

# Supporting Information for Optimization-based Synthesis of Stochastic Biocircuits with Statistical Specifications \*

Yuta Sakurai    Yutaka Hori<sup>†</sup>

## S. 1 Model of stochastic chemical reactions

We consider a chemical reaction network that consists of  $n \in \mathbb{N}_+$  species of molecules and  $r$  types of reactions. We denote the copy number of each molecular species by  $\mathbf{x} = [x_1, x_2, \dots, x_n]^T \in \mathbb{N}_0^n$  and stoichiometry of the  $i$ -th reaction by  $\mathbf{s}_i = [s_{i1}, s_{i2}, \dots, s_{in}]^T \in \mathbb{Z}^n$  ( $i = 1, 2, \dots, r$ ). The stoichiometry vector represents the increment of the copy number  $\mathbf{x}$  by each chemical reaction.

We define  $P_{\mathbf{x}}(t)$  as the probability that there are  $\mathbf{x}$  molecules at time  $t$ . Then, the dynamics of  $P_{\mathbf{x}}(t)$  can be modeled by the following Chemical Master Equation (CME) [4],

$$\frac{dP_{\mathbf{x}}(t)}{dt} = \sum_{i=1}^r \{w_i(\mathbf{x} - \mathbf{s}_i)P_{\mathbf{x}-\mathbf{s}_i}(t) - w_i(\mathbf{x})P_{\mathbf{x}}(t)\}, \quad (\text{A.1})$$

where  $w_i(\mathbf{x})$  is the propensity function (reaction rate) associated with the  $i$ -th chemical reaction and is defined by

$$w_i(\mathbf{x}) = \lim_{\Delta t \rightarrow 0} \frac{P_{\mathbf{x}+\mathbf{s}_i}(t + \Delta t) - P_{\mathbf{x}}(t)}{\Delta t} \quad (i = 1, 2, \dots, r). \quad (\text{A.2})$$

We assume that all of the  $r$  reactions are elementary. Note that non-elementary reactions can always be decomposed into elementary reactions by considering intermediate molecular species [2]. For elementary reactions, the propensity function  $w_i(\mathbf{x})$  becomes a polynomial of  $x_j$ 's [3].

---

\*Published in Journal of The Royal Society Interface

<sup>†</sup>The authors contributed equally to this work. Correspondence should be addressed to Yutaka Hori. Department of Applied Physics and Physico-Informatics, Keio University, 223-8522 Kanagawa, Japan. Email: yhori@appi.keio.ac.jp

## S. 2 Transcription and translation model in Fig. 2

### Monte-Carlo simulations of mRNA and protein copy numbers

We consider a model of transcription and translation shown in Fig. 2A. In this model, the copy numbers of mRNA and protein are the variables of the biocircuit ( $n = 2$ ). There are four reactions, transcription, translation and degradation of mRNA and protein ( $r = 4$ ). We denote the copy numbers of mRNA and protein by  $x_1$  and  $x_2$ , respectively, and define  $\boldsymbol{x} = [x_1, x_2]^T$ . Specific forms of the propensity functions  $w_i(\boldsymbol{x})$  and the stoichiometry vectors  $\boldsymbol{s}_i$  ( $i = 1, 2, 3, 4$ ) are defined as summarized in Table B.1, where  $D_{T_1}$  is a constant representing the copy number of DNA.

Table B.1: Propensity functions and stoichiometry vectors for the transcription and translation model

Reaction $i$	$w_i(\boldsymbol{x})$	$\boldsymbol{s}_i$
1	$k_1 D_T$	$[1, 0]^T$
2	$k_2 x_1$	$[-1, 0]^T$
3	$k_3 x_1$	$[0, 1]^T$
4	$k_4 x_2$	$[0, -1]^T$

### S. 3 Derivation of moment equations

In this section, we derive the model of moment dynamics based on the CME (A.1). We define the raw moments of random variables  $x_j$ 's by

$$m_{\alpha} := \mathbb{E} \left[ \prod_{j=1}^n x_j^{\alpha_j} \right] = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \cdots \sum_{x_n=0}^{\infty} \prod_{j=1}^n x_j^{\alpha_j} P_{\mathbf{x}}(t), \quad (\text{A.3})$$

where  $\alpha := [\alpha_1, \alpha_2, \dots, \alpha_n] \in \mathbb{N}_0^n$ . We refer to the sum of the entries of  $\alpha$  as the order of the moment and denote by  $\deg(m_{\alpha}) := \sum_j \alpha_j$ .

To derive a moment equation, we multiply the model (A.1) by  $\prod_j x_j^{\alpha_j}$  and take the sum of the equations for all non-negative values of  $x_j$ 's to obtain

$$\frac{d}{dt} \left\{ \sum_{\mathbf{x}} \left( \prod_{j=1}^n x_j^{\alpha_j} \right) P_{\mathbf{x}}(t) \right\} = \sum_{\mathbf{x}} \left( \prod_{j=1}^n x_j^{\alpha_j} \right) \sum_{i=1}^r \{ w_i(\mathbf{x} - \mathbf{s}_i) P_{\mathbf{x}-\mathbf{s}_i}(t) - w_i(\mathbf{x}) P_{\mathbf{x}}(t) \}, \quad (\text{A.4})$$

where  $\sum_{\mathbf{x}}$  represents the sum for all nonnegative  $x_j$ 's, that is,

$$\sum_{\mathbf{x}} := \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \cdots \sum_{x_n=0}^{\infty}. \quad (\text{A.5})$$

Using the fact that  $P_{\mathbf{x}}(t) = 0$  if  $x_j < 0$  for some  $j = 1, 2, \dots, n$ , (A.4) can be rewritten as

$$\frac{d}{dt} \left\{ \sum_{\mathbf{x}} \left( \prod_{j=1}^n x_j^{\alpha_j} \right) P_{\mathbf{x}}(t) \right\} = \sum_{\mathbf{x}} \sum_{i=1}^r \left\{ \left( \prod_{j=1}^n (x_j + s_{ji}^{\alpha_j}) \right) w_i(\mathbf{x}) P_{\mathbf{x}}(t) - \left( \prod_{j=1}^n x_j^{\alpha_j} \right) w_i(\mathbf{x}) P_{\mathbf{x}}(t) \right\}. \quad (\text{A.6})$$

Since  $w_i(\mathbf{x})$  in (A.6) is a polynomial of  $x_j$  ( $j = 1, 2, \dots, n$ ), (A.4) leads to the following differential equation of raw moment dynamics

$$\frac{d}{dt} m_{\alpha} = \sum_{\beta_1} \sum_{\beta_2} \cdots \sum_{\beta_n} A_{\alpha, \beta} m_{\beta} \quad (\alpha \in \mathbb{N}_0^n), \quad (\text{A.7})$$

where  $A_{\alpha, \beta}$  are the coefficients of the monomials of

$$\sum_{i=1}^r \left\{ (x_1 + s_{i1})^{\alpha_1} (x_2 + s_{i2})^{\alpha_2} \cdots (x_n + s_{in})^{\alpha_n} - x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \right\} w_i(\mathbf{x}). \quad (\text{A.8})$$

Note that the range of the summation in (A.7) depends on the degree of the polynomial (A.8). In many cases, the dynamics of a moment of order  $\ell$  are characterized by the moments of order  $\ell + 1$  or higher, *i.e.*, there are moments  $m_{\beta}$  satisfying  $\deg(m_{\beta}) > \deg(m_{\alpha})$  in the right-hand side of (A.7).

In general, it follows from (A.8) that the right-hand side of (A.7) satisfies

$$0 \leq \deg(m_\beta) \leq \deg(m_\alpha), \quad (\text{A.9})$$

if and only if  $\deg(w_i(\mathbf{x})) = 0$  or  $1$ . In other words, the dynamics of the  $\nu$ -th order moment are expressed by the  $\nu + 1$ -th or lower order moments if and only if the propensity functions are affine in  $\mathbf{x}$ . Otherwise,  $\deg(m_\beta) > \deg(m_\alpha)$  holds, and we need infinitely many raw moments to determine  $m_\alpha$ .

To derive a general form of the truncated moment equation, we consider a finite subset of the equations (A.7) by limiting the order of the moments in the left-hand side up to  $\mu$ , which we designate as the truncation order of the moment equation. Specifically, let  $\mathbf{m}_{\alpha(\mu)}$  be a column vector that is composed of all the elements of the set  $\{m_\alpha \mid \deg(m_\alpha) = \mu\}$  and define the following vectors of moments

$$\mathbf{m} := [\mathbf{m}_{\alpha(0)}, \mathbf{m}_{\alpha(1)}, \dots, \mathbf{m}_{\alpha(\mu)}]^\top, \quad (\text{A.10})$$

$$\mathbf{u} := \mathbf{m}_{\alpha(\mu+1)}. \quad (\text{A.11})$$

Then, (A.7) can be rewritten as

$$\frac{d\mathbf{m}}{dt} = A\mathbf{m} + B\mathbf{u} \quad (\text{A.12})$$

Substituting  $d\mathbf{m}/dt = \mathbf{0}$ , we obtain the steady state moment equation

$$\mathbf{0} = A\mathbf{m} + B\mathbf{u}. \quad (\text{A.13})$$

It follows from (A.9) that  $B = 0$  if and only if  $\deg(w_i(\mathbf{x})) = 0$  or  $1$ . Thus, we can uniquely determine the solution of the linear equation  $A\mathbf{m} = 0$  (if the equation is not degenerated). On the other hand,  $B \neq 0$  holds when the transition rates  $w_i(\mathbf{x})$  are quadratic or higher order, resulting in the equation (A.13) that depends on higher order moments  $\mathbf{u}$ . Note that the linear equation (A.13) is underdetermined, and the solution cannot be determined uniquely unless there are additional constraints.

## S. 4 Moment conditions and derivation of Proposition

In this section, we show the derivation of the Proposition, which is a semi-algebraic condition that the moments of probability distributions must satisfy, and discuss some properties of the conditions.

Let  $\mathbf{x}^p$  be the vector that consists of all of the monomial bases satisfying  $\deg(\prod_j x_j^{p_j}) = p$ , and  $(H)_{p,q}$  be a matrix of the form

$$(H)_{p,q} := \mathbb{E}[\mathbf{x}^p (\mathbf{x}^q)^\top], \quad (\text{A.14})$$

where  $\mathbb{E}[\mathbf{x}^p (\mathbf{x}^q)^\top]$  represents the entry-wise expected value of the matrix  $\mathbf{x}^p (\mathbf{x}^q)^\top$ . In a similar manner, we define  $(H_k)_{p,q}$  by

$$(H_k)_{p,q} := \mathbb{E}[x_k \mathbf{x}^p (\mathbf{x}^q)^\top] \quad (k = 1, 2, \dots, n). \quad (\text{A.15})$$

It should be noted that the entries of the matrices  $(H)_{p,q}$  and  $(H_k)_{p,q}$  can be written with  $m_\beta$  satisfying

$$\deg(m_\beta) = \begin{cases} p + q & \text{for } (H)_{p,q} \\ p + q + 1 & \text{for } (H_k)_{p,q} \end{cases}. \quad (\text{A.16})$$

Using the matrices  $(H)_{p,q}$  and  $(H_k)_{p,q}$ , we define the following block Hankel matrices  $H^{(\gamma_1)}(\mathbf{m}, \mathbf{u})$  and  $H_k^{(\gamma_2)}(\mathbf{m}, \mathbf{u})$  ( $k = 1, 2, \dots, n$ ).

$$H^{(\gamma_1)}(\mathbf{m}, \mathbf{u}) := \begin{bmatrix} (H)_{0,0} & (H)_{0,1} & \cdots & (H)_{0,\gamma_1} \\ (H)_{1,0} & (H)_{1,1} & \cdots & (H)_{1,\gamma_1} \\ \vdots & \vdots & \ddots & \vdots \\ (H)_{\gamma_1,0} & (H)_{\gamma_1,1} & \cdots & (H)_{\gamma_1,\gamma_1} \end{bmatrix}, \quad (\text{A.17})$$

$$H_k^{(\gamma_2)}(\mathbf{m}, \mathbf{u}) := \begin{bmatrix} (H_k)_{0,0} & (H_k)_{0,1} & \cdots & (H_k)_{0,\gamma_2} \\ (H_k)_{1,0} & (H_k)_{1,1} & \cdots & (H_k)_{1,\gamma_2} \\ \vdots & \vdots & \ddots & \vdots \\ (H_k)_{\gamma_2,0} & (H_k)_{\gamma_2,1} & \cdots & (H_k)_{\gamma_2,\gamma_2} \end{bmatrix}, \quad (\text{A.18})$$

where  $\gamma_1$  and  $\gamma_2$  are determined as follows.

$$\gamma_1 = \begin{cases} (\mu + 1)/2 & \text{(if } \mu \text{ is odd)} \\ \mu/2 & \text{(if } \mu \text{ is even)} \end{cases}, \quad (\text{A.19})$$

$$\gamma_2 = \begin{cases} (\mu - 1)/2 & \text{(if } \mu \text{ is odd)} \\ \mu/2 & \text{(if } \mu \text{ is even)} \end{cases}. \quad (\text{A.20})$$

Using  $H^{(\gamma_1)}(\mathbf{m}, \mathbf{u})$  and  $H_k^{(\gamma_2)}(\mathbf{m}, \mathbf{u})$ , the following Proposition A.1 shows linear matrix inequality (LMI) conditions that the moments of some (probability) measure on  $\mathbb{R}_+^n$  must satisfy.

**Proposition A.1.** [5, 6] The entries of a given vector  $[\mathbf{m}^T, \mathbf{u}^T]$  constitute moments of some measure defined on the positive orthant  $\mathbb{R}_+^n := \{\mathbf{x} \in \mathbb{R}^n \mid x_k \geq 0 \ (k = 1, 2, \dots, n)\}$  only if

$$H^{(\gamma_1)}(\mathbf{m}, \mathbf{u}) \geq O, \quad (\text{A.21})$$

$$H_k^{(\gamma_2)}(\mathbf{m}, \mathbf{u}) \geq O \quad (k = 1, 2, \dots, n). \quad (\text{A.22})$$

The conditions (A.21) and (A.22) are called a moment condition and localized moment conditions, respectively. Combining Proposition A.1 with (A.12), we have the Proposition in the main text.

**Proposition.** Consider the CME (A.1). The steady state moments of the probability distribution satisfies

$$\begin{aligned} \mathbf{0} &= A\mathbf{m} + B\mathbf{u}, \\ H^{(\gamma_1)}(\mathbf{m}^*, \mathbf{u}^*) &\geq O, \\ H_k^{(\gamma_2)}(\mathbf{m}^*, \mathbf{u}^*) &\geq O \quad (k = 1, 2, \dots, n). \end{aligned} \quad (\text{A.23})$$

This proposition implies that the moments of the stochastic chemical reaction governed by (A.1) lie in the semi-algebraic set specified by (A.23).

The semi-algebraic set (A.23) is non-increasing in the truncation order  $\mu$ , since the solution space of the moment equation (A.12) with the truncation order  $\mu + 1$  is included in that with the truncation order  $\mu$ , and

$$H^{(\gamma_1+1)}(\mathbf{m}, \mathbf{u}) \geq O \Rightarrow H^{(\gamma_1)}(\mathbf{m}, \mathbf{u}) \geq O \quad \text{and} \quad H_k^{(\gamma_1+1)}(\mathbf{m}, \mathbf{u}) \geq O \Rightarrow H_k^{(\gamma_1)}(\mathbf{m}, \mathbf{u}) \geq O \quad (\text{A.24})$$

hold. It should be noted that the conditions (A.21) and (A.22) with  $\gamma_1, \gamma_2 \rightarrow \infty$  become both necessary and sufficient for the sequence  $(m_\beta)_{\beta \in \mathbb{N}_0}$  to be the moments of some measure on  $\mathbb{R}_+$  in the case of  $n = 1$  random variable ( $n = 1$  molecule) [5]. In general, however, the set (A.23) is not guaranteed to be a strictly decreasing set with  $\mu$ , implying that the convergence of the variable  $\mathbf{m}$  to a single value is not guaranteed.

## S. 5 Optimization programs for computing bounds of descriptive statistics

The Proposition in Section S.4 implies that the lower and the upper bounds of the steady state statistics, say  $f(\mathbf{m}, \mathbf{u})$ , can be calculated by solving the following optimization problems, respectively.

$$\begin{aligned} & \min f(\mathbf{m}, \mathbf{u}) \\ & \text{subject to (A.23),} \end{aligned} \tag{A.25}$$

$$\begin{aligned} & \max f(\mathbf{m}, \mathbf{u}) \\ & \text{subject to (A.23).} \end{aligned} \tag{A.26}$$

In what follows, we show more specific forms of optimization programs for computing the bounds of statistics.

### Upper and lower bounds of mean

The objective function  $f(\mathbf{m}, \mathbf{u})$  is obviously linear in this case. Thus, the problem falls into the class of semidefinite programming. For example, if one wants to compute the mean of the copy number of the  $j$ -th molecule, the objective function is

$$f(\mathbf{m}, \mathbf{u}) = \mathbb{E}[x_j] = m_{e_j}, \tag{A.27}$$

where  $e_j$  is the row vector whose  $j$ -th entry is 1 and 0 elsewhere. More generally, we can compute linear combinations of mean values by solving the following optimization problem.

$$\begin{aligned} & \min \quad \mathbf{c}^T \mathbf{m}_1 \\ & \text{subject to (A.23),} \end{aligned}$$

where  $\mathbf{c}$  is a given constant vector and  $\mathbf{m}_1 = \mathbb{E}[\mathbf{x}]$ , which is a vector that contains all of the first order raw moments. Note that flipping the sign of  $\mathbf{c}$  leads to the maximization problem.

### Upper bound of variance

We consider to compute the upper bound of

$$f(\mathbf{m}, \mathbf{u}) = \mathbf{c}^T \mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T] \mathbf{c}, \tag{A.28}$$

where  $\mathbf{c}$  is a given constant vector,  $\mathbb{E}[\mathbf{x}] = \mathbf{m}_1$ , and  $\mathbb{E}[\mathbf{x}\mathbf{x}^T]$  is a matrix whose entries are  $m_\alpha$  with  $\deg(m_\alpha) = 2$ . The function (A.28) represents linear combinations of the variance. Note that the variance of the copy number of the  $j$ -th molecule is given by

$$f(\mathbf{m}, \mathbf{u}) = \mathbf{e}_j^T \mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T] \mathbf{e}_j = m_{2e_j} - m_{e_j}^2 \tag{A.29}$$

by letting  $\mathbf{c} = \mathbf{e}_j$ . Since the second term of the objective function  $f(\mathbf{m}, \mathbf{u}) = m_{2\mathbf{e}_j} - m_{\mathbf{e}_j}^2$  is quadratic, the optimization problem (A.26) is seemingly not semidefinite programming. However, we can use Schur complement [1] to convert the optimization problem (A.26) into an equivalent semidefinite program.

**Theorem 1.** The solution of the optimization problem (A.26) with the objective function (A.28) is the same as that of the following optimization problem.

$$\begin{aligned} & \max_{V, \mathbf{m}, \mathbf{u}} && \mathbf{c}^T V \mathbf{c} \\ & \text{subject to} && \text{(A.23)} \end{aligned} \tag{A.30}$$

$$\begin{bmatrix} \mathbb{E}[\mathbf{x}\mathbf{x}^T] - V & \mathbb{E}[\mathbf{x}] \\ \mathbb{E}[\mathbf{x}]^T & 1 \end{bmatrix} \geq O,$$

where  $V$  is a symmetric matrix whose size is the same as  $\mathbb{E}[\mathbf{x}\mathbf{x}^T]$ .

**Proof** Consider a matrix  $V$  such that  $\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T] \geq V$ . It then follows that

$$\mathbf{c}^T \mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T] \mathbf{c} \geq \mathbf{c}^T V \mathbf{c}. \tag{A.31}$$

This implies that the solution of the maximization problem gives the maximum value of the function  $f(\mathbf{m}, \mathbf{u})$  in (A.28). Applying Schur complement [1] to

$$(\mathbb{E}[\mathbf{x}\mathbf{x}^T] - V) - \mathbb{E}[\mathbf{x}]\mathbb{E}[\mathbf{x}]^T \geq O, \tag{A.32}$$

which is equivalent to  $\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^T] \geq V$ , we have an equivalent semidefinite inequality condition

$$\begin{bmatrix} \mathbb{E}[\mathbf{x}\mathbf{x}^T] - V & \mathbb{E}[\mathbf{x}] \\ \mathbb{E}[\mathbf{x}]^T & 1 \end{bmatrix} \geq O. \tag{A.33}$$

□

## Lower bound of variance

The lower bound of the variance is not convex and thus the minimization problem of the variance cannot be formulated as a semidefinite program. In this work, we computed the lower bound of the variance in Fig. 5(C) by solving multiple SDPs. Specifically, we first solved the maximization problem of the mean value  $\mathbb{E}[x_6]$ , whose optimal value is denoted by  $(\mathbb{E}[x_6])_{\max}$ . Then, we solved the minimization problem of  $\mathbb{E}[x_6^2] - (\mathbb{E}[x_6])_{\max} \mathbb{E}[x_6]$ . Both optimization problems can be formulated as SDP by simply using the linear objective functions  $\mathbb{E}[x_6]$  and  $\mathbb{E}[x_6^2] - (\mathbb{E}[x_6])_{\max} \mathbb{E}[x_6]$ , respectively.



## Upper bound of the coefficient of variation

The coefficient of variation (CV) of the copy number of the  $j$ -th molecule is defined by

$$f(\mathbf{m}, \mathbf{u}) = \frac{\sqrt{m_{2e_j} - m_{e_j}^2}}{m_{e_j}}. \quad (\text{A.34})$$

The following theorem shows a semidefinite program to compute the upper bound of the CV.

**Theorem 2.** The upper bound of the CV (A.34),  $\text{CV}_{\max}$ , is given by

$$\text{CV}_{\max} = \sqrt{\frac{1}{\gamma} - 1}, \quad (\text{A.35})$$

where  $\gamma$  is the solution of the following minimization problem.

$$\begin{aligned} \min_{\gamma, \mathbf{m}, \mathbf{u}} \quad & \gamma \\ \text{subject to} \quad & (\text{A.23}) \\ & \begin{bmatrix} m_{2e_j} & m_{e_j} \\ m_{e_j} & \gamma \end{bmatrix} \geq O \\ & m_{e_j} > 0 \end{aligned}$$

**Proof** We consider to compute the upper bound of  $f(\mathbf{m}, \mathbf{u})$  by minimizing  $\gamma$  subject to (A.23) and

$$f(\mathbf{m}, \mathbf{u}) = \frac{\sqrt{m_{2e_j} - m_{e_j}^2}}{m_{e_j}} \geq \sqrt{\frac{1}{\gamma} - 1} \quad (\text{A.36})$$

with  $0 < \gamma \leq 1$ . Taking the square of both sides, we have equivalent conditions

$$m_{2e_j} - \frac{1}{\gamma} m_{e_j}^2 \geq 0, \quad 0 < \gamma \leq 1, \quad (\text{A.37})$$

$$m_{e_j} > 0. \quad (\text{A.38})$$

The inequality (A.37) then leads to an equivalent linear matrix inequality (LMI)

$$\begin{bmatrix} m_{2e_j} & m_{e_j} \\ m_{e_j} & \gamma \end{bmatrix} \geq O, \quad \gamma \leq 1. \quad (\text{A.39})$$

by Schur complement. Thus, the maximum value of  $f(\mathbf{m}, \mathbf{u})$  is obtained by minimizing  $\gamma$  subject to (A.23), (A.38) and (A.39). Since  $H^{(1)}(\mathbf{m}, \mathbf{u}) \geq O$  implies  $m_{2e_j} - m_{e_j}^2 \geq 0$  by taking Schur complement,  $\gamma \leq 1$  is redundant, which concludes the proof.  $\square$

### Lower bound of the coefficient of variation

The minimization problem of the coefficient of variation is not convex. Thus, it cannot be directly obtained by SDP. In this work, we independently computed the upper bound of the denominator and the lower bound of the numerator of (A.34) to obtain the lower bound of the coefficient of variation. Note that each optimization problem is SDP.

### Upper bound of Fano factor

Fano factor of the copy number of the  $j$ -th molecule is given by

$$f(\mathbf{m}, \mathbf{u}) = \frac{m_{2e_j} - m_{e_j}^2}{m_{e_j}}. \quad (\text{A.40})$$

To derive a semidefinite program for computing the upper bound of  $f(\mathbf{m}, \mathbf{u})$ , we define normalized moment vectors  $\mathbf{m}' := \mathbf{m}/m_{e_j}$  and  $\mathbf{u}' := \mathbf{u}/m_{e_j}$ , and a scalar  $m'_{2e_j} := m_{2e_j}/m_{e_j}$ . Note that one of the entries of the vector  $\mathbf{m}'$  is one.

**Theorem 3.** The upper bound of the Fano factor (A.40) is given by the solution of the following maximization problem.

$$\begin{aligned} & \max_{\gamma, \mathbf{m}', \mathbf{u}', m_{e_j}} && \gamma \\ & \text{subject to} && A\mathbf{m}' + B\mathbf{u}' = 0 \\ & && H^{(\gamma_1)}(\mathbf{m}', \mathbf{u}') \geq O, \quad H_k^{(\gamma_2)}(\mathbf{m}', \mathbf{u}') \geq O \quad (k = 1, 2, \dots, n) \\ & && m'_{2e_j} - m_{e_j} \geq \gamma, \\ & && m_{e_j} > 0 \end{aligned} \quad (\text{A.41})$$

**Proof** The upper bound of the Fano factor (A.40) is given by the maximum value of  $\gamma$  subject to (A.23) and

$$f(\mathbf{m}, \mathbf{u}) = \frac{m_{2e_j} - m_{e_j}^2}{m_{e_j}} = m'_{2e_j} - m_{e_j} \geq \gamma. \quad (\text{A.42})$$

Normalizing the vectors  $\mathbf{m}$  and  $\mathbf{u}$  of the inequality conditions (A.23) by  $m_{e_j} (> 0)$ , we have

$$\begin{aligned} & A\mathbf{m}' + B\mathbf{u}' = 0 \\ & H^{(\gamma_1)}(\mathbf{m}', \mathbf{u}') \geq O, \quad H_k^{(\gamma_2)}(\mathbf{m}', \mathbf{u}') \geq O \quad (k = 1, 2, \dots, n), \end{aligned}$$

which concludes the proof. □

## Largest confidence ellipsoids

We consider the largest confidence ellipsoids as illustrated in Fig. S.1 based on the covariance matrix of the  $j_1$ -th and the  $j_2$ -th molecule. The confidence ellipsoids show confidence intervals of molecular copy numbers assuming the distribution is Gaussian and are useful to examine the correlation between the copy numbers of two molecules.

The covariance matrix of the  $j_1$ -th and the  $j_2$ -th molecules is given by

$$\begin{aligned} \text{Cov} &:= \mathbb{E} \left[ \left( \mathbf{x}_{j_1 j_2} - \mathbb{E}[\mathbf{x}_{j_1 j_2}] \right) \left( \mathbf{x}_{j_1 j_2} - \mathbb{E}[\mathbf{x}_{j_1 j_2}] \right)^{\top} \right] \\ &= \begin{bmatrix} m_{2e_{j_1}} - m_{e_{j_2}}^2 & m_{e_{j_1}+e_{j_2}} - m_{e_{j_1}} m_{e_{j_2}} \\ m_{e_{j_1}+e_{j_2}} - m_{e_{j_1}} m_{e_{j_2}} & m_{2e_{j_1}} - m_{e_{j_2}}^2 \end{bmatrix}, \end{aligned} \quad (\text{A.43})$$

where  $\mathbf{x}_{ij} = [x_{j_1} \ x_{j_2}]^{\top}$ . Note that the covariance matrix is always positive semidefinite from its definition. The confidence ellipsoid is then obtained by computing the eigenvalues and eigenvectors of the matrix Cov. Specifically, the eigenvalues  $\lambda_1$  and  $\lambda_2$  give the length of minor and major axes and the associated eigenvectors give the directions of the two axes. The following semidefinite program calculates the area of the largest confidence ellipsoid and its associated covariance matrix Cov.

**Theorem 4.** The area of the largest confidence ellipsoid is the fourth power of the solution of the following optimization problem. The associated covariance matrix Cov is given by  $M_2 - M_1 M_1^{\top}$ .

$$\begin{aligned} & \max_{\gamma, L, m, u} && \gamma \\ & \text{subject to} && (\text{A.23}) \\ & && \begin{bmatrix} M_2 & L & M_1 \\ L^{\top} & I & \mathbf{0} \\ M_1^{\top} & \mathbf{0}^{\top} & 1 \end{bmatrix} \geq O \\ & && \begin{bmatrix} L_{11} & \gamma \\ \gamma & L_{22} \end{bmatrix} \geq O, \end{aligned} \quad (\text{A.44})$$

where  $M_1, M_2$  are defined by

$$M_1 = [m_{e_{j_1}} \ m_{e_{j_2}}]^{\top}, \quad M_2 = \begin{bmatrix} m_{2e_{j_1}} & m_{e_{j_1}+e_{j_2}} \\ m_{e_{j_1}+e_{j_2}} & m_{2e_{j_2}} \end{bmatrix},$$

and the variable  $L \in \mathbb{R}^{2 \times 2}$  is a lower triangular matrix.  $L_{11}$  and  $L_{22}$  represent the (1,1) and (2,2) entries of the matrix  $L$ , respectively.

**Proof** Applying Schur complement [1] to the inequality (A.44), we have

$$M_2 - LL^{\top} - M_1 M_1^{\top} = \text{Cov} - LL^{\top} \geq O. \quad (\text{A.45})$$

This implies

$$\det(\text{Cov}) \geq \{\det(L)\}^2 \tag{A.46}$$

since there is an orthogonal matrix  $U$  that diagonalizes the symmetric matrix  $\text{Cov}$  by  $U^T(\text{Cov})U = \text{diag}(\lambda_1, \lambda_2)$ , and

$$\det(\text{Cov}) = \lambda_1\lambda_2 \geq \det(U^T LL^T U) = \det(LL^T) = \{\det(L)\}^2. \tag{A.47}$$

In fact, we can always find a lower triangular matrix  $L$  that satisfies  $\text{Cov} = LL^T$  by Cholesky decomposition, and thus the equality is always attained by maximizing the determinant of  $LL^T$ , which is given by  $(L_{11}L_{22})^2$  since  $L$  is a lower triangular matrix. Thus,  $\det(\text{Cov})$  is the fourth power of the solution  $\gamma$  of the following optimization program.

$$\begin{aligned} & \max_{\gamma, L, m, u} \quad \gamma \\ & \text{subject to} \quad (A.23) \\ & \quad \quad \quad (A.44) \\ & \quad \quad \quad (L_{11}L_{22}) \geq \gamma^2, \quad L_{11} \geq 0, \quad L_{22} \geq 0 \end{aligned} \tag{A.48}$$

The proof is complete by applying Schur complement to (A.48). □

## S. 6 Transfer curve of the repressor protein

We investigated the relation between the mean copy number of the reporter  $\mathbb{E}[x_6]$  and the repressor  $\mathbb{E}[x_2]$  for the reaction model in Fig. 4(A). The mean copy number of the reporter is given by

$$\mathbb{E}[x_6] = \frac{k_7 k_9}{k_8 k_{10}} \mathbb{E}[x_4], \quad (\text{A.49})$$

where  $\mathbb{E}[x_4]$  follows the moment equation

$$\frac{d}{dt} \mathbb{E}[x_4] = k_6 D_{T_2} - k_6 \mathbb{E}[x_4] - k_5 \mathbb{E}[x_2 x_4]. \quad (\text{A.50})$$

Equation (A.50) implies that  $\mathbb{E}[x_4]$  depends on the second order moment  $\mathbb{E}[x_2 x_4]$ , and thus, higher order moments are necessary to fully characterize  $\mathbb{E}[x_4]$ . In other words, the mean copy number of the reporter  $\mathbb{E}[x_6]$  cannot be determined simply from the mean copy number of the repressor  $\mathbb{E}[x_2]$  but it requires higher order statistics.

This is in contrast with the deterministic modeling, where the steady state concentration of the free DNA  $x_4$  is expressed by the Michaelis-Menten equation

$$x_4 = \frac{K}{K + x_2} D_{T_2} \quad (\text{A.51})$$

with  $K = k_6/k_5$ .

In Fig. S.3, we characterized the reporter expression level versus the repressor copy number by solving the SDP for  $\mathbb{E}[x_4]$ . The figure clearly illustrates that the reporter copy number deviates from the Michaelis-Menten kinetics with the maximum relative error of 6%, suggesting that the simple Michaelis-Menten kinetics is erroneous especially when the biocircuit is highly stochastic.

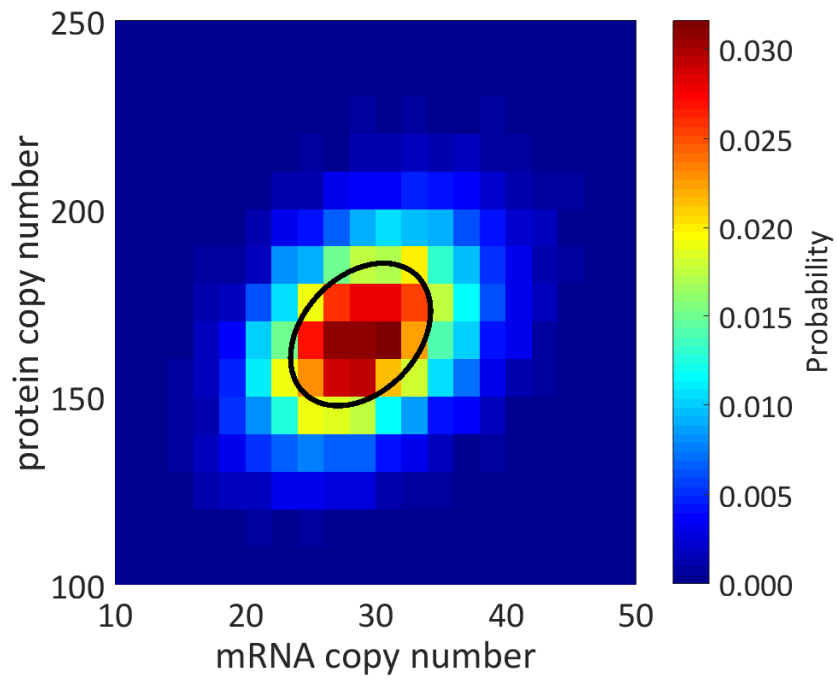


Figure S.1: A confidence ellipsoid of mRNA and protein copy numbers in the transcription-translation biocircuit in Fig. 2A.

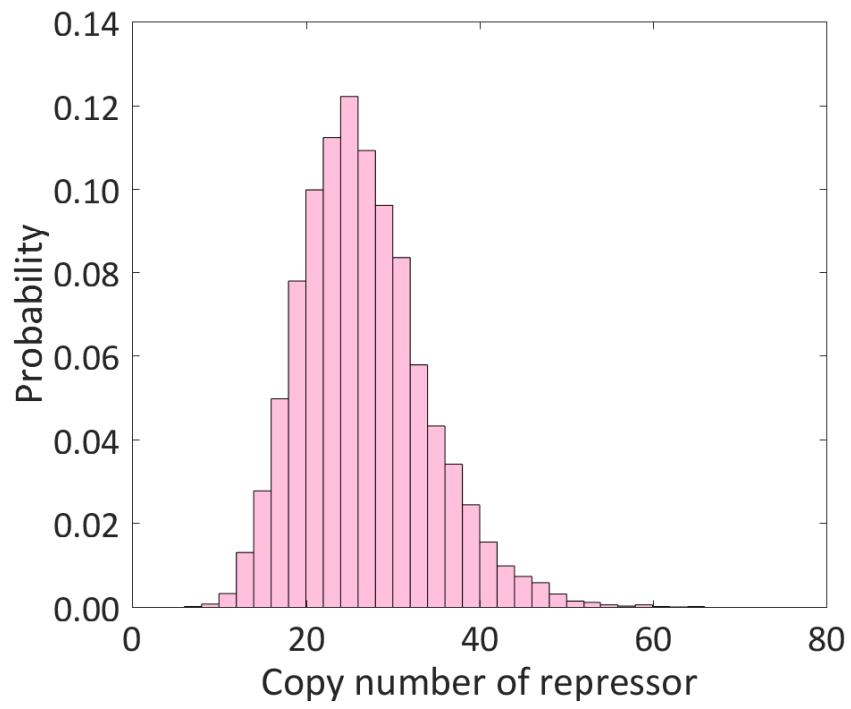


Figure S.2: The steady state distribution of the repressor copy number

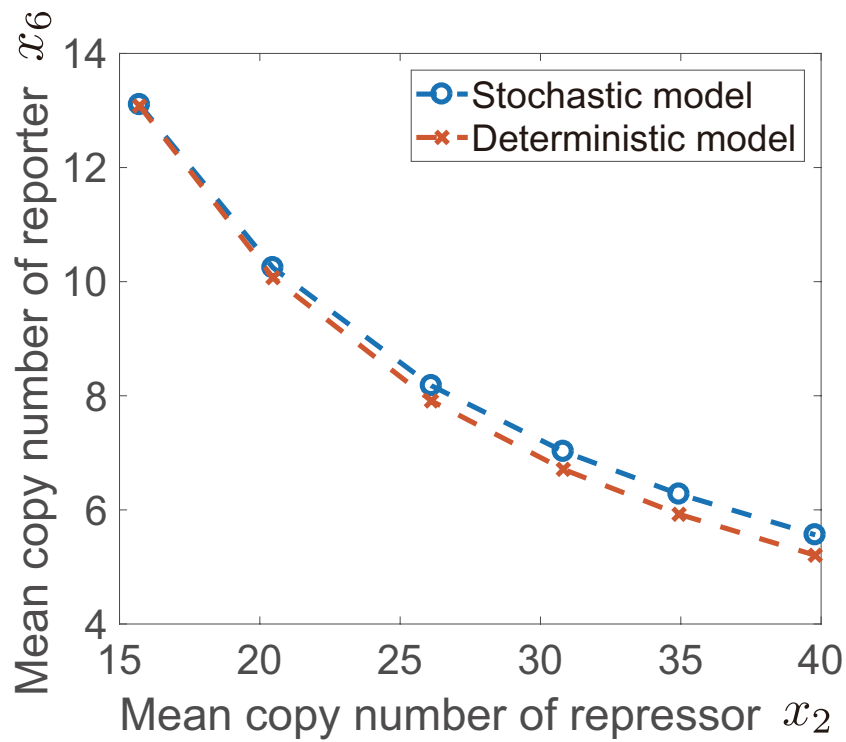


Figure S.3: A transfer curve showing the mean copy number of the reporter protein with respect to the repressor (stochastic model). The expression level of the reporter is slightly higher than that predicted by deterministic Michaelis-Menten kinetics.



## References

- [1] S. P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [2] E. T. Denisov, O. M. Sarkisov, and G. I. Likhteshtein. *Chemical kinetics: fundamentals and new developments*. Elsevier, 2003.
- [3] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, 1976.
- [4] D. T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A*, 188(1–3):404–425, 1992.
- [5] H. J. Landau. *Moments in Mathematics*. American Mathematical Society, 1987.
- [6] J. A. Shohat and J. D. Tamarkin. *The problem of moments*. American Mathematical Society, 1943.