On Projection-Based Model Reduction of Biochemical Networks
Part II: The Stochastic Case

Aivar Sootla[1] and James Anderson[2]

Abstract— In this paper, we consider the problem of model order reduction of stochastic biochemical networks. In particular, we reduce the order of (the number of equations in) the Linear Noise Approximation of the Chemical Master Equation, which is often used to describe biochemical networks. In contrast to other biochemical network reduction methods, the presented one is projection-based. Projection-based methods are powerful tools, but the cost of their use is the loss of physical interpretation of the nodes in the network. In order alleviate this drawback, we employ structured projectors, which means that some nodes in the network will keep their physical interpretation. For many models in engineering, finding structured projectors is not always feasible; however, in the context of biochemical networks it is much more likely as the networks are often (almost) monotonic. To summarise, the method can serve as a trade-off between approximation quality and physical interpretation, which is illustrated on numerical examples.

Index Terms— model order reduction; structured model order reduction; stochastic averaging principle; linear noise approximation; chemical master equation

I. INTRODUCTION AND PRELIMINARIES

In mathematical biology, one of the most common approaches to model reduction of deterministic dynamical systems is to apply Tikhonov's theorem [1], this is also referred to as time-scale separation or the quasi-steady state assumption. In order to apply the method, it is assumed that some variables reach steady state faster than others. Typically, it is said that the system operates on a slow and a fast time-scale. Tikhonov's theorem provides conditions for a systematic elimination of the fast variables from the system without affecting the steady state solution for the slow variables. A technique similar to time-scale separation has also been developed for Stochastic Differential Equations (SDEs) and is known as the stochastic averaging principle [2]. The main idea of averaging is integrating out the fast variables from the system, which in the deterministic case reduces to time-scale separation under certain assumptions.

The framework based on time-scale separation has also received considerable attention in control theory [3], however, projection-based methods (see, e.g. [4], [5]) have now become the preferred method. This shift occurred because the latter proved itself more powerful than the former. In any projection-based method first an appropriate state-space transformation is computed, which can create a larger number of equations operating on a faster time scale. Only then tools similar to time-scale separation are applied. Having the freedom to apply such transformations results in the following advantages: existence of error bounds (in some cases), automatic selection of reduced states, higher (input-output) approximation quality.

In the context of biochemical networks, we are dealing with the so-called Chemical Master Equation (CME), a solution to which is a continuous-time Markov chain with an infinite number of states. To the authors' best knowledge there are no tractable time-scale separation or averaging methods applicable to CMEs. Hence reduced order modelling is typically applied to different approximations of the CME such as Finite State Projection [6], [7], Linear Noise Approximation (LNA) [8] or macroscopic deterministic models of reaction networks [9], [10]. In this paper, we will compute reduced order models for the LNA, which is a collection of two systems: a deterministic one describing macroscopic concentrations and an SDE driven by a Gaussian noise describing the fluctuations about the macroscopic concentrations. We will use the techniques described in [8] for reducing the number of equations in the system, but first we will employ control theoretic projection methods in order to create better conditions for averaging. The focus of this paper is towards the stochastic model reduction, in Part I [11] the details of the structured projectors are given and proofs included[1].

Applying a state-space transformation results in the

[1]Whilst both papers are self contained the first paper is available on line at http://arxiv.org/abs/1403.3579

states of the reduced order model which are composed of linear combinations of the states of the full order model. Hence the new states lack physical interpretation and sparsity patterns in the network are lost, which constitutes a major drawback of the projection-based methods. In order to limit the loss of physical interpretation, we partition the states into two groups: one group of states is left intact, and the other states are linear combinations of some original states. The transformations are computed according to the partitioning based on the covariance matrix of fluctuations about the macroscopic concentrations of the model. This covariance matrix approximates the statistics of the solution to the CME under certain conditions [12]. Therefore, the reduced order model can potentially capture additional information about the stochastic nature of the model in comparison with other approaches.

The idea of using structured transformations in order to preserve the physical meaning of the states is not new (cf. [13]). Moreover, the class of models, for which such transformations can be computed, is not rich and not many necessary conditions for their existence are known. However, in the context of biochemical networks for the systems with monotone dynamics such transformations can always be computed. In Part I [11] we formally define the concept of monotonicity, for the definition see [14].

In this paper, we notice that the proposed method can be applied to a more general class of networks. Many biochemical networks have a tendency to be nearly monotone [15], that is if a small number of edges are removed, then the network becomes monotone. Since the monotonicity implies the existence of structured projectors, one can make a conjecture on the existence of such projectors for a large class of biochemical networks. To illustrate this observation, we apply the method to a model of yeast glycolysis [16], which becomes monotone after removing just three reactions (out of more than thirty). For this model, structured projectors exist for any state-space partitioning, which is a very strong property.

The paper is organised as follows. In the next subsections, we introduce the CME, LNA and stochastic modeling tools. In Section II, the proposed model reduction method is described, which is validated on numerical examples in Section III.

A. Modelling Biochemical Networks

Biochemical networks are typically modelled by a continuous time infinite Markov chain, probability distribution function which is computed by a Chemical Master Equation (CME):

$$\frac{\partial \mathbb{P}(n,t)}{\partial t} = \Omega \sum_{i=1}^{R} (\hat{f}(n - S_i, \Omega) \mathbb{P}(n - S_i, t) -$$
$$\hat{f}(n, \Omega) \mathbb{P}(n, t), \quad (1)$$

where R is the number of reactions; column vectors S_i form a stoichiometry matrix S; $\hat{f}$ is a vector containing the reaction rates $\hat{f}_i$; n is a vector containing the number of molecules n_j of species j; Ω is a volume of a compartment where reactions are occurring; finally, $\mathbb{P}(n, t)$ is the probability of the vector of the number of molecules equal to n at time t. This equation cannot be solved analytically except for a handful of cases and numerical simulations are extremely expensive. In order to lower the complexity of simulations, different approximations of a CME are often derived, for example, the Linear Noise Approximation (LNA) [12]. The major assumption in the LNA is as follows:

$$\frac{n}{\Omega} = x + \Omega^{-1/2}\eta,$$

where x is a vector of macroscopic concentrations of the species, η is a vector of stochastic fluctuations about x. Additionally, if we assume a large number of reactions occurring per unit time, it can be shown, that the fluctuations η and macroscopic concentrations x obey the following equations:

$$\dot{\eta} = J(x)\eta + \Omega^{-1/2}SF(x)\Gamma, \quad (2)$$
$$\dot{x} = Sf(x), \quad (3)$$

where J is the Jacobian of $Sf(x)$, $Sf(x)$ and $J(x)\eta$ are the first terms in the Taylor expansion of $\hat{f}(n, \Omega)$, Γ is a Gaussian noise, and $F = \text{diag}(\sqrt{f(x)})$. Note that the matrices J, F do not depend on the fluctuations η, but depend only on the macroscopic concentrations x, which is computed using deterministic differential equations. The covariance, X, of η is computed by solving (or simulating) the Lyapunov differential equation:

$$JX + XJ^T + \Omega^{-1}SF^2S = \frac{dX}{dt}. \quad (4)$$

As a final remark, note that the solution to (2), (3) is a vector Markov process with elements $y_i(t)$, which at every time t are normally distributed with mean equal to $x_i(t)$, variance equal to $X_{ii}(t)$ and covariances $\text{cov}(y_i(t)y_j(t))$ equal to $X_{ij}(t)$.

The macroscopic reaction rates are not affected at all by the fluctuations dynamics, hence the time-scale separation can be applied directly to (3). In order to reduce the order of the fluctuation dynamics, we apply a method [8], which can be seen as a version of the

averaging principle. This result is presented here for completeness. Consider the system:

$$\dot{\eta}_s = J_{ss}\eta_s + J_{sf}\eta_f + S_sF\Gamma,$$
$$\varepsilon\dot{\eta}_f = J_{fs}\eta_s + J_{ff}\eta_f + S_fF\Gamma, \tag{5}$$

where

$$J = \begin{pmatrix} J_{ss} & J_{fs} \\ J_{sf} & J_{ff} \end{pmatrix} \quad S = \begin{pmatrix} S_s \\ S_f \end{pmatrix}$$

Note that the s and f subscripts denote slow and fast states respectively. Then the reduced order linear noise approximation of the fluctuations corresponding to the slow dynamics η_s can be obtained as follows:

$$\dot{\eta}_s = (J_{ss} - J_{sf}J_{ff}^{-1}J_{fs})\eta_f + (S_s - J_{sf}J_{ff}^{-1}S_f)F\Gamma. \tag{6}$$

II. Model Reduction Methods

In this paper, it is proposed to employ stochastic averaging coupled with a projection approach in order to obtain a reduced order model. First, a particular state-space transformation T will be applied to species n in CME (1) resulting in the new species m such that $m = Tn$. The transformation T is computed based on the covariance matrix of η. Finally, we simply apply averaging as described above in order to reduce the fast species in m.

A. CME under a State-Space Transformation

First, we show that based on the solution of the CME in the species m, it is possible to reconstruct the solution in the species n.

Lemma 1: Assume the variable n satisfies the Master equation:

$$\frac{\partial \mathbb{P}(n,t)}{\partial t} = \Omega \sum_{j=1}^{R} (\hat{f}(n - S_i, \Omega) - \hat{f}(n, \Omega))\mathbb{P}(n,t).$$

Then the variable $m = Tn$ satisfies the following equation:

$$\frac{\partial \mathbb{P}(m,t)}{\partial t} = \Omega \sum_{j=1}^{R} (\tilde{f}(m - \tilde{S}_i, \Omega) - \tilde{f}(m, \Omega))\mathbb{P}(m,t),$$

where T is an invertible constant matrix, $\tilde{S} = TS$, $\tilde{f}(m, \Omega) = f(T^{-1}m, \Omega)$.

Proof: Consider the term $(\hat{f}(n - S_i, \Omega) - \hat{f}(n, \Omega))$ and make a substitution $n = T^{-1}m$. Now

$$\hat{f}(n - S_i, \Omega) - \hat{f}(n, \Omega) = \hat{f}(T^{-1}m - S_i, \Omega) - $$
$$\hat{f}(T^{-1}m, \Omega) = \hat{f}(T^{-1}(m - TS_i), \Omega) - $$
$$\hat{f}(T^{-1}m, \Omega) = \tilde{f}(m - \tilde{S}_i, \Omega) - \tilde{f}(m, \Omega)$$

Finally, in order to prove the claim, note that $\mathbb{P}(n,t)$ is equal to $\mathbb{P}(m,t)$. ∎

B. Structured Transformations

In order to compute a transformation T we require an output y as a function of x according to classical methods [4], [5]. The output in our case can be chosen simply as observations of the states of particular interest. Assume that we are interested in the behaviour of the first l species, then define the matrix $C := \begin{pmatrix} I_l & 0_{l,k} \end{pmatrix}$, where $l + k$ is the total number of species. Now the outputs are chosen as follows: the output of the macroscopic concentrations y^d is equal to Cx, and the output for the fluctuations y^f is equal to $C\eta$.

We propose to compute a transformation T based on the covariances in a steady-state x_{ss} of (3). Hence, the obtained approximation will be around the steady-state x_{ss}. Consider a system:

$$\dot{\nu} = A\nu + B\Gamma$$
$$y^f = C\nu \tag{7}$$

where the drift matrix $A = J(x_{ss})$ is asymptotically stable and $B = \Omega^{-1/2}SF(x_{ss})$. Assume also that the system (7) is partitioned as follows

$$\nu = \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \; A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \; B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \; C^T = \begin{pmatrix} I_l \\ 0_{k,l} \end{pmatrix}, \tag{8}$$

where $\nu_1 \in \mathbb{R}^l$, $\nu_2 \in \mathbb{R}^k$, and the matrices A, B and C are partitioned according to the same dimensions. The next step is to compute structured Gramians, which are obtained as solutions to Lyapunov inequalities

$$AP + PA^T + BB^T \leq 0$$
$$QA + A^TQ + C^TC \leq 0 \tag{9}$$

subject to the same partitioning as the states:

$$P = \begin{pmatrix} P_{11} & 0_{l,k} \\ 0_{k,l} & P_{22} \end{pmatrix} \quad Q = \begin{pmatrix} Q_{11} & 0_{l,k} \\ 0_{k,l} & Q_{22} \end{pmatrix} \tag{10}$$

Note that P is an approximation of X from (4) around the steady-state x_{ss}. Given our assumption about the importance of the first l species, the transformation T is composed as follows:

$$T = \begin{pmatrix} I_l & 0_{l,k} \\ 0_{l,k} & T_{22} \end{pmatrix}$$

where T_{22} is such that

$$T_{22}^{-1}P_{22}T_{22}^{-T} = T_{22}^T Q_{22}T_{22} = \Sigma_{22},$$

and Σ_{22} is diagonal. According to standard tools [13], we choose the states to truncate according to the magnitude of the values of the diagonal of Σ_{22}. Assume r states are to be reduced, let W_{22} be the first $k - r$ columns of T_{22}, while W_{22}^r are the rest r columns of T_{22}. Let also V_{22} be the first $k-r$ columns of T_{22}^{-1}, while

V_{22}^r are the rest r columns of T_{22}^{-1}. Now, the projectors can be obtained as follows

$$W = \begin{pmatrix} I_l & 0_{l,k-r} \\ 0_{k-r,l} & W_{22} \end{pmatrix} \quad W_r = \begin{pmatrix} 0_{l,r} \\ W_{22}^r \end{pmatrix}$$

$$V = \begin{pmatrix} I_l & 0_{l,k-r} \\ 0_{k-r,l} & V_{22} \end{pmatrix} \quad V_r = \begin{pmatrix} 0_{l,r} \\ V_{22}^r \end{pmatrix} \qquad (11)$$

The existence of block-diagonal solutions to (10) cannot be guaranteed for general linear systems. It is known, however, that some classes of systems admit block-diagonal solutions to the Lyapunov inequalities. One such class is positive systems, that is systems with A matrix with non-negative off-diagonal entries (these matrices are called Metzler), B and C matrices with non-negative entries. A generalisation of positive systems to the non-linear case are monotone systems. This implies that for monotone systems such Gramians always exist. We discuss this further in Part I [11].

C. Reduced Order Models and Model Comparison

We assume that the full order model is defined as follows:

$$\begin{aligned} \dot{x} &= Sf(x), & y^d &= Cx \\ \dot{\eta} &= J(x)\eta + \Omega^{-1/2}SF(x)\Gamma, & y^f &= C\eta \\ x(0) &= x_0, \quad \eta(0) = 0 \end{aligned}$$

where y^d is "the output" of the macroscopic concentrations, y^f is "the output" of the fluctuations with a constant matrix C and x_0 is the initial state. Let z be a new variable equal to Tx, let also z_r, η_r be the species to be removed from the model, and z_m, η_m the states of the reduced order model. Now the equations approximating the full order dynamics (2-3) can be computed as follows:

$$\begin{aligned} \dot{z}_m &= S_r f_r(z_m), \\ y_r^d &= C_r z + D_r, \\ \dot{\eta}_m &= J_r(z_m)\eta_m + \Omega^{-1/2}S_r F_r(z_m)\Gamma, \qquad (12) \\ y_r^f &= C_r \eta \\ z_m(0) &= V x_0, \quad \eta_m(0) = 0 \end{aligned}$$

where $S_r = V^T S$, $f_r(z) = f(Wz + W_r z_r)$, $F_r(z) = F(Wz + W_r z_r)$, $J_r(z) = V^T J(Wz + W_r z_r)W$, $C_r = CW$ and $D_r = CW_r z_r$. Note that [8] fits the framework in (12) with the identity transformation T, and projectors W and V reducing particular states. We will take [8] as a baseline and compare it to the proposed method.

We compare separately the error in the macroscopic dynamics (mean) and the fluctuations (variance), since their dynamic models are decoupled. The error $y^d - y_r^d$ in macroscopic dynamics is computed by perturbing the initial state x_0 from the steady-state x_{ss} and measured

(a) Depiction of the network
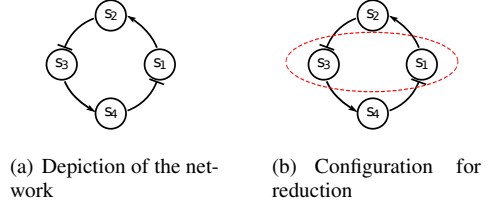
(b) Configuration for reduction

Fig. 1. Toy Example. In the left panel, the network is schematically depicted. In the right panel, the configuration for reduction is depicted. In this configuration species S_1 and S_3 are grouped together, while reducing one state

in L_1, L_2 and L_∞ norms (which for completeness we now define):

$$\begin{aligned} \|u\|_p &= \left(\int_{-\infty}^{\infty} |u(t)|_p^p dt \right)^{\frac{1}{p}} \quad \text{for } p \in \{1,2\}, \\ \|u\|_\infty &= \text{ess sup}_t |u(t)|_\infty. \end{aligned}$$

A comparison in terms of the fluctuations η is performed by computing the covariance matrix of the outputs y and y_r. For the full order model this matrix is computed as

$$\text{cov}^f = \text{cov}(y^f(y^f)^T) = C\text{cov}(\eta\eta^T)C^T = CXC^T$$

where X satisfies the Lyapunov equation (4). Similarly, the covariance matrix for the reduced order models cov_r^f can be computed.

III. EXAMPLES

A. Toy Example.

The first network we consider consists of only four species, see Figure 1(a). One can interpret the species S_1 and S_3 as mRNA, and S_2 and S_4 as the corresponding proteins. We refer the reader to Part I [11] for a full interpretation of the model.

$$\begin{aligned} \dot{m}_i &= \frac{c_{i1}}{1 + p_j^2} - c_{i2}m_i \\ \dot{p}_i &= c_{i3}m_i - c_{i4}p_i \end{aligned}$$

where c_{i1} are constants, m_1, m_2 are species S_1, S_3, p_1, p_2 are species S_2, S_4 and $i \in \{1,2\}$. We compare the simulation results for the full order model, the reduced order model obtained by [8], and the reduced order model obtained from reduction according to the configuration in Figure 1(b) with parameters

$$c_{1.} = c_{2.} = \begin{pmatrix} 3 & 4 & 1 & 0.2 \end{pmatrix}.$$

The method from [8] and the presented method produce very similar deterministic models, simulation of which is depicted in Figure 2(a). But the computation of
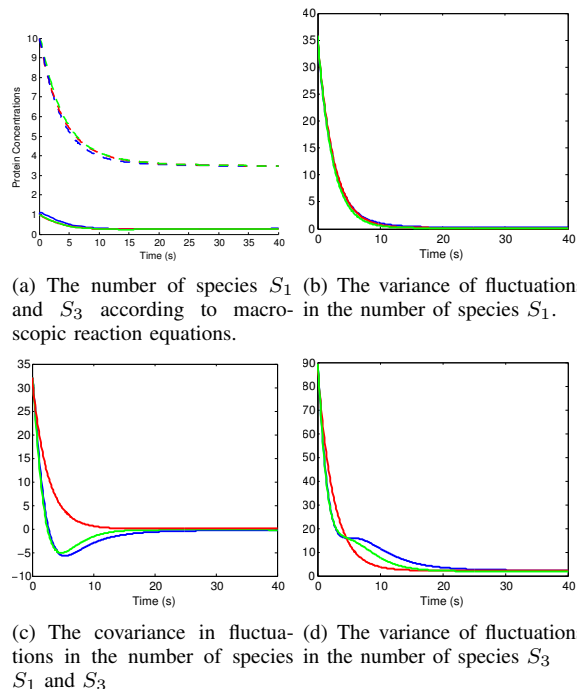
(a) The number of species S_1 and S_3 according to macroscopic reaction equations.

(b) The variance of fluctuations in the number of species S_1.

(c) The covariance in fluctuations in the number of species S_1 and S_3

(d) The variance of fluctuations in the number of species S_3

Fig. 2. The number of species S_1 and S_3 according to macroscopic (deterministic) reaction equations and the covariance matrix of these species. In all figures, the blue lines are obtained by simulating the full order model, the red lines by simulating the model obtained using [8], and the green lines by simulating the reduced order model with states lumped according to the configuration in Figure 1(b). In Figure 2(a) The dashed lines are the number of species of S_1, and the solid lines are the number of species of S_3. Even though macroscopic concentrations of species are almost the same for all the models, the covariances of the model obtained by [8] are quite different from the full order model covariances.

the covariance matrix of the fluctuations η, paints a different picture. The presented method provided the fluctuations with statistics very close the full order model statistics, which is not the case for the statistics of the model obtained by [8], see Figure 2. The presented reduction method uses the covariance matrices in order to compute the reduced order models. Moreover, the proposed method targets for reduction the species with small variances in the fluctuations about the macroscopic species concentrations. This outlines a major advantage of the proposed method.

B. Kinetic Model of Yeast Glycolysis. Non-Monotone Dynamics

The model was published in [16] and consists of twelve metabolites and four boundary fluxes. We model the network's response to change of glucose in the system as in [9]. Again we refer the reader to Part I [11] for a full model description, we point out that the Jacobian of the system is not Metzler, but there are only
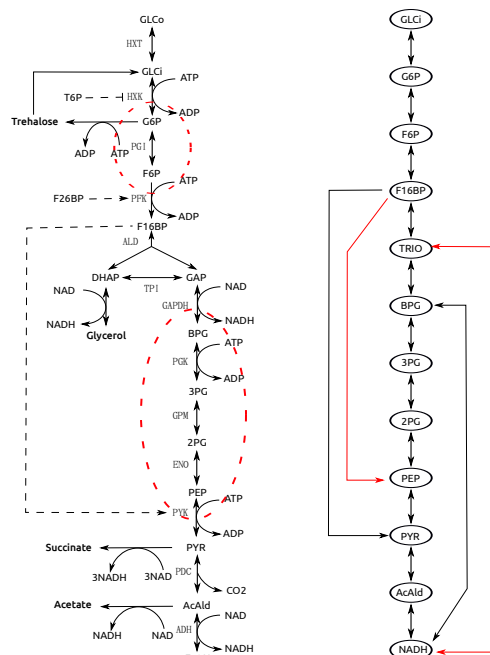

Fig. 3. Kinetic model of yeast glycolysis. In the left panel the biochemical graph is depicted. In the right panel a graph of dynamic interactions between metabolites. If the red connections are removed, the dynamics of the network would become monotone.

three negative off-diagonal entries.

Using this fact, it was not a great surprise that a linearised model around a steady-state would have block-diagonal Gramians with a sparsity pattern according to some state partitioning. However, the existence of diagonal Gramians was surprising. This meant that without any reservation we could approximate any group of states, while preserving the other states.

The simulation results are presented in Table I for various reduction configurations. We compare only the errors in the macroscopic concentrations. We apply [8] to metabolite concentrations, while using the proposed method we try to lump those metabolites in one state, so that the number of reduced states is similar in both cases. The first three rows of each sub-table in Table I can be compared directly, and it is clear that the proposed method performs better in terms of quality than [8].

The proposed method is also more flexible in terms of reduction choices. In the forth row of Table I-B, the region {BPG-PEP} contains four metabolites; however, we reduced only two states after computing the state-space transformation. In the fifth row, in the region {GLCi-F6P}, which contains three metabolites, we reduce just one state and this provides us with the best model among all the reduction attempts.

TABLE I

REDUCTION OF THE GLYCOLYSIS MODEL. THE ERROR OF THE
OUTPUT IS GIVEN IN DIFFERENT NORMS.

TABLE I-A. APPXOIMATION RESULTS OBTAINED BY USING [8]

States \ Error	L_1	L_2	L_∞
F6P, 2PG, PEP	1.21	23.6	0.98
F6P, 3PG, 2PG, PEP	1.56	33.6	1.55
G6P, F6P, 3PG, 2PG, PEP	2.05	36.3	1.59

TABLE I-B. REDUCTION BY $\{k_1, k_2\}$ STATES IN EVERY REGION

Lumped Region(s)	$\{k_1, k_2\}$	L_1	L_2	L_∞
{G6P, F6P}, {3PG-PEP}	{1, 2}	1.17	24.8	1.01
{G6P, F6P}, {BPG-PEP}	{1, 3}	1.41	22.9	0.86
{GLCi-F6P}, {BPG-PEP}	{2, 3}	1.23	18.7	0.64
{G6P, F6P}, {BPG-PEP}	{1, 2}	1.40	22.8	0.86
{GLCi-F6P}, {BPG-PEP}	{1, 2}	0.03	0.56	0.02

IV. CONCLUSION AND DISCUSSION

This paper provides a proof of concept of using projection-based methods for modelling stochastic biological systems. We illustrate on numerical examples that the proposed method is competitive with averaging in terms of estimation of mean and covariance matrix of a process. In combination with Part 1 [11] we have shown how to construct the structured projectors and shown that for monotone dynamical system we can preserve structure, and (locally) monotonicity.

An additional inquiry is required into existence of structured Gramians for the biochemical networks. In general, diagonal Gramians do not exist for all models; however, block-diagonal ones may be more common. In any case, even the initial intuition provides a great justification for the proposed reduction method.

Another interesting topic is an extension to non-constant transformations T (non-linear or time-varying). A non-linear structured balancing can be potentially applied to a Chemical Langevin Equation (CLE), which is a more representative approximation of CME in comparison to LNA [12].

An important topic for the proposed model reduction method is partitioning of a network. Some intuition can be gained through biological insights, for example, one can collapse a whole pathway into a couple of states. One can also employ tools from metabolic control analysis [17]. On the other hand, we can employ purely theoretical tools such as [18]–[21]. Currently, we are not investigating the problem of partitioning; however, it is an important future work direction.

REFERENCES

[1] A. Tikhonov, "Systems of differential equations containing small parameters in the derivatives," *Mat. Sbornik*, vol. 73, no. 3, pp. 575–586, 1952.

[2] R. Z. Khasminskii, "On the principle of averaging for the itô stochastic differential equations," *Kybernetika (Czechoslovakia)*, vol. 4, pp. 260–279, 1968.

[3] P. Kokotovic, H. K. Khalil, and J. O'Reilly, *Singular perturbation methods in control: analysis and design*. SIAM, 1987, vol. 25.

[4] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Autom. Control*, vol. 26, no. 1, pp. 17–32, Feb 1981.

[5] J. M. Scherpen, "Balancing for nonlinear systems," *Systems and Control Letters*, vol. 21, pp. 143–153, 1993.

[6] B. Munsky and M. Khammash, "The finite state projection algorithm for the solution of the chemical master equation," *J. Chem. Phys.*, vol. 124, no. 4, p. 044104, 2006.

[7] C. D. Pahlajani, P. J. Atzberger, and M. Khammash, "Stochastic reduction method for biological chemical kinetics using time-scale separation," *Journal of theoretical biology*, vol. 272, no. 1, pp. 96–112, 2011.

[8] P. Thomas, R. Grima, and A. V. Straube, "Rigorous elimination of fast stochastic variables from the linear noise approximation using projection operators," *Physical Review E*, vol. 86, no. 4, p. 041110, 2012.

[9] S. Rao, A. van der Schaft, K. van Eunen, B. M. Bakker, and B. Jayawardhana, "Model-order reduction of biochemical reaction networks," in *Proc. European Control Conf.*, Zurich, Switzerland, July 2012.

[10] M. Sunnåker, G. Cedersund, and M. Jirstrand, "A method for zooming of nonlinear models of biochemical systems," *BMC systems biology*, vol. 5, no. 1, p. 140, 2011.

[11] A. Sootla and J. Anderson, "On projection-based model reduction of biochemical networks Part I: The deterministic case," in *To appear in the 53rd Conf. Decision Control*, 2014.

[12] E. Wallace, D. Gillespie, K. Sanft, and L. Petzold, "Linear noise approximation is valid over limited times for any chemical system that is sufficiently large," *Systems Biology, IET*, vol. 6, no. 4, pp. 102–115, 2012.

[13] H. Sandberg and R. M. Murray, "Model reduction of interconnected linear systems using structured gramians," in *Proc. of the IFAC World Congress*, Seoul, Korea, July 2008, pp. 8725–8730.

[14] D. Angeli and E. Sontag, "Monotone control systems," *IEEE Trans. Autom. Control*, vol. 48, no. 10, pp. 1684–1698, 2003.

[15] E. D. Sontag, "Monotone and near-monotone biochemical networks," *Systems and Synthetic Biology*, vol. 1, no. 2, pp. 59–87, 2007.

[16] K. van Eunen, J. A. Kiewiet, H. V. Westerhoff, and B. M. Bakker, "Testing biochemistry revisited: how in vivo metabolism can be understood from in vitro enzyme kinetics," *PLoS Comp. Biol.*, vol. 8, no. 4, p. e1002483, 2012.

[17] E. Klipp, R. Herwig, A. Kowald, C. Wierling, and H. Lehrach, *Systems biology in practice: concepts, implementation and application*. John Wiley & Sons, 2008.

[18] A. Sootla and A. Rantzer, "Scalable positivity preserving model reduction using linear energy functions," in *Proc. Conf. Decision Control*, Dec. 2012, pp. 4285–4290.

[19] J.-C. Delvenne, M. T. Schaub, S. N. Yaliraki, and M. Barahona, "The stability of a graph partition: A dynamics-based framework for community detection," in *Dynamics On and Of Complex Networks, Volume 2*. Springer, 2013, pp. 221–242.

[20] M. Sezer and D. Šiljak, "Nested ε-decompositions and clustering of complex systems," *Automatica*, vol. 22, no. 3, pp. 321–331, 1986.

[21] J. Anderson and A. Papachristodoulou, "A decomposition technique for nonlinear dynamical system analysis," *IEEE Trans. Autom. Control*, vol. 57, no. 6, pp. 1516–1521, 2012.