while the other of the limit cycle. For a comparison, we also show in blue a representative sample path generated by applying the Gillespie algorithm on (11). It can be seen that the stochastic solution spends significantly more time near the deterministic equilibrium point. To gain a clearer picture, we display in Figures 2(a), and (b), the joint, and the $x_1$-marginal, stationary PMFs, respectively, underlying network (11), which have been obtained numerically for the same parameter values as in Figure 2(c). In (b), one can notice that the PMF is bimodal, but the left peak, corresponding to the limit cycle, is significantly smaller than the right peak, which corresponds to the stable equilibrium point.

We now apply Algorithm 1 on network (11) to achieve two goals. Firstly, we balance the sizes of the two peaks of the stationary PMF from Figure 2(b), thereby forcing the stochastic system to spend comparable amounts of time at the two deterministic attractors. Secondly, we reverse the situation shown in Figure 2(b), by making the left PMF peak significantly larger than the right one, thereby forcing the stochastic system to spend most of the time near the limit cycle. We could achieve the goals by introducing species $\bar{s}_1, \bar{s}_2$ into (11), and using suitable basis zero-drift networks. We take a simpler approach, by mapping (11) to $\mathcal{R}^1(\bar{s}_1, \bar{s}_2) \cup \mathcal{R}^2(\bar{s}_2) \cup (\mathcal{R}^3)_{\bar{s}_1=10}^1(\bar{s}_2, \bar{s}_2) \cup \mathcal{R}^3_{\bar{s}_1=0}(\bar{s}_2, \bar{s}_2)$, which is given by equation (SI18) in SI Text. For our purposes, only one of $\bar{s}_1, \bar{s}_2$ is sufficient, since the stochastic dynamics of $s_1$ and $s_2$ are coupled. We have chosen $\bar{s}_2$ for convenience, since $x_2$-state-space may be truncated at a lower value, $C_2 = 180$, than $x_1$-state-space (see also Figure 2(a)). The $x_3$-component of the deterministic limit cycle satisfies $x_2 \in (10, 30)$. Correspondingly, we introduce two zero-drift networks: $\mathcal{R}^1_{C_2=10}(s_2, \bar{s}_2)$, and $\mathcal{R}^3_{C_2=0}(s_2, \bar{s}_2)$, which redistribute the PMF from $x_2 \in [0, 10]$, and from $x_2 \in [30, C_2]$, respectively, to the limit cycle region, $x_2 \in (10, 30)$. We fix the scaled coefficient $K_{C_2=10}$ to a large value (so that the PMF is nearly zero for $x_2 \in [0, 10]$), and vary the coefficient $K_{C_2=0}$ which redistributes the PMF from the deterministic equilibrium point to the limit cycle. Network $\mathcal{R}^3_{C_2=0}(\bar{s}_2)$ is necessary for the preservation of the deterministic dynamics of (11) under the application of Algorithm 1.

In Figures 2(d), and (e), we show the joint, and $x_1$-marginal, stationary PMFs for an intermediate value of $K_{C_2=0}$, when the PMF is partially redistributed from $x_2 \in [30, C_2]$ to $x_2 \in (10, 30)$, so that the two peaks in (e) are of comparable sizes. In Figure 2(f), we show a representative sample path, obtained by applying the Gillespie algorithm on network (SI18) from SI Text, together with the deterministic trajectories obtained by solving (SI17). One can notice that the stochastic system now spends significantly more time near the limit cycle, when compared to (c). In Figures 2(f)–(g), we show analogous plots, but for a sufficiently large value of $K_{C_2=0}$, when the PMF is almost completely redistributed from $x_2 \in [30, C_2]$ to $x_2 \in (10, 30)$. Now, in contrast to Figures 2(a)–(c), the PMF becomes essentially unimodal, and concentrated around the limit cycle. Let us note that the red trajectories from Figures 2(f) and (i) were generated by numerically solving the deterministic model of network (11), given by (SI17). For our purposes, it is not necessary to solve the corresponding (stiff) deterministic model of network (SI18). The reason is that Algorithm 1 does not influence the deterministic equilibrium points of a given reaction network, regardless of the choice of the kinetic algorithm parameters. For example, while the deterministic limit cycle is not necessarily preserved for the algorithm parameters chosen in Figure 2(i), the enclosed deterministic unstable focus is necessarily preserved. Thus, the blue sample path corresponds to noise-induced oscillations either near a deterministic limit cycle, or near a deterministic unstable focus.

4. Summary

In this paper, we have presented the noise-control algorithm, which is given as Algorithm 1. The algorithm maps an input chemical reaction network to output networks, all under mass-action kinetics, by introducing appropriate additional species and reactions, such that the output networks satisfy the following two properties. Firstly, the output networks have the same deterministic model as the input network, in appropriate limits of some of the parameters (rate coefficients) introduced by the algorithm. Secondly, controllable state-dependent noise is introduced into the stochastic model of the output networks. Thus, Algorithm 1 may be used to control the intrinsic noise of a given reaction network under mass-action kinetics, while
preserving the deterministic dynamics. Let us note that the asymptotic conditions for the algorithm parameters are necessary for preservation of the time-dependent deterministic solutions. However, the time-independent deterministic solutions (the deterministic equilibrium points), which capture important features of the deterministic dynamics, are preserved under the algorithm even if the asymptotic conditions are not satisfied.

The algorithm has been applied to a test problem, taking the form of the one-species production-decay system given by (1). Using analytical and numerical methods, we have shown that the additional intrinsic noise, introduced by the algorithm, may be used to favorably modify the stationary probability mass function at arbitrary points in the state-space, as demonstrated in Figure 1. For example, in Figure 1(b), the noise is added to the whole interior of the state-space, while in (e) only at a single point, in both cases resulting in noise-induced bimodality. On the other hand, in Figure 1(h), by adding the noise to specific points in the state-space, the network is redesigned to display noise-induced trimodality. As shown in Figures 1(c), (f), (i), the blue stochastic trajectories display multistability, while the red deterministic ones remain monostable.

The algorithm has also been applied to a more challenging problem, taking the form of the two-species production-decay system given by (11), which, for the parameters taken in this paper, at the deterministic level displays a bistability involving an equilibrium point and a limit cycle [21, 22]. At the stochastic level, the system is significantly more likely to be found near the equilibrium point, as demonstrated in Figures 2(a)–(c). We have used the algorithm to redesign network (11), so that the stochastic system spends comparable amounts of time near the two attractors, as demonstrated in Figures 2(d)–(f). The network was also redesigned to display noise-induced oscillations, which is shown in Figures 2(g)–(i).

The controllable state-dependent noise is generated by Algorithm 1 using the zero-drift networks (19) and (20). Any nonnegative function, defined on a bounded discrete domain, may be represented by a linear combination of propensity functions induced by an appropriate union of the zero-drift networks. Thus, choosing suitable zero-drift networks, the algorithm may control the intrinsic noise at arbitrary points in the state-space of the stochastic dynamics of reaction networks. The cost of such a precision in noise-control is a larger number of reactants in the underlying zero-drift networks. However, while the high-molecular reactions introduced by the algorithm are more expensive to synthesize, they do not limit applicability of Algorithm 1 to synthetic biology. The reason for this is that such reactions may always be broken down into sets of up-to bi-molecular reactions, with asymptotically equivalent deterministic and stochastic dynamics [40, 41]. In particular, a zero-drift network, involving reactions of order \((n + \bar{n})\), may be broken down into \(2(n + \bar{n}) - 2\) reactions of up-to second-order, which may be readily mapped to DNA-based physical networks.

Algorithm 1 may constitute a qualitatively novel finding which will facilitate the progress of DNA computing [24]. In particular, a hybrid approach for constructing DNA-based reaction networks may be used: the deterministic model may be used to guide the construction of reaction networks, and then Algorithm 1 may be applied to favorably reprogram the intrinsic noise in the stochastic model, while preserving the mean-field behavior. The algorithm may be of critical importance when the synthetic networks involve species at low copy-numbers, since then the stochastic effects may play a significant role [31, 32, 14, 19, 22, 28, 29, 30], uncontrollably contaminating the performance of the synthetic networks. In such circumstances, Algorithm 1 may be used for controlling the stochastic effects, enriching the DNA-based synthetic systems with novel, noise-induced functionalities.

## 5. Methods

Let us consider the mass-action reaction network \(\mathcal{R}\) given by

\[
\mathcal{R}(s_1, \ldots, s_N) : \sum_{i=1}^{N} c_{ij} s_i \rightarrow \sum_{i=1}^{N} c'_{ij} s_i, \quad j \in \{1, \ldots, M\},
\]

(12)

where \(s_1, \ldots, s_N\) are the reacting species, \(k_j\) the reaction rate coefficients, and \(c_{ij}, c'_{ij}\) the stoichiometric coefficients. Let us denote by \(e_j, e'_j \in \mathbb{N}_0^N\) the vectors of the stoichiometric coefficients of reaction \(j\), and \(\Delta x_j = e'_j - e_j\).

The deterministic model of reaction network (12) is given by the following system of ordinary-differential equations (ODEs), also known as the reaction rate equations [23, 25]:

\[
\frac{dx}{dt} = \sum_{j=1}^{M} k_j x^{e_j} \Delta x_j, \quad i \in \{1, \ldots, N\}.
\]

(13)

Here, \(x = x(t) \in \mathbb{R}_+^N\) is the vector of species concentrations, i.e. \(x_i(t)\) is the concentration of species \(s_i\) at time \(t\), and \(x^{e_j} = \prod_{l=1}^{N} x_l^{e_{jl}}\), with the convention that \(0^0 \equiv 1\).

The stochastic model of reaction network (12) is given by the following system of difference-differential equations, also known as the chemical master equation (CME) [23, 25, 27]:

\[
\frac{\partial}{\partial t} p(x,t) = \mathcal{L} p(x,t) = \sum_j (E^{-\Delta x_j} - 1)(\alpha_j(x)p(x,t)).
\]

(14)

Here, \(p(x,t)\) is the probability mass function (PMF), i.e. the probability that the vector of copy-numbers \(X = X(t) \in \mathbb{N}_0^N\) of species \(s_1, \ldots, s_N\) at time \(t\) is given by \(x\). Linear operator \(\mathcal{L}\) is called the forward operator, and step operator \(E^{-\Delta x_j}\) is such that \(E^{-\Delta x_j} p(x,t) = p(x - \Delta x_j, t)\). Function \(\alpha_j(x)\) is the propensity function [23, 25] of the \(j\)-th reaction from (12), and is given by

\[
\alpha_j(x) = k_j x^{e'_j} = k_j \prod_{l=1}^{N} x_l^{e'_{jl}},
\]

(15)
where $x_l^{c_j}$ denotes a falling factorial of $x_l$, i.e. $x_l^{c_j} \equiv x_l(x_l - 1) \cdots (x_l - c_j + 1)$.

Acknowledgments

The authors would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme “Stochastic Dynamical Systems in Biology: Numerical Methods and Applications”, where work on this paper was undertaken. The authors would also like to thank John J. Tyson (Department of Biology, Virginia Polytechnic Institute and State University, USA) for a discussion on a possible realization of network (5) via a bifunctional histidine kinase/phosphatase from [39]. This work was supported by EPSRC grant no EP/K032208/1. This work was partially supported by a grant from the Simons Foundation. Konstantinos C. Zygalakis was supported by the Alan Turing Institute under the EPSRC grant EP/N510129/1. David F. Anderson would like to acknowledge the NSF grant NSF-DMS-1318832, and Army Research Office grant W911NF-14-1-0401. Radek Erban would also like to thank the Royal Society for a University Research Fellowship.

References

[1] Endy, D., 2005. Foundations for Engineering Biology. Nature, 483: 449–453.

[2] Andrianantoandro, E., Basu, S., Karig, D. K., Weiss, R., 2006. Synthetic biology: new engineering rules for an emerging discipline. Molecular Systems Biology, 2: 2006.0028.

[3] Abil, Z., Xiong, X., Zhao, H., 2015. Synthetic Biology for Therapeutic Applications. Molecular Pharmaceutics, 12(2): 322–331.

[4] Anderson, J. C., Clarke, E. J., Arkin, A. P., Voigt, C. A., 2006. Environmentally Controlled Infection of Cancer Cells by Engineered Bacteria. Journal of Molecular Biology, 355(4): 619–627.

[5] Benenson, Y., Gil, B., Ben-Dor, U., Adar, R., and Shapiro, E., 2004. An Autonomous Molecular Computer for Logical Control of Gene Expression. Nature, 429: 423–429.

[6] Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K., and Riggs, A. D., 1979. Expression in Escherichia Coli of Chemically Synthesized Genes for Human Insulin. Proc. Natl Acad. Sci. USA, 76(1): 106–110.

[7] Ro, D., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J., Chang, M. C. Y., Withers, S. T., Shiba, Y., Sarpong, R., and Keasling, J. D., 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature, 440: 940–943.

[8] Westfall, P. J., et al., 2012. Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. Proc. Natl Acad. Sci. USA, 109: E111–E118.

[9] Widmaier, D. M., Tullman-Ercek, D., Mirsky, E. A., Hill, R., Govindarajan, S., Minshull, J., and Voigt, C. A., 2009. Engineering the Salmonella Type III Secretion System to Export Spider Silk Monomers. Molecular Systems Biology, 5(309).

[10] Sedlak, M., and Ho, W. Y., 2004. Production of Ethanol from Cellulosic Biomass Hydrolysates Using Genetically Engineered Saccharomyces Yeast Capable of Cofermenting Glucose and Xylose. Applied Biochemistry and Biotechnology, 114(1): 403–416.

[11] Ball, P. 2005. Synthetic Biology for Nanotechnology. Nanotechnology, 16: R1–R8.

[12] Jungmann, R., Renner, S., and Simmel, F. C., 2008. From DNA Nanotechnology to Synthetic Biology. Applied Biochemistry and Biotechnology, 2(2): 99–109.

[13] Menezes, A. A., Cumbers, J., Hogan, J. A., Arkin, A. P., 2015. Towards synthetic biological approaches to resource utilization on space missions. Journal of The Royal Society Interface, 12(102): 20140715.

[14] Elowitz, M. B., Leibler, S., 2000. A Synthetical Oscillatory network of Transcriptional Regulators. Nature, 403: 335–338.

[15] Deamer, D., 2005. A giant step towards artificial life? Trends in Biotechnology, 23(7): 336–338.

[16] Glass, J. I, Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M., III, C. A. H., Smith, H. O., and Venter, J. C., 2006. Essential Genes of a Minimal Bacterium. Proceedings of the National Academy of Sciences, 103(2): 425–430.

[17] Gibson, D. G., Benders, G. A., Andrews-Pfannkoch, C., Denisova, E. A., Baden-Tillson, H., Zaveri, J., Stockwell, T. B., Brownley, A., Thomas, D. W., Algire, M. A., Merryman, C., Young, L., Noskov, V. N., Glass, J. I., Venter, J. C., III, C. A. H., Smith, H. O., 2008. Complete Chemical Synthesis, Assembly, and Cloning of a Mycoplasma genitalium Genome. Science’s STKE, 319(5867): 1215–1220.

[18] Gibson, D.G., et al., 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. Science, 329(5987): 52–56.

[19] Gardner, T. S., Cantor, C. R., and Collins, J. J., 2000. Construction of a Genetic Toggle Switch in Escherichia Coli. Nature, 403: 339–342.

[20] You, L., Cox III, R. S., Weiss, R., and Arnold, F. H., 2004. Programmed Population Control by Cell-Cell Communication and Regulated Killing. Nature, 428: 868–871.
Plesa, T., Vejchodský, T., and Erban, R., 2016. Chemical Reaction Systems with a Homoclinic Bifurcation: An Inverse Problem. Journal of Mathematical Chemistry, doi:10.1007/s10910-016-0656-1.

Plesa, T., Vejchodský, T., and Erban, R., 2016. Test Models for Statistical Inference: Two-Dimensional Reaction Systems Displaying Limit Cycle Bifurcations and Bistability, chapter contribution submitted to Stochastic Dynamical Systems, Multiscale Modeling, Asymptotics and Numerical Methods for Computational Cellular Biology, Ed. D. Holcman, available as https://arxiv.org/abs/1607.07738.

Érdi, P., Tóth, J. Mathematical Models of Chemical Reactions. Theory and Applications of Deterministic and Stochastic Models. Manchester University Press, Princeton University Press, 1989.

Soloveichik, D., Seeling, G., Winfree, E., 2010. DNA as a Universal Substrate for Chemical Kinetics. Proceedings of the National Academy of Sciences, 107(12): 5393–5398.

Anderson, D. F., Kurtz, T. G. Stochastic Analysis of Biochemical Systems. Springer, 2015.

Gillespie, D., 1977. Exact Stochastic Simulation of Coupled Chemical Reactions. Journal of Physical Chemistry, 81(25): 2340–2361.

Van Kampen, N. G. Stochastic Processes in Physics and Chemistry. Elsevier, 2007.

Erban, R., Chapman, S. J., Kevrekidis, I. and Vejchodský, T., 2009. Analysis of a stochastic chemical system close to a SNIPER bifurcation of its mean-field model. SIAM Journal on Applied Mathematics, 70(3): 984–1016.

Duncan, A., Liao, S., Vejchodský, T., Erban, R., Grima, R., 2015. Noise-induced multistability in chemical systems: discrete vs continuum modelling. Physical Review E, 91, 042111.

Yates, C., Erban, R., Escudero, C., Couzin, I., Buhl, J., Kevrekidis, I., Maini, P., and Sumpter, D., 2009. Inherent Noise can Facilitate Coherence in Collective Swarm Motion. Proceedings of the National Academy of Sciences, 106(14): 5464–5469.

Vilar, J. M. G., Kueh, H. Y., Barkai, N. and Leibler, S., 2002. Mechanisms of Noise-resistance in Genetic Oscillators. Proceedings of the National Academy of Sciences of the United States of America, 99(9): 5988–5992.

Dublanche, Y., Michalodimitrakis, K., Kummerer, N., Foglerini, M. and Serrano, L., 2006. Noise in Transcription Negative Feedback Loops: Simulation and Experimental Analysis. Molecular Systems Biology, 2(41): E1–E12.

Kurtz, T. G., 1972. The Relationship between Stochastic and Deterministic Models for Chemical Reactions. Journal of Chemical Physics, 57: 2976–2978.

Srinivas, N. Programming Chemical Kinetics: Engineering Dynamic Reaction Networks with DNA Strand Displacement. PhD Thesis, California Institute of Technology, Pasadena, California, 2015.

Soltanoughi, D., Cook, M., Winfree, E., Bruck, J., 2008. Computation with Finite Stochastic Chemical Reaction Networks. Natural Computing, 7(4): 615–633.

Ohkubo, J., Shnerb, N., and Kessler, D. A., 2008. Transition Phenomena Induced by Internal Noise and Quasi-Absorbing State. Journal of the Physical Society of Japan, 77, 044002.

Biancalani, T., Dyson, L., and McKane, A. J., 2014. Noise-Induced Bistable States and Their Mean Switching Time in Foraging Colonies. Physical Review Letters, 112, 038101.

Saito, N., and Kaneko, K., 2015. Theoretical Analysis of Discreteness-Induced Transition in Autocatalytic Reaction Dynamics. Physical Review Letters, 91, 022707.

Subramanian, K., Paul, M. R., Tyson, J. J., 2013. Potential Role of a Bistable Histidine Kinase Switch in the Asymmetric Division Cycle of Caulobacter crescentus. PLOS Computational Biology, 9, e1003221.

Wilhelm, T., 2000. Chemical systems consisting only of elementary steps - a paradigm for nonlinear behavior. Journal of Mathematical Chemistry, 27: 71–88.

Plesa, T., 2017. Stochastic Approximation of High-Molecular by Bi-molecular Reactions. In preparation.

Klonowski, W., 1983. Simplifying principles for chemical and enzyme reaction kinetics. Biophysical Chemistry, 18(3): 73–87.

Supplementary Information (SI) Text

The Deterministic Dynamics of Network $\mathbb{R}^1 \cup \mathbb{R}^2$ in the Limit $\mu \to 0$

Let us analyse system (4) in the asymptotic limit $\mu \to 0$. It follows from the Tikhonov theorem [42] that the ODE for $y$, given by second equation in (4), reduces to the algebraic equation $y = (c - x)^{-1}$ as $\mu \to 0$. Substituting the algebraic equation into (4) results in

$$\frac{dx}{dt} = k_1 - k_2 x,$$

$$x(0) = x_0, \text{ as } \mu \to 0.$$  \hspace{1cm} (SI1)

Initial value problems (2) and (SI1) have the same form, and let us denote their solutions by $\tilde{x}(t; x_0)$ and $x(t; x_0)$, respectively. Then, choosing $c \geq \max_{x \in [0, \infty)} \tilde{x}(t; x_0) < \infty$ and $x_0 = \tilde{x}_0$, ensures that concentration of auxiliary species $\tilde{x}$ is nonnegative, $\tilde{x}(t) = c - x(t) \geq 0$, and that the solutions of (2) and (4) are asymptotically equivalent in the limit $\mu \to 0$. 

The Stochastic Dynamics of Network $\mathcal{R}^1 \cup \mathcal{R}_1^2 \cup \mathcal{R}_1^3$ in the Limit $\mu \to 0$

The chemical master equation (CME) [27] induced by network $\mathcal{R}^1 \cup \mathcal{R}_1^2 \cup \mathcal{R}_1^3$ is given by

$$ \frac{\partial}{\partial t} p(x,y,t) = \left( \mathcal{L}^1 + 1 + \frac{1}{\mu} \mathcal{L}_1^2 + K_{1,1} \mathcal{L}_1^3 \right) p(x,y,t), \tag{SI2} $$

where $x(t), y(t) \in \mathbb{N}_0$ are copy-numbers of species $s, l_1$ from (3), respectively, with

$$ \mathcal{L}^1 = k_1(E^{-1} - 1)((C - x)y) + k_2(E^+ - 1)x, $$
$$ \mathcal{L}_1^2 = (E^{-1} - 1) + (C - x)(E^+ - 1)y, $$
$$ \mathcal{L}_1^3 = (E^{-1} + E^+ - 2)\beta_{1,1}(x), \tag{SI3} $$

and $K_{1,1}, \beta_{1,1}(x)$ are given in (8). Operators $\mathcal{L}^1, \mathcal{L}_1^2, \mathcal{L}_1^3$ are induced by subnetworks $\mathcal{R}^1, \mathcal{R}_1^2, \mathcal{R}_1^3$, respectively.

Let us analyse system (SI2) in the limit $\mu \to 0$, and consider the following power-series expansion:

$$ p(x,y,t) = p_0(x,y,t) + \mu p_1(x,y,t) + \ldots + \mu^i p_i(x,y,t) + \ldots, \tag{SI4} $$

with $i \geq 2$. Substituting (SI4) into (SI2), and equating terms of equal powers in $\mu$, the following system of equations is obtained:

$$ \mathcal{O} \left( \frac{1}{\mu} \right) : -\mathcal{L}^2 p_0(x,y,t) = 0, $$
$$ \mathcal{O}(1) : -\mathcal{L}^2 p_1(x,y,t) = (\mathcal{L}^1 + K_{1,1} \mathcal{L}_1^3 - \frac{\partial}{\partial t})p_0(x,y,t). \tag{SI5} $$

Order $1/\mu$ equation. A suitable form of the zero-order approximation of the PMF follows from the Bayes theorem: $p_0(x,y,t) = p_0(y|x)p_0(x,t)$, where $p_0(y|x)$ is the stationary PMF of $y$ conditional on $x$, while $p_0(x,t)$ is the marginal PMF of $x$. Substituting $p_0(x,y,t) = p_0(y|x)p_0(x,t)$ into the first equation in (SI5), with $i, x$ fixed, leads to $-\mathcal{L}^2 p_0(y|x) = 0$. It follows that $p_0(y|x)$ is a Poisson distribution with parameter $(C-x)^{-1}$, so that the zero-order PMF is given by

$$ p_0(x,y,t) = \left( \frac{1}{y!} \left( \frac{1}{(C-x)} \right)^y \exp \left( -\frac{1}{(C-x)} \right) \right) p_0(x,t). \tag{SI6} $$

Order 1 equation. Substituting (SI6) into the second equation in (SI5), summing over all the possible states $y \in \mathbb{N}_0$, using (SI3), and equalities $\sum_y y p_0(y|x) = (C-x)^{-1}$ and $\sum_x p_0(y|x) = 1$, one obtains the effective CME, given by

$$ \frac{\partial}{\partial t} p_0(x,t) = (\mathcal{L} + K_{1,1} \mathcal{L}_1^3) p_0(x,t), \tag{SI7} $$

where $\mathcal{L}$ is the forward operator corresponding to network (1), and has the following form

$$ \mathcal{L} = k_1(E^{-1} - 1) + k_2(E^+ - 1)x. \tag{SI8} $$

Limit $K_{1,1} \to 0$

Setting the left-hand side (LHS) to zero, and taking $K_{1,1} = 0$ in (SI7), and assuming $C$ is fixed to a sufficiently large value, it follows that the stationary PMF is a Poisson distribution with parameter $k_1/k_2$ [27]:

$$ p_0(x) = \begin{cases} \left( \frac{a}{C} \right)^x \exp \left( -\frac{a}{C} \right), & \text{if } x \in [0,C], \\ 0, & \text{otherwise}. \end{cases} \tag{SI9} $$

Limit $K_{1,1} \to \infty$

Let us substitute the power-series expansion

$$ p_0(x) = f_0(x) + \frac{1}{K_{1,1}} f_1(x) + \ldots + \left( \frac{1}{K_{1,1}} \right)^i f_i(x) + \ldots, \tag{SI10} $$

with $i \geq 2$, into (SI7) with the LHS set to zero, and consider the limit $K_{1,1} \to \infty$. Then, equating terms of equal powers in $1/K_{1,1}$, one obtains:

$$ \mathcal{O}(1) : -\mathcal{L}^2 f_0(x) = 0, $$
$$ \mathcal{O} \left( \frac{1}{K_{1,1}} \right) : -\mathcal{L}_1^3 f_1(x) = \mathcal{L}_1^3 f_0(x). \tag{SI11} $$

Order 1 equation. The solution to the first equation in (SI11) is given by

$$ f_0(x) = \begin{cases} 1, & \text{if } x = 0, \\ \frac{a}{C}, & \text{if } x = C, \\ 0, & \text{otherwise}. \end{cases} \tag{SI12} $$

where $a \in \mathbb{R}_+$ is an arbitrary constant.

Order $1/K_{1,1}$ equation. Multiplying the second equation in (SI11) by $x$, and summing over $x \in \mathbb{N}_0$, with the convention that $f_0(x) = 0$ and $f_1(x) = 0$ for $x \notin [0,C]$, one obtains the solvability condition $0 = \sum_{x=0}^{\infty} x \mathcal{L}_1^3 f_0(x)$, which implies $a = k_1/k_2$. Substituting $a$ into (SI12) leads to the zero-order approximation of the stationary PMF:

$$ f_0(x) = \begin{cases} 1 - \frac{k_1}{k_2}, & \text{if } x = 0, \\ \frac{k_1}{k_2}, & \text{if } x = C, \\ 0, & \text{otherwise}. \end{cases} \tag{SI13} $$

Zero-Drift Networks $\mathcal{R}_3^{n,\bar{n}}$

The propensity function of reactions underlying $\mathcal{R}_3^{n,\bar{n}}(x,\bar{n})$, $n,\bar{n} \in \mathbb{N}_0$, and $(n+\bar{n}) \leq C$, is given by $K_{n,\bar{n}} \beta_{n,\bar{n}} : [0,C] \to \mathbb{R}_+$, with

$$ K_{n,\bar{n}} = M_{n,\bar{n}} \beta_{n,\bar{n}}, \tag{SI14} $$

and

$$ \beta_{n,\bar{n}}(x) = (M_{n,\bar{n}})^{-1} \prod_{i=0}^{n-1} x^i \prod_{i=0}^{\bar{n}-1} ((C - i) - x), \tag{SI15} $$
where the scaling factor \( M_{n, \bar{n}} \) is introduced to approximately normalize \( \beta_{n, \bar{n}}(x) \), and is given by

\[
M_{n, \bar{n}} = \prod_{i=0}^{n-1} \left( \frac{n}{n+\bar{n}} C - i \right) \prod_{i=0}^{\bar{n}-1} \left( \frac{\bar{n}}{n+\bar{n}} C - i \right). \tag{SI16}
\]

Here, we take the convention \( \prod_{i=0}^{N} f(i) = 1 \) if \( N < 0 \), where \( f(i) \) is an arbitrary function of \( i \). Function \( \beta_{n, \bar{n}}(x) \) is nonzero on the interval \([n, C-\bar{n}]\), with the single maximum approximately at \( Cn/(n+\bar{n}) \).

**Interior zero-drift networks.** Zero-drift network \( \mathcal{R}_{n, \bar{n}}^3(s, \bar{s}) \), with \( n, \bar{n} \neq 0 \), satisfies (19), and the propensity function of its reactions, which is proportional to (SI15), is nonzero only in the interior of the state-space. Since the propensity function of \( \mathcal{R}_{n, \bar{n}}^3(s, \bar{s}) \), with \( n, \bar{n} \neq 0 \), attains its maximum in the interior of the domain, we call the network an interior zero-drift network.

**Boundary zero-drift networks.** Network \( \mathcal{R}_{n, \bar{n}}^3(s, \bar{s}) \), satisfying (20), is a zero-drift network in the limit \( \mu_{n, \bar{n}} \to 0 \). Furthermore, in the same limit, the first two reactions from (20) have the same propensity function, which is proportional to (SI15) with \( n = 0 \), and which is nonzero at the left boundary point, \( x = 0 \). Similarly, network \( \mathcal{R}_{n, 0}^3(s, \bar{s}; \tilde{B}, k_{n, 0}, \mu_{n, 0}) \) is a zero-drift network as \( \mu_{n, 0} \to 0 \), and its first two reactions have the same propensity function, which is nonzero at the right boundary point, \( x = C \). Since networks with \( n = 0 \) (respectively, \( \bar{n} = 0 \)) generate propensity functions with the maximum values at the left (respectively, right) boundary point, we call such networks left (respectively, right) boundary zero-drift networks.

**Basis zero-drift networks.** Stoichiometric coefficients \( n, \bar{n} \) control the support of the intrinsic noise, which network \( \mathcal{R}_{n, \bar{n}}^3 \) introduces into the stochastic dynamics, via the control of support of the compact function (SI15). The larger the sum \((n + \bar{n})\) is, with \((n + \bar{n}) \leq C\), the smaller the support of (SI15), and, hence, one obtains a more precise noise-control. In the special case when \( n + \bar{n} = C \), the propensity function (SI15) is nonzero only at a single point in the state-space, \( x = n \). We call networks \( \mathcal{R}_{n, \bar{n}}^3(s, \bar{s}) \), with \( n + \bar{n} = C \), basis zero-drift networks, and the corresponding propensity functions basis propensity functions. Any nonnegative function, defined on a bounded discrete domain, may be represented by a suitable linear combination of the basis propensity functions.

**The Deterministic Model for Network \( \mathcal{R} \)**

The deterministic model of network (11) is given by

\[
\begin{align*}
\frac{dx_1}{dt} &= k_1 + k_2 x_1 + k_3 x_1^2 - k_4 x_1 x_2 - k_5 x_1^2 x_2 + k_6 x_1 x_2^2, \\
\frac{dx_2}{dt} &= k_7 - k_8 x_2 + k_9 x_1 x_2 + k_{10} x_2^2 - k_{11} x_2^3, \tag{SI17}
\end{align*}
\]

where \( x_1 = x_1(t), x_2 = x_2(t) \) are the concentrations of species \( s_1, s_2 \), respectively, at time \( t \).
Figure 1. Panels (a), (d) and (g) display propensity functions $\beta_{1,1}(x)$, $\beta_{5,10}(x)$ and $\beta(x) \equiv \beta_{0,15}(x) + \beta_{2,9}(x) + \beta_{8,5}(x) + \beta_{12,0}(x)$, respectively. Panels (b), (e) and (h) display the stationary PMF of networks $\hat{R}^1 \cup \hat{R}^2 \cup \hat{R}^3_{1,1}$, $\hat{R}^1 \cup \hat{R}^2 \cup \hat{R}^3_{5,10}$ and $\hat{R}^1 \cup \hat{R}^2 \cup (\hat{R}^3_{0,15} \cup \hat{R}^3_{2,9} \cup \hat{R}^3_8 \cup \hat{R}^3_{12,0})$, respectively, where $\hat{R}^1 \cup \hat{R}^2$ is given by (3), while the rest of the (zero-drift) networks are as given in second step of Algorithm 1. In (h), $K = K_{0,15} = K_{2,9} = K_{8,5} = K_{12,0}$. Panels (c), (f), and (i) display in blue the sample paths, corresponding to the PMFs shown in (b), (e) and (h), respectively, and were obtained by applying the Gillespie algorithm on the underlying networks. Also shown in red are the deterministic trajectories, obtained by numerically solving the corresponding deterministic models. The dimensionless parameters are fixed to: $k_1 = 2.5$, $k_2 = 0.5$, $\mu = 10^{-3}$, $C = 15$, and the state-space for species $I$ is bounded in (b), (e) and (h) by 50. In (b) and (e), the two-species stationary chemical master equation (CME) was numerically solved, while in (h) the boundary zero-drift networks are taken in the asymptotic limits $\mu_{0,15}, \mu_{12,0} \to 0$. The blue and red trajectories from panel (i) were generated with $(\mu_{0,15})^{-1} M_{0,15} = (\mu_{12,0})^{-1} M_{12,0} = 10^7$. The trajectories from (c), (f) and (i) were all initiated at the deterministic equilibrium, $X(0) = 5$. 
Figure 2. Panel (a) displays the joint stationary PMF of network (11), while (d) and (g) display the stationary PMFs of network (SI18) from SI Text for \((K_{0.C2} = 10^{18}, K_{30,0} = 10^{10})\) and \((K_{0.C2} = 10^{18}, K_{30,0} = 10^{10})\), respectively, with the rest of the parameters being the same. Panels (b), (e) and (h) display the \(x_1\)-marginal PMFs corresponding to (a), (b) and (c), respectively. Panels (c), (f) and (i) display in blue the sample paths, corresponding to the PMFs shown in (b), (e) and (h), respectively, and were obtained by applying the Gillespie algorithm on the underlying networks. Also shown in red are two deterministic trajectories, one initiated near the equilibrium point, while the other near the limit cycle, obtained by numerically solving equation (SI17) from SI Text. The dimensionless parameters are fixed to: \(k_1 = 4, k_2 = 1.408, k_3 = 0.0518, k_4 = 0.164, k_5 = 3.1 \times 10^{-3}, k_6 = 4.8 \times 10^{-3}, k_7 = 4, k_8 = 8, k_9 = 0.16, k_{10} = 0.104, k_{11} = 2.1 \times 10^{-3}\). In (a)–(b), (d)–(e) and (g)–(h), the stationary chemical master equation (CME) is numerically solved, with the state-space truncated to \((x_1, x_2) \in [0, C_1] \times [0, C_2]\), where \(C_1 = 300, C_2 = 180\), and \(\mu, \mu_{0.C2}, \mu_{30,0} \to 0\). The blue sample paths from panels (f) and (i) were generated with \((\mu^{-1}, (\mu_{0.C2}^{-1})^{-1} M_{0.C2}^{-1}, (\mu_{30,0}^{-1})^{-1} M_{30,0}^{-1}) = (10^3, 10^{20}, 2 \times 10^{10})\) and \((\mu^{-1}, (\mu_{0.C2}^{-1})^{-1} M_{0.C2}^{-1}, (\mu_{30,0}^{-1})^{-1} M_{30,0}^{-1}) = (10^3, 10^{20}, 10^{20})\), respectively. The blue trajectories from (c), (f) and (i) were all initiated near the deterministic limit cycle.
Input: Let the input reaction network be given by

\[ \tilde{\mathcal{R}}(s_1, \ldots, s_N) : \sum_{i=1}^{N} c_{ij} s_i \xrightarrow{k_j} \sum_{i=1}^{N} c'_{ij} s_i, \quad j \in \{1, \ldots, M\}, \]  

(16)

where \( s_1, \ldots, s_N \), are the species, \( k_j \) the reaction rate coefficients, and \( c_{ij}, c'_{ij} \) the stoichiometric coefficients.

(1) Step: Reaction network \( \tilde{\mathcal{R}} \), given by (16), is mapped to a pairwise conservative network \( \tilde{\mathcal{R}}^1 \) given by

\[ \tilde{\mathcal{R}}^1(s_1, \ldots, s_N, \tilde{s}_1, \ldots, \tilde{s}_N) : \sum_{i=1}^{N} \left( c_{ij} s_i + (\Delta x_{ij} \tilde{s}_i) \right) \times 1_{\mathbb{N}}(\Delta x_{ij}) \xrightarrow{k_j} \sum_{i=1}^{N} \left( c'_{ij} s_i - (\Delta x_{ij} \tilde{s}_i) \right) \times 1_{\mathbb{N}}(-\Delta x_{ij}) + I_i^{\Delta x_{ij}} \times 1_{\mathbb{N}}(\Delta x_{ij}), \quad j \in \{1, \ldots, M\}. \]  

(17)

Here, \( \tilde{s}_i, I_i^{\Delta x_{ij}} \) are additional species, \( \Delta x_{ij} = (c'_{ij} - c_{ij}) \), and \( 1_{\mathbb{N}}(\cdot) \) is the indicator function of the natural numbers.

(2) Step: For each species \( I_i^{\Delta x_{ij}} \), a drift-corrector network is constructed, \( \mathcal{R}^2_{\Delta x_{ij}}(\tilde{s}_i) = \mathcal{R}^2_{\Delta x_{ij}}(\tilde{s}_i; I_i^{\Delta x_{ij}}, \mu) \), given by

\[ \mathcal{R}^2_{\Delta x_{ij}}(\tilde{s}_i) : \quad \otimes \xrightarrow{1/\mu} I_i^{\Delta x_{ij}}, \]  

\[ \Delta x_{ij} \tilde{s}_i + I_i^{\Delta x_{ij}} \xrightarrow{1/\mu} \Delta x_{ij} \tilde{s}_i, \]  

(18)

where \( 0 \leq \mu \ll 1 \).

(3) Step: For each species \( \tilde{s}_i \), a union of zero-drift networks may be constructed. Let \( n, \bar{n} \in \mathbb{N}_0 \), and \((n + \bar{n}) \leq C_i \). Network \( \mathcal{R}^3_{n, \bar{n}}(s_i, \tilde{s}_i) = \mathcal{R}^3_{n, \bar{n}}(s_i, \tilde{s}_i; k_{n, \bar{n}}^i) \), with \( n, \bar{n} \neq 0 \), is given by

\[ \mathcal{R}^3_{n, \bar{n}}(s_i, \tilde{s}_i) : \quad ns_i + \bar{n}\tilde{s}_i \xrightarrow{k_{n, \bar{n}}^i} (n+1)s_i + (\bar{n}-1)\tilde{s}_i, \]  

\[ ns_i + \bar{n}\tilde{s}_i \xrightarrow{1/k_{n, \bar{n}}^i} (n-1)s_i + (\bar{n}+1)\tilde{s}_i. \]  

(19)

Network \( \mathcal{R}^3_{0, \bar{n}}(s_i, \tilde{s}_i) = \mathcal{R}^3_{0, \bar{n}}(s_i, \tilde{s}_i; B_i, k_{0, \bar{n}}^{0, \bar{n}}) \), with \( \bar{n} \neq 0 \), is given by

\[ \mathcal{R}^3_{0, \bar{n}}(s_i, \tilde{s}_i) : \quad \bar{n}\tilde{s}_i \xrightarrow{k_{0, \bar{n}}^{0, \bar{n}}} s_i + (\bar{n} - 1)\tilde{s}_i, \]  

\[ C_i s_i + B_i \xrightarrow{k_{0, \bar{n}}^{0, \bar{n}}} (C_i - 1)s_i + \tilde{s}_i + B_i, \]  

\[ \bar{n}\tilde{s}_i \xrightarrow{1/k_{0, \bar{n}}^{0, \bar{n}}} \bar{n}\tilde{s}_i + B_i, \]  

\[ C_i s_i + B_i \xrightarrow{1/k_{0, \bar{n}}^{0, \bar{n}}} C_i s_i, \]  

(20)

where \( 0 \leq \mu_{0, \bar{n}} \ll 1 \), and \( B_i \) is an additional species. Network \( \mathcal{R}^3_{n, 0} = \mathcal{R}^3_{0, n}(\tilde{s}_i, s_i; \bar{B}_i, k_{n, 0}^{i, k} \mu_{n, 0}), \) is given by

Output: An output reaction network \( \mathcal{R} \) is given by

\[ \mathcal{R} = \mathcal{R}^1 \cup \mathcal{R}^2 \cup \mathcal{R}^3, \]  

(21)

where \( \mathcal{R}^2 = \cup_{\Delta x_{ij}} \mathcal{R}^2_{\Delta x_{ij}}(\tilde{s}_i) \), and \( \mathcal{R}^3 = \cup_{(n, \bar{n})} \mathcal{R}^3_{n, \bar{n}}(s_i, \tilde{s}_i) \).

Algorithm 1. The noise-control algorithm.