BMC
Systems Biology

RESEARCH ARTICLE **Open Access**

A multiscale approximation in a heat shock response model of E. coli

Hye-Won Kang

Abstract

Background: A heat shock response model of *Escherichia coli* developed by Srivastava, Peterson, and Bentley (2001) has multiscale nature due to its species numbers and reaction rate constants varying over wide ranges. Applying the method of separation of time-scales and model reduction for stochastic reaction networks extended by Kang and Kurtz (2012), we approximate the chemical network in the heat shock response model.

Results: Scaling the species numbers and the rate constants by powers of the scaling parameter, we embed the model into a one-parameter family of models, each of which is a continuous-time Markov chain. Choosing an appropriate set of scaling exponents for the species numbers and for the rate constants satisfying balance conditions, the behavior of the full network in the time scales of interest is approximated by limiting models in three time scales. Due to the subset of species whose numbers are either approximated as constants or are averaged in terms of other species numbers, the limiting models are located on lower dimensional spaces than the full model and have a simpler structure than the full model does.

Conclusions: The goal of this paper is to illustrate how to apply the multiscale approximation method to the biological model with significant complexity. We applied the method to the heat shock response model involving 9 species and 18 reactions and derived simplified models in three time scales which capture the dynamics of the full model. Convergence of the scaled species numbers to their limit is obtained and errors between the scaled species numbers and their limit are estimated using the central limit theorem.

Keywords: Multiscale, Markov chains, Chemical reaction, Reaction networks, Heat shock

Background

Stochasticity may play an important role in biochemical systems. For example, stochasticity may be beneficial to give variability in gene expression, to produce population heterogeneity, and to adjust or respond to fluctuations in environment [1]. We are interested in local dynamics of biochemical networks involving some species with a small number of molecules so that the system is assumed to be well-mixed and relative fluctuations of small species numbers may play a role in the system dynamics.

The conventional stochastic model for the well-stirred biochemical network is based on the chemical master equation. The chemical master equation governs the evolution of the probability density of species numbers and is expressed as the balanced equation between influx and outflux of the probability density. When the biochemical network involves many species or bimolecular reactions, it is rarely possible to obtain an exact solution of the master equation in a closed form. Instead of searching for the solution of the master equation, stochastic simulation algorithms are used to obtain the temporal evolution of the species numbers. For example, Gillespie's Stochastic Simulation Algorithm (SSA, or the direct method) is well known [2,3] and provides a realization of the exact trajectory of the sample path for the species numbers. As the biochemical network has more species and reactions, SSA becomes computationally expensive and more efficient algorithms were suggested by many authors [4-6]. The detailed review of stochastic simulation methods, stochastic approximations, and hybrid simulation methods is given in [7]. For models with well-separated time scales, numerous authors suggested stochastic simulation algorithms for biochemical reaction networks by assuming that "fast" subnetworks have reached a "partial equilibrium" [6] or a "quasi-steady state" [4]. Using

Correspondence: kang.235@mbi.osu.edu
Mathematical Biosciences Institute, Ohio State University, Columbus, OH, USA

these assumptions, the approximate stochastic simulation algorithms involve a reduced number of species or reactions.

On the other hand, Ball et al. [8] described the state of the biochemical reaction network in the well-stirred system directly using stochastic equations for species numbers, and suggested an approximation of the reaction network via limiting models derived using different scalings for the species numbers and for the reaction rate constants. Kang and Kurtz [9] extended this multiscale approximation method and gave a systematic way to obtain limiting models in the time scales of interest. Conditions are given to help identify appropriate values for a set of scaling exponents which determine the time scale of each species and reaction. Using this method, nonstationary behavior of biochemical systems can be analyzed. Moreover, application of the method is flexible in the sense that the method does not require the exact parameter values but gives approximations valid for a range of parameter values. More recently, Crude et al. [10] also proposed a reduction method to derive simplified models with preserving stochastic properties and with key parameters using averaging and hybrid simplification.

The multiscale approximation method in [9] requires consideration of magnitude of both species numbers and rate constants of the reactions involving the corresponding species. When a moderately fast reaction involves two species, one with a small number of molecules and the other with a large number of molecules, the effects of this reaction on these species are different. Net molecule changes of species with large numbers due to the reaction is less noticeable than those of species with small numbers. Therefore, though the same reaction governs these species, their time scales may be different from each other. Letting N_0 be a fixed constant and choosing a large value for N_0, for example $N_0 = 100$, we express magnitudes of species numbers and reaction rate constants in terms of powers of N_0 with different scaling exponents. For instance, 1 to 10 molecules are expressed as $1 \times N_0^0$ to $10 \times N_0^0$ molecules, 500 to 800 molecules are rewritten as $5 \times N_0$ to $8 \times N_0$ molecules, and 0.0002 sec becomes $2 \times N_0^{-2}$ sec. Assuming N_0 is large, we replace N_0 by a large parameter N and stochastic equations for species numbers are expressed in terms of N. Then, N is an analogue of $1/\epsilon$ where ϵ is a small parameter in perturbation theory.

A specific time scale of interest is expressed in terms of a power of N, and its exponent contributes to reaction rates due to change of variables in time. For each species (or linear combination of species), we compare a power of N for the species number and those for reaction rates involving this species. Consider a case when the power for the species number is larger than those for the rates of all reactions where the species is involved. Then net molecule changes due to the reactions are not

large enough to be noticeable in this time scale, and the species number is approximated as constant. Next, consider a case when the power for the species number is smaller than those for some reaction rates involving the species. In this case, the species number fluctuates very rapidly due to the fast reactions in this time scale, and the averaged behavior of the species number can be described in terms of other species numbers. The method of averaging is similar to approximation of one variable in terms of others using a quasi-steady state assumption. Last, when the power for the species number is equal to those for the rates of reactions where the species is involved, the scaled species number is approximated by a nondegenerate limit describing nonstationary behavior of the species number in the specific time scale of interest. The limit could be described in various kinds of variables: a continuous time Markov chain, a deterministic model given by a system of ordinary differential equations, or a hybrid model with both discrete and continuous variables. Since some of the scaled species numbers are approximated as constants or the averaged behavior of some species numbers is expressed in terms of other variables, dimension of species in the approximation of the biochemical network is reduced.

In the multiscale approximation method, scaling exponents for species numbers and for reaction rate constants are not uniquely determined, since the choice of values for the exponents is flexible. For example, 0.005 sec can be expressed as $0.5 \times N_0^{-1}$ or $5 \times N_0^{-1.5}$ when $N_0 = 100$. The goal in this method is to find an appropriate set of scaling exponents to obtain a nondegenerate limit of the scaled species numbers. Orders of magnitude of species numbers in the propensities affect reaction rates, and reaction rates contribute to determining rates of net molecule changes of the species involved in the reactions. Since species numbers and reaction rates interact, it is not easy to determine scaling exponents for all species numbers and reaction rate constants so that the limits of the scaled species numbers become balanced.

Kang and Kurtz [9] introduced *balance conditions* for the scaling exponents, which help to determine values for a set of exponents. The key idea in these conditions is that for each species (or linear combination of species) the maximum of scaling exponents in the rates of the reactions where this species is produced should be the same as that in the rates of the reactions where this species is consumed, i.e. maximal production and consumption rates of the species should be balanced in the order of magnitude. In case the maximums of scaling exponents for productions and consumptions are not balanced for some species, an increase or decrease of the scaled species number can be described by its limit during a certain time period. However after this time period, the scaled species number will either become zero

or blow up to infinity. Therefore, if some of the scaled species numbers are not balanced due to a difference between orders of magnitude of production and consumption rates, the chosen scaling is valid up to a certain time scale. After this time scale, we need to choose different values for scaling exponents. In each time scale of interest we derive a limiting model including a subset of species and reactions, which is used to approximate the state of the full reaction network. The multiscale approximation method is applicable in case some of reaction rates are not known accurately, since the chosen scaling is applicable in some ranges of the parameters. Therefore, based on the behavior of the limiting models, we may be able to estimate behavior for a range of parameter values without performing a huge number of stochastic simulations.

The paper [9] included several simple examples of biochemical networks involving two to four species, and derived limiting models in each time scale of interest. To apply this method, more scaling exponents must be determined as the biochemical network involves more species or reactions. Therefore, it is challenging to apply the method to complex biochemical systems and to determine appropriate values for scaling exponents so that the corresponding limiting models preserve important dynamical features of the full system. One of the goals of this paper is to illustrate how to apply this method to an example with significant complexity. In this paper, using a significantly complicated biochemical network, we derive limiting models, show convergence of the scaled species numbers to their limit, and estimate the error analytically between the scaled species numbers and their limit. We analyze a heat shock response model of *Escherichia coli* (*E. coli*) developed by Srivastava, Peterson, and Bentley in [11]. The model involves 9 species and 18 reactions with significant complexity as shown in Figure 1, and it has various time scales due to wide ranges of species numbers and reaction rate constants. Because of various scales involved, this model has been used as an example to show accuracy of the stochastic simulation algorithms which are developed to increase computational efficiency using the multiscale nature of the chemical reaction network [12,13]. Another version of a heat shock response model of *E. coli* is studied in [6] using an accelerated SSA that also exploits the multiscale nature of the system.

Applying the multiscale approximation method to the heat shock response model of *E. coli*, we derive limiting models in three time scales of our interests, which approximate the full network given in Figure 1. Denote ∅ as species we are not interested in. Let S_i represent the *i*th species and S_{23} be addition of species S_2 and S_3. $A \to B$ denotes a reaction where one molecule of species A is converted to one molecule of species B. In the early stage
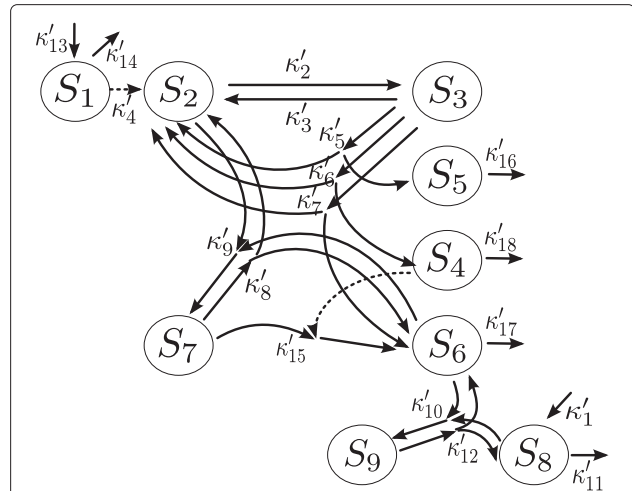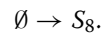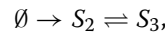

Figure 1 A chemical reaction network in the heat shock response model of E. coli. A dotted line represents the effect of the species acting as catalysts. κ_k'''s represent stochastic reaction rate constants.

of time period of order 1 sec, we obtain the following reduced network:

$$\emptyset \to S_2 \rightleftharpoons S_3,$$

$$\emptyset \to S_8.$$

The reduced network in the early stage has very simple structure without any bimolecular reactions, and all reactions involved are either production from a source or conversion. Moreover, the reduced network is well separated into two due to independence of S_8 from S_2 and S_3.
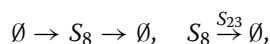
In the medium stage of time period of order 100 sec, the full network is reduced to

$$\emptyset \to S_{23},$$

$$\emptyset \to S_6 \xrightarrow{S_8} \emptyset, \quad \emptyset \xrightarrow{S_{23}} S_6,$$

$$S_7 \to S_6,$$

$$\emptyset \to S_8,$$
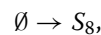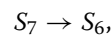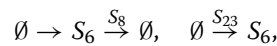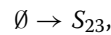
where a species over the arrow accelerates or inhibits the corresponding reaction. The reaction does not change this species number, but the propensity of the corresponding reaction is a function of this species number. In this time scale, conversion between S_2 and S_3 occurs very frequently and S_2 and S_3 play a role as a single "virtual" species rather than separate species. The species numbers of S_{23} and S_8 are described as two independent birth processes and the species number of S_7 is governed by conversion. In this time scale, the species number of S_8 is normalized and treated as a continuous variable. The interesting thing is that the

behavior of the species S_8 which rapidly increases in time is well approximated in both first and second time scales.

In the late stage of time period of order 10,000 sec, we get a reduced network with more species involved than those in the previous time scales. However, the reduced network is still much simpler than the full network in Figure 1. At this time scale, we get

$$\emptyset \to S_1 \to \emptyset,$$

$$\emptyset \xrightarrow{S_1} S_{23} \xrightarrow{S_8,S_9} \emptyset,$$

$$\emptyset \xrightarrow{S_{23}} S_4 \to \emptyset,$$

$$\emptyset \xrightarrow{S_{23}} S_5 \to \emptyset,$$

$$\emptyset \to S_8 \to \emptyset, \quad S_8 \xrightarrow{S_{23}} \emptyset,$$

$$\emptyset \xrightarrow{S_{23}} S_9.$$
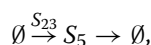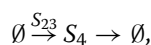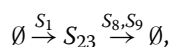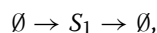
As we see in Figure 1, the full network involves reactions with more than two reactants or products. However, all reactions in the reduced network at the times of order 10,000 sec consist of either production or degradation of each species, though most of the species (6 species out of 9) are involved in the reduced model. As in the medium stage of time period, S_2 and S_3 play a role as a single species. In the early and medium stages of time period propensities are in a form following the law of mass action, while in the late stage of time period the propensity for degradation of S_{23} is a nonlinear function of the species numbers similar to the reaction rate appearing in the Michaelis-Menten approximation for an enzyme reaction. The nonlinear function involves the species numbers of S_{23}, S_8, and S_9, which come from averaging of the species numbers of S_2 and S_6 which fluctuate rapidly in the third time scale. Similarly, the propensity of catalytic degradation of S_8 is not proportional to the number of molecules of S_8.

In the late stage of time period of order 10,000 sec, we study the error between the scaled species numbers and their limit analytically using the central limit theorem derived in [14] and show that the error is of order 10^{-1}.

Methods

In the next several sections, we apply the multiscale approximation to the heat shock response model of *E. coli* and derive the limiting models. The multiscale approximation method is described in terms of the following steps so that the method can be applied to the general cases.

1. Write a chemical reaction network involving s_0 species and r_0 reactions in the form of

$$\sum_{i=1}^{s_0} v_{ik} S_i \longrightarrow \sum_{i=1}^{s_0} v'_{ik} S_i, \qquad k = 1, \cdots, r_0,$$
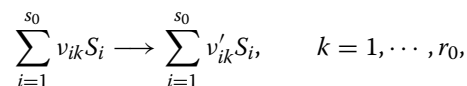
where v_{ik} and v'_{ik} are nonnegative integers. Rearrange the reactions so that the reaction rate constants are decreasing monotonically as k gets large.

2. Derive a system of stochastic equations for species numbers.

   (a) Letting $X_i(t)$ be the number of molecules of species S_i at time t, the corresponding stochastic equation is

   $$X_i(t) = X_i(0) + \sum_{k=1}^{r_0} R_k^t(\lambda_k(X))(v'_{ik} - v_{ik}), \quad i = 1, \cdots s_0,$$

   where $R_k^t(\cdot)$ counts the number of times that the kth reaction occurs up to time t.

   (b) $\lambda_k(x)$ is determined by a stochastic version of mass action kinetics, and is expressed as a product of the rate constant and the numbers of molecules of reactants. If the kth reaction is second-order ($\sum_{i=1}^{s_0} v_{ik} = 2$) with different types of reactants, $\lambda_k(x) = \kappa'_k x_p x_q$. When the reactants are two molecules of the same species, $\lambda_k(x) = \kappa'_k x_p(x_p - 1)$.

3. Derive a system of stochastic equations for the normalized species numbers after a time change, $Z^{N,\gamma}(t)$.

   (a) In the equation for $X_i(t)$ obtained in Step 2 (a), replace X_i by $Z_i^{N,\gamma}$ and divide reaction terms by N^{α_i}. In the kth reaction term, put $N^{\gamma + \rho_k}$ in the propensity and replace $\lambda_k(X)$ by $\hat{\lambda}_k(Z^{N,\gamma})$. Then, we have

   $$Z_i^{N,\gamma}(t) = Z_i^{N,\gamma}(0) + N^{-\alpha_i} \sum_{k=1}^{r_0} R_k^t(N^{\gamma+\rho_k}\hat{\lambda}_k(Z^{N,\gamma}))$$
   $$\times (v'_{ik} - v_{ik}), \quad i = 1, \cdots s_0.$$

   (b) In the equation in Step 3 (a), $\rho_k = \beta_k + \sum_{j=1}^{s_0} \alpha_j v_{jk}$.

   (c) In the most reactions, $\hat{\lambda}_k$ is obtained by replacing κ'_k by κ_k in λ_k. In case the kth reaction is second-order with reactants of the same species, $\lambda_k(x) = \kappa'_k x_p(x_p - 1)$ is replaced by $\hat{\lambda}_k(z) = \kappa_k z_p(z_p - N^{-\alpha_p})$.

4. Write a set of species balance equations and their time-scale constraints.

(a) Define Γ_i^+ and Γ_i^- as subsets of reactions where the species number of S_i increases or decreases every time the reaction occurs. Comparing ρ_k's for $k \in \Gamma_i^+$ and those for $k \in \Gamma_i^-$, set the balance equations as

$$\max_{k \in \Gamma_i^+} \rho_k = \max_{k \in \Gamma_i^-} \rho_k, \qquad i = 1, \cdots, s_0.$$

(b) Time-scale constraints are given as

$$\gamma \le \max_{k \in \Gamma_i^+ \cup \Gamma_i^-} \rho_k, \qquad i = 1, \cdots, s_0.$$

5. Find a minimum set of linear combinations of species whose maximum of collective production (or consumption) rates may be different from that of one of any species. We construct a minimum set of linear combinations of species by selecting a linear combination of species if any reaction term involving the species consisting of the linear combination is canceled in the equation for the linear combination of species.

6. For each selected linear combination of species, write a collective species balance equation and its time-scale constraint. They are obtained similarly to the ones in Step 4 using subsets of reactions where the number of molecules of linear combinations of species either increases or decreases instead of using Γ_i^+ and Γ_i^-.

7. Select a large value for N_0 and choose an appropriate set of α_i's and β_k's so that

   (a) the species number X_i and the reaction rate constant κ_k' are approximately of orders $N_0^{\alpha_i}$ and $N_0^{\beta_k}$;
   (b) the normalized species number $Z_i^{N,\gamma}$ and the scaled reaction rate constant κ_k are of order 1;
   (c) most of the balance equations obtained in Steps 4 and 6 are satisfied;
   (d) β_k's are monotone decreasing among each class of reactions which have the same number of molecules of reactants.

8. Plugging the chosen values for α_i's and β_k's in the time-scale constraints obtained in Steps 4 and 6, compute an upper bound (denoted as γ_0) for a time-scale exponent. Then, the chosen set of exponents α_i's and β_k's can be used for γ satisfying $\gamma \le \gamma_0$. For $\gamma > \gamma_0$, select another set of exponents α_i's and β_k's using Steps 7 and 8.

9. Using each set of values for α_i's and β_k's, identify a natural time scale exponent of each species (denoted as γ_i for species S_i) so that γ_i satisfies

$$\max_{k \in \Gamma_i^+ \cup \Gamma_i^-} (\gamma_i + \rho_k) = \alpha_i, \qquad i = 1, \cdots, s_0.$$

We collect γ_i's with the same values, whose species are in the same time scales in the approximation.

10. Modify α_i's and β_k's so that the conditions in Step 7 are satisfied and that γ_i's are divided into appropriate number of values, which gives the number of time scales, $N^\gamma = N^{\gamma_i}$, we are interested in.

11. For each chosen γ, derive a limiting equation for each species S_i with $\gamma_i = \gamma$. Using the stochastic equation obtained in Step 3 (a), we let N go to infinity.

    (a) For $k \in \Gamma_i^+ \cup \Gamma_i^-$, the kth reaction term converges to zero if $\alpha_i > \gamma + \rho_k$.
    (b) If $\alpha_i = \gamma + \rho_k$, the kth reaction term appears as a limit in the limiting equation. The limit of the kth reaction term is discrete if $\alpha_i = 0$, while it is a continuous variable with the limit of its propensity if $\alpha_i > 0$.
    (c) There is no k satisfying $\alpha_i < \gamma + \rho_k$ in the equation for species S_i with $\gamma = \gamma_i$ due to the definition of γ_i given in Step 9.

12. In the limiting equation for each species S_i with $\gamma_i = \gamma$, we approximate propensities in the reaction terms. Suppose that the normalized species number for S_j appears in the propensities.

    (a) If $\gamma_j > \gamma$, the limit of the normalized species number for S_j is its initial value.
    (b) If $\gamma_j = \gamma$, the limit of the normalized species number for S_j appears as a variable in the propensities in the limiting equation.
    (c) If $\gamma_j < \gamma$, the limit of the normalized species number for S_j is expressed as a function of the limits of the normalized species numbers for S_i with $\gamma_i = \gamma$. The function for S_j is obtained by dividing the equation for S_j by $N^{\max_{k \in \Gamma_j^+ \cup \Gamma_j^-} (\gamma + \rho_k) - \alpha_j}$ and letting N go to infinity.

13. If a limiting model is not closed, consider limiting equations for some linear combinations of species selected in Step 5 whose natural time scale exponents are equal to the chosen γ.

The method for multiscale approximation described above can be applied to general chemical reaction networks containing different scales in species numbers and reaction rate constants. We can apply the method in case the rates of chemical reactions are determined by law of mass action and when there is no species whose number is either zero or infinity at all times. As given in [9], in the reaction network involving $\emptyset \rightarrow S_1$, $\emptyset \rightarrow S_2$, $\emptyset \rightarrow S_3$, $S_1 + S_2 \rightarrow \emptyset$, and $S_1 + S_3 \rightarrow \emptyset$, convergence of the limit for the scaled species numbers may not be guaranteed at some time scales. Suppose that production rate of S_1 is

larger than that of S_2 but with the same order of magnitude, and that production rate of S_3 is much smaller than those of S_1 and S_2. Then, $X_1(t)$ may blow up to infinity and $X_2(t)$ may go to zero at some time scales. In this case, the method is not applicable.

Results and discussion

Model description

We analyze a heat shock response model of *E. coli* developed by Srivastava, Peterson, and Bentley [11]. The heat shock response model gives a simplified mechanism occurring in the *E. coli* to respond to high temperature. Heat causes unfolding, misfolding, or aggregation of proteins, and cells overcome the heat stress by producing heat shock proteins, which refold or degrade denatured proteins. In *E. coli*, σ^{32} factors play an important role in recovery from the stress under the high temperature. σ^{32} factors catalyze production of the heat shock proteins such as chaperon proteins and other proteases. In this model, *J* denotes a chaperon complex, *FtsH* represents a σ^{32}-regulated stress protein, and *GroEL* is a σ^{32}-mediated stress response protein.

σ^{32} factors are in three different forms, free σ^{32} protein, σ^{32} combined with RNA polymerase ($E\sigma^{32}$), and σ^{32} combined with a chaperon complex (σ^{32}-*J*). Under the normal situation without stress, most of the σ^{32} factors combine with chaperon complexes and form σ^{32}-*J*. A chaperon complex *J* keeps σ^{32} factors in an inactive form, and σ^{32} factors can directly respond to the stress by changing into different forms. When there exist σ^{32} factors combined with chaperon complexes, *FtsH* catalyzes degradation of σ^{32} factors. Thus, if enough σ^{32}-regulated stress proteins are produced, σ^{32} factors are degraded.

Not only σ^{32} factors, but recombinant proteins also require chaperon complexes to form a complex so that denatured protein can be fixed. Therefore, σ^{32} factors and recombinant proteins compete to bind chaperon complexes, and different levels of binding affinity of recombinant proteins to chaperon complexes change the evolution of the system state. In the model, we assume that σ^{32} factors and recombinant proteins have the same affinity to bind to chaperon complexes. The system is sensitive to the amount and forms of σ^{32} factors: a small decrease of σ^{32} factors causes a large reduction of production of chaperon complexes and σ^{32}-regulated stress proteins, and the ratio of three different forms of σ^{32} factors determines system dynamics in the stress response [11]. The total initial number of molecules of σ^{32} factors in each cell is small [11] (also see initial values for S_2, S_3, and S_7 which are 1, 1, and 7 in Table 1), and the stochastic model is appropriate to be considered.

The model involves 9 species and 18 reactions. Denote s_0 as the number of species and r_0 as the number of reactions. Let $X(t)$ be a state vector whose ith component

Table 1 Species in the heat shock response model of E. coli and their initial values

X_1	=	# of S_1	σ^{32} mRNA	$X_1(0)$	=	10
X_2	=	# of S_2	σ^{32} protein	$X_2(0)$	=	1
X_3	=	# of S_3	$E\sigma^{32}$	$X_3(0)$	=	1
X_4	=	# of S_4	*FtsH*	$X_4(0)$	=	93
X_5	=	# of S_5	*GroEL*	$X_5(0)$	=	172
X_6	=	# of S_6	*J*	$X_6(0)$	=	54
X_7	=	# of S_7	σ^{32}-*J*	$X_7(0)$	=	7
X_8	=	# of S_8	Recombinant protein	$X_8(0)$	=	50
X_9	=	# of S_9	*J*-Recombinant protein	$X_9(0)$	=	0

represents the number of molecules of species S_i at time t for $i = 1, \cdots, s_0$. Define a random process which counts the number of times that the kth reaction occurs by time t as

$$R_k^t(\lambda_k(X)) \equiv Y_k\left(\int_0^t \lambda_k(X(s))\,ds\right), \quad k = 1, \cdots, r_0,$$

where $\lambda_k(X)$ is the propensity of the kth reaction and the Y_k's are independent unit Poisson processes. Therefore, $R_k^t(\cdot)$ is a nonnegative integer-valued random process increasing by 1. As $\lambda_k(\cdot)$ gets large, the moment when $R_k^t(\lambda_k(\cdot))$ increases becomes more frequent. Let ν_{ik} (ν'_{ik}) be the number of molecules of S_i that are consumed (produced) in the kth reaction. Define ν_k (ν'_k) as an s_0-dimensional vector whose ith component is ν_{ik} (ν'_{ik}). Then, $X(t)$ is given as

$$X(t) = X(0) + \sum_{k=1}^{r_0} R_k^t(\lambda_k(X))(\nu'_k - \nu_k). \qquad (1)$$

That is, species numbers at time t are expressed in terms of their initial values and sum of the number of times that each reaction occurs multiplied by net molecule changes in the corresponding reaction. In our model, the system of equations are derived using a set of reactions in Table 2 as:

$$X_1(t) = X_1(0) + R_{13}^t(\kappa'_{13}) - R_{14}^t(\kappa'_{14}X_1),$$

$$\begin{aligned} X_2(t) = X_2(0) &+ R_3^t(\kappa'_3X_3) + R_4^t(\kappa'_4X_1) + R_5^t(\kappa'_5X_3) \\ &+ R_6^t(\kappa'_6X_3) + R_7^t(\kappa'_7X_3) + R_8^t(\kappa'_8X_7) - R_2^t(\kappa'_2X_2) \\ &- R_9^t(\kappa'_9X_2X_6), \end{aligned}$$

$$\begin{aligned} X_3(t) = X_3(0) &+ R_2^t(\kappa'_2X_2) - R_3^t(\kappa'_3X_3) - R_5^t(\kappa'_5X_3) \\ &- R_6^t(\kappa'_6X_3) - R_7^t(\kappa'_7X_3), \end{aligned}$$

$$X_4(t) = X_4(0) + R_6^t(\kappa'_6X_3) - R_{18}^t(\kappa'_{18}X_4),$$

$$X_5(t) = X_5(0) + R_5^t(\kappa'_5X_3) - R_{16}^t(\kappa'_{16}X_5),$$

$$\begin{aligned} X_6(t) = X_6(0) &+ R_7^t(\kappa'_7X_3) + R_8^t(\kappa'_8X_7) + R_{12}^t(\kappa'_{12}X_9) \\ &+ R_{15}^t(\kappa'_{15}X_4X_7) - R_9^t(\kappa'_9X_2X_6) - R_{10}^t(\kappa'_{10}X_6X_8) \\ &- R_{17}^t(\kappa'_{17}X_6), \end{aligned}$$

Table 2 Reactions in the heat shock response model of E. coli

	Reaction	Transition
R1	$\varnothing \xrightarrow{\text{gene}} S_8$	Recombinant protein synthesis
R2	$S_2 \longrightarrow S_3$	Holoenzyme association
R3	$S_3 \longrightarrow S_2$	Holoenzyme disassociation
R4	$\varnothing \xrightarrow{S_1} S_2$	σ^{32} translation
R5	$S_3 \xrightarrow{\text{gene}} S_2 + S_5$	*GroEL* synthesis
R6	$S_3 \xrightarrow{\text{gene}} S_2 + S_4$	*FtsH* synthesis
R7	$S_3 \xrightarrow{\text{gene}} S_2 + S_6$	*J*-production
R8	$S_7 \longrightarrow S_2 + S_6$	σ^{32}-*J*-disassociation
R9	$S_2 + S_6 \longrightarrow S_7$	σ^{32}-*J*-association
R10	$S_6 + S_8 \longrightarrow S_9$	Recombinant protein-*J* association
R11	$S_8 \longrightarrow \varnothing$	Recombinant protein degradation
R12	$S_9 \longrightarrow S_6 + S_8$	Recombinant protein-*J* disassociation
R13	$\varnothing \xrightarrow{\text{gene}} S_1$	σ^{32} transcription
R14	$S_1 \longrightarrow \varnothing$	σ^{32} mRNA decay
R15	$S_7 \xrightarrow{S_4} S_6$	σ^{32} degradation
R16	$S_5 \longrightarrow \varnothing$	*GroEL* degradation
R17	$S_6 \longrightarrow \varnothing$	*J*-disassociation
R18	$S_4 \longrightarrow \varnothing$	*FtsH* degradation

In Reaction 5, 6, and 7, we assume that the number of molecules of each gene is 1 and that these reactions are effectively unimolecular. Similarly, Reactions 1 and 13 are treated as production from a source.

$$
\begin{aligned}
X_7(t) &= X_7(0) + R_9^t(\kappa_9' X_2 X_6) - R_8^t(\kappa_8' X_7) - R_{15}^t(\kappa_{15}' X_4 X_7), \\
X_8(t) &= X_8(0) + R_1^t(\kappa_1') + R_{12}^t(\kappa_{12}' X_9) - R_{10}^t(\kappa_{10}' X_6 X_8) \\
&\quad - R_{11}^t(\kappa_{11}' X_8), \\
X_9(t) &= X_9(0) + R_{10}^t(\kappa_{10}' X_6 X_8) - R_{12}^t(\kappa_{12}' X_9).
\end{aligned}
\tag{2}
$$

κ_k' represents the stochastic reaction rate constant for the kth reaction, and their values from [11] are given in Table 3.

We derive the limiting models in three time scales, which approximate a full network in a certain time period involving a subset of species and reactions. In what follows, Z_i^{γ} is a limit of the scaled species number of S_i at some time scales depending on γ, and as γ gets larger the times are in the later stage. Note that the exponent γ in Z_i^{γ} does not imply $(Z_i)^{\gamma}$ but it shows dependence of Z_i^{γ} on γ. Let κ_k be a scaled reaction rate constant for the kth reaction. In the first time scale (when the times are in the early stage), the subnetwork governed by

$$
\begin{aligned}
Z_2^0(t) &= Z_2^0(0) + R_3^t(\kappa_3 Z_3^0) + R_4^t(\kappa_4 Z_1^0(0)) - R_2^t(\kappa_2 Z_2^0), \\
Z_3^0(t) &= Z_3^0(0) + R_2^t(\kappa_2 Z_2^0) - R_3^t(\kappa_3 Z_3^0), \\
Z_8^0(t) &= Z_8^0(0) + R_1^t(\kappa_1),
\end{aligned}
\tag{3}
$$

Table 3 Stochastic reaction rate constants in the heat shock response model of E. coli

Rates		Rates	
κ_1'	4.00×10^{0}	κ_{10}'	3.62×10^{-4}
κ_2'	7.00×10^{-1}	κ_{11}'	9.99×10^{-5}
κ_3'	1.30×10^{-1}	κ_{12}'	4.40×10^{-5}
κ_4'	7.00×10^{-3}	κ_{13}'	1.40×10^{-5}
κ_5'	6.30×10^{-3}	κ_{14}'	1.40×10^{-6}
κ_6'	4.88×10^{-3}	κ_{15}'	1.42×10^{-6}
κ_7'	4.88×10^{-3}	κ_{16}'	1.80×10^{-8}
κ_8'	4.40×10^{-4}	κ_{17}'	6.40×10^{-10}
κ_9'	3.62×10^{-4}	κ_{18}'	7.40×10^{-11}

We convert deterministic rate constants in [11] using the volume of *E. coli* which is assumed to be $1.5 \times 10^{-15} L$.

approximates the network when the times are of order 1 sec. Denote Z_{23}^1 as the limit of the addition of the scaled species numbers for S_2 and S_3. In the second time scale (when the times are in the medium stage), the subnetwork governed by

$$
\begin{aligned}
Z_{23}^1(t) &= Z_{23}^1(0) + R_4^t(\kappa_4 Z_1^1(0)), \\
Z_6^1(t) &= Z_6^1(0) + R_7^t\left(\frac{\kappa_2 \kappa_7}{\kappa_2 + \kappa_3} Z_{23}^1\right) + R_{12}^t(\kappa_{12} Z_9^1(0)) \\
&\quad + R_{15}^t(\kappa_{15} Z_4^1(0) Z_7^1) - R_{10}^t(\kappa_{10} Z_6^1 Z_8^1), \\
Z_7^1(t) &= Z_7^1(0) - R_{15}^t(\kappa_{15} Z_4^1(0) Z_7^1), \\
Z_8^1(t) &= Z_8^1(0) + \kappa_1 t,
\end{aligned}
\tag{4}
$$

approximates the network at the times of order 100 sec. In the third time scale, set the limit of the averaged scaled species numbers of fast-fluctuating species S_2, S_3, and S_6 as

$$
\begin{aligned}
\bar{Z}_2^2(t) &\equiv \frac{\kappa_3}{\kappa_2 + \kappa_3} Z_{23}^2(t), \\
\bar{Z}_3^2(t) &\equiv \frac{\kappa_2}{\kappa_2 + \kappa_3} Z_{23}^2(t), \\
\bar{Z}_6^2(t) &\equiv \frac{\kappa_7 \bar{Z}_3^2(s) + \kappa_{12} Z_9^2(s)}{\kappa_{10} Z_8^2(s)}.
\end{aligned}
$$

When the times are in a late stage, the subnetwork governed by

$$
\begin{aligned}
Z_1^2(t) &= Z_1^2(0) + R_{13}^t(\kappa_{13}) - R_{14}^t(\kappa_{14} Z_1^2), \\
Z_{23}^2(t) &= Z_{23}^2(0) + \int_0^t \left[\kappa_4 Z_1^2(s) - \frac{\kappa_3 \kappa_9}{\kappa_2 + \kappa_3} Z_{23}^2(s) \right. \\
&\quad \left. \times \left(\frac{\frac{\kappa_2 \kappa_7}{\kappa_2 + \kappa_3} Z_{23}^2(s) + \kappa_{12} Z_9^2(s)}{\kappa_{10} Z_8^2(s)} \right) \right] ds \\
&\equiv Z_{23}^2(0) + \int_0^t \left[\kappa_4 Z_1^2(s) - \kappa_9 \bar{Z}_2^2(s) \bar{Z}_6^2(s) \right] ds,
\end{aligned}
$$

$$Z_4^2(t) = Z_4^2(0) + \int_0^t \left(\frac{\kappa_2 \kappa_6}{\kappa_2 + \kappa_3} Z_{23}^2(s) - \kappa_{18} Z_4^2(s) \right) ds$$

$$\equiv Z_4^2(0) + \int_0^t \left(\kappa_6 \bar{Z}_3^2(s) - \kappa_{18} Z_4^2(s) \right) ds,$$

$$Z_5^2(t) = Z_5^2(0) + \int_0^t \left(\frac{\kappa_2 \kappa_5}{\kappa_2 + \kappa_3} Z_{23}^2(s) - \kappa_{16} Z_5^2(s) \right) ds$$

$$\equiv Z_5^2(0) + \int_0^t \left(\kappa_5 \bar{Z}_3^2(s) - \kappa_{16} Z_5^2(s) \right) ds,$$

$$Z_8^2(t) = Z_8^2(0) + \int_0^t \left(\kappa_1 - \frac{\kappa_2 \kappa_7}{\kappa_2 + \kappa_3} Z_{23}^2(s) - \kappa_{11} Z_8^2(s) \right) ds$$

$$\equiv Z_8^2(0) + \int_0^t \left(\kappa_1 - \kappa_7 \bar{Z}_3^2(s) - \kappa_{11} Z_8^2(s) \right) ds,$$

$$Z_9^2(t) = Z_9^2(0) + \int_0^t \frac{\kappa_2 \kappa_7}{\kappa_2 + \kappa_3} Z_{23}^2(s)\, ds$$

$$\equiv Z_9^2(0) + \int_0^t \kappa_7 \bar{Z}_3^2(s)\, ds,$$

$$\tag{5}$$

approximates the network at the times of order 10,000 sec. Detailed derivation is given in the later sections. Note that it is possible to identify different numbers of time scales depending on the scaling of the species numbers and reaction rate constants. In the heat shock response model of *E. coli*, it is possible to obtain approximate models with two or four time scales. However, if the number of time scales are too many, the limiting model in each time scale may involve one species and a few number of reactions and the model in this case may not be interesting to consider.

Derivation of the scaled models

The stochastic equations given in Equations (2) describe temporal evolution of the species numbers. For example, the equations for species S_2 and S_3 are

$$X_2(t) = X_2(0) + R_3^t \left(\kappa_3' X_3 \right) + R_4^t \left(\kappa_4' X_1 \right) + R_5^t \left(\kappa_5' X_3 \right)$$
$$+ R_6^t \left(\kappa_6' X_3 \right) + R_7^t \left(\kappa_7' X_3 \right) + R_8^t \left(\kappa_8' X_7 \right) - R_2^t \left(\kappa_2' X_2 \right)$$
$$- R_9^t \left(\kappa_9' X_2 X_6 \right),$$

$$\tag{6a}$$

$$X_3(t) = X_3(0) + R_2^t \left(\kappa_2' X_2 \right) - R_3^t \left(\kappa_3' X_3 \right) - R_5^t \left(\kappa_5' X_3 \right)$$
$$- R_6^t \left(\kappa_6' X_3 \right) - R_7^t \left(\kappa_7' X_3 \right).$$

$$\tag{6b}$$

In Equation (6), species numbers of S_2 and S_3 are determined by the times when reactions occur and by the number of times that reactions happen. On the other

hand, reaction time and frequency are determined by propensities which are some functions of species numbers. Therefore, reaction rates and species numbers interact one another. Reaction rates vary from $O(10^{-11})$ to $O(1)$ as we see in Table 3, and species numbers in this model are from $O(1)$ to $O(10^4)$ as we see later in the simulation of the full network. We express each species number and rate constant in terms of powers of a common number with different weights on exponents. Define $N_0 = 100$ as a fixed unitless constant used to express the magnitude of the species numbers and the reaction rate constants. Define α_i for $i = 1, \cdots, s_0$ and β_k for $k = 1, \cdots, r_0$ as the scaling exponent for species S_i and for the reaction rate constant κ_k'. We express the reaction rate constants in a form of $N_0^{\beta_k} \kappa_k$ where κ_k is of order 1 and is determined so that $\kappa_k' = N_0^{\beta_k} \kappa_k$. For example, we have $\kappa_6' = 4.88 \times 10^{-3}$ and we can choose $\beta_6 = -1$ so that the reaction rate is expressed as $\kappa_6' = 0.488 \times N_0^{\beta_6}$. Assuming that N_0 is large, we replace N_0 by N and express the process as $X_i^N(t)$ to show dependence of the species numbers on N. Note that $\{X^N(t)\}$ is a family of processes depending on N and $X_i^N(t) = X_i(t)$ when $N = N_0$. Then, the equation for X_3^N is given as

$$X_3^N(t) = X_3^N(0) + R_2^t \left(N^{\beta_2} \kappa_2 X_2^N \right) - R_3^t \left(N^{\beta_3} \kappa_3 X_3^N \right)$$
$$- R_5^t \left(N^{\beta_5} \kappa_5 X_3^N \right) - R_6^t \left(N^{\beta_6} \kappa_6 X_3^N \right) - R_7^t \left(N^{\beta_7} \kappa_7 X_3^N \right),$$

where $X_3^N(0)$ is defined later so that $X_3^N(0) = X_3(0)$ when $N = N_0$. Since the numbers of molecules of species are in different orders of magnitude, we scale the number of molecules of the ith species by N^{α_i} and set a normalized species number as

$$Z_i^N(t) = N^{-\alpha_i} X_i^N(t).$$

The ith species number may have different orders of magnitude at different times so α_i may have different values for different time scales. Now, we set the initial values as

$$X_i^N(0) \equiv \left\lfloor \left(\frac{N}{N_0} \right)^{\alpha_i} X_i(0) \right\rfloor,$$

$$\tag{7}$$

so that $X_i^{N_0}(0) = X_i(0)$ and $\lim_{N \to \infty} Z_i^N(0) = \lim_{N \to \infty} N^{-\alpha_i} X_i^N(0) = N_0^{-\alpha_i} X_i(0)$.

Next, we scale the propensities of reactions by replacing X_i^N by $N^{\alpha_i} Z_i^N$ and replacing κ_k' by $N^{\beta_k} \kappa_k$. For example, consider the 9th reaction term in (6a). Replacing κ_9' by $N^{\beta_9} \kappa_9$, X_2^N by $N^{\alpha_2} Z_2^N$, and X_6^N by $N^{\alpha_6} Z_6^N$, the 9th reaction term becomes

$$R_9^t(\kappa_9' X_2^N X_6^N) = R_9^t(N^{\beta_9 + \alpha_2 + \alpha_6} \kappa_9 Z_2^N Z_6^N).$$

$$\tag{8}$$

For simplicity, set $\rho_9 = \beta_9 + \alpha_2 + \alpha_6$ and define a scaling exponent in the propensity of the kth reaction term as

$$\rho_k \equiv \beta_k + \nu_k \cdot \alpha,$$

where $\alpha = (\alpha_1, \cdots, \alpha_{s_0})^T$ and $\nu_k = (\nu_{1k}, \cdots, \nu_{s_0k})^T$. Here, ν_{ik} gives the number of molecules of species S_i consumed in the kth reaction. Then, (8) is rewritten as

$$R_9^t\left(\kappa_9' X_2^N X_6^N\right) = R_9^t\left(N^{\rho_9}\kappa_9 Z_2^N Z_6^N\right).$$

Dividing (6a) by N^{α_2} and (6b) by N^{α_3} and scaling the propensities, we get

$$
\begin{aligned}
Z_2^N(t) = Z_2^N(0) + N^{-\alpha_2}&\Big[R_3^t\left(N^{\rho_3}\kappa_3 Z_3^N\right) + R_4^t\left(N^{\rho_4}\kappa_4 Z_1^N\right) \\
&+ R_5^t\left(N^{\rho_5}\kappa_5 Z_3^N\right) + R_6^t\left(N^{\rho_6}\kappa_6 Z_3^N\right) \\
&+ R_7^t\left(N^{\rho_7}\kappa_7 Z_3^N\right) + R_8^t\left(N^{\rho_8}\kappa_8 Z_7^N\right) \\
&- R_2^t\left(N^{\rho_2}\kappa_2 Z_2^N\right) - R_9^t\left(N^{\rho_9}\kappa_9 Z_2^N Z_6^N\right)\Big],
\end{aligned}
$$

(9a)

$$
\begin{aligned}
Z_3^N(t) = Z_3^N(0) + N^{-\alpha_3}&\Big[R_2^t\left(N^{\rho_2}\kappa_2 Z_2^N\right) - R_3^t\left(N^{\rho_3}\kappa_3 Z_3^N\right) \\
&- R_5^t\left(N^{\rho_5}\kappa_5 Z_3^N\right) - R_6^t\left(N^{\rho_6}\kappa_6 Z_3^N\right) \\
&- R_7^t\left(N^{\rho_7}\kappa_7 Z_3^N\right)\Big].
\end{aligned}
$$

(9b)

For each reaction, ρ_k is given in terms of α_i and β_k in the Additional file 1: Table S1.

We are interested in dynamics of species numbers $Z_2^N(t)$ and $Z_3^N(t)$ in various stages of time period. In the early stage of time period, normalized species numbers of S_2 and S_3 are very close to their scaled initial values, since these species numbers have not changed yet. In the medium stage of time period, the normalized species numbers of S_2 and S_3 are asymptotically equal to non-constant limits. In the late stage of time period, the normalized species numbers of S_2 and S_3 fluctuate very rapidly and their averaged behavior is captured in terms of some function of other species numbers.

We want to express the time scale of each species in terms of power of N. First, we express order of magnitude of a specific time period of interest as a power of N with a time scale exponent γ. Applying a time change by replacing t by $N^\gamma t$ in $Z_i^N(t)$, we define a variable for the normalized species numbers after a time change as

$$Z_i^{N,\gamma}(t) \equiv N^{-\alpha_i} X_i^N(tN^\gamma) = Z_i^N(tN^\gamma). \tag{10}$$

Then, $Z_i^{N,\gamma}(t)$ gives a normalized species number at the times of order N^γ. A *natural time scale* of S_i is the time when $Z_i^{N,\gamma}(t)$ has a nonzero finite limit which is not constant and of order 1.

Changing a time variable by replacing t by $N^\gamma t$ in (9a) and (9b), the normalized species numbers of S_2 and S_3 after a time change satisfy

$$
\begin{aligned}
Z_2^{N,\gamma}(t) = Z_2^{N,\gamma}(0) + N^{-\alpha_2}&\Big[R_3^t\left(N^{\gamma+\rho_3}\kappa_3 Z_3^{N,\gamma}\right) \\
&+ R_4^t\left(N^{\gamma+\rho_4}\kappa_4 Z_1^{N,\gamma}\right) + R_5^t\left(N^{\gamma+\rho_5}\kappa_5 Z_3^{N,\gamma}\right) \\
&+ R_6^t\left(N^{\gamma+\rho_6}\kappa_6 Z_3^{N,\gamma}\right) + R_7^t\left(N^{\gamma+\rho_7}\kappa_7 Z_3^{N,\gamma}\right) \\
&+ R_8^t\left(N^{\gamma+\rho_8}\kappa_8 Z_7^{N,\gamma}\right) - R_2^t\left(N^{\gamma+\rho_2}\kappa_2 Z_2^{N,\gamma}\right) \\
&- R_9^t\left(N^{\gamma+\rho_9}\kappa_9 Z_2^{N,\gamma} Z_6^{N,\gamma}\right)\Big],
\end{aligned}
$$

(11a)

$$
\begin{aligned}
Z_3^{N,\gamma}(t) = Z_3^{N,\gamma}(0) + N^{-\alpha_3}&\Big[R_2^t\left(N^{\gamma+\rho_2}\kappa_2 Z_2^{N,\gamma}\right) \\
&- R_3^t\left(N^{\gamma+\rho_3}\kappa_3 Z_3^{N,\gamma}\right) - R_5^t\left(N^{\gamma+\rho_5}\kappa_5 Z_3^{N,\gamma}\right) \\
&- R_6^t\left(N^{\gamma+\rho_6}\kappa_6 Z_3^{N,\gamma}\right) - R_7^t\left(N^{\gamma+\rho_7}\kappa_7 Z_3^{N,\gamma}\right)\Big],
\end{aligned}
$$

(11b)

where N^γ in each propensity comes from the change of the time variable. Here, the initial values may depend on γ, since we can choose different values for α_i for each γ due to changes in order of magnitude of species numbers in time. The stochastic equations after scaling and a time change for all species are given in the Additional file 1: Section 1.

Balance conditions

Our goal is to approximate dynamics of the full network in the heat shock response model of *E. coli* in specific times of interest in terms of simplified subnetworks preserving significant biological features. In each time period of interest, we obtain a nondegenerate limiting model which is not equal to zero and does not blow up to infinity. In this section, we introduce balance conditions which help us to choose appropriate values for the scaling exponents α_i's and β_k's so that the limit is nonzero finite. For each time period of interest of order N_0^γ where $N_0 = 100$, we choose values for scaling exponents so that orders of magnitude of the species number for S_i and the kth reaction rate constant are about $N_0^{\alpha_i}$ and $N_0^{\beta_k}$, respectively. That is,

$$Z_i^{N_0,\gamma}(t) = \frac{X_i^{N_0}(tN_0^\gamma)}{N_0^{\alpha_i}} = O(1),$$

$$\kappa_k = \frac{\kappa_k'}{N_0^{\beta_k}} = O(1).$$

It is natural to choose β_k's in monotone decreasing manner in k, since κ_k''s are in monotone decreasing order

as shown in Table 3. In Table 3, the production rates from a source are the rates per second. The unimolecular reaction rates are the rates per molecule per second while the bimolecular reaction rates are the rates per a pair of molecules per second. Since the reaction rates are expressed in different units, we separate rate constants into three classes based on the number of reactants and assume that monotonicity of β_k's holds in each class of reactions. In other words, we choose β_k's so that

$$\beta_1 \geq \beta_{13},$$
$$\beta_2 \geq \beta_3 \geq \beta_4 \geq \beta_5 \geq \beta_6 \geq \beta_7 \geq \beta_8 \geq \beta_{11} \geq \beta_{12} \geq \beta_{14}$$
$$\geq \beta_{16} \geq \beta_{17} \geq \beta_{18}, \quad \text{and}$$
$$\beta_9 \geq \beta_{10} \geq \beta_{15}.$$

Next, in order to make the normalized specie number $Z_i^{N,\gamma}(t)$ balanced, it is required that the rates of production and consumption of S_i should be in the same order of magnitude. If the order of magnitude of production rate is larger than that of consumption, the normalized species number asymptotically goes to infinity. In the opposite case, the normalized species number asymptotically becomes zero. Therefore, for each species S_i, we set the balance equation for α_i's and β_k's so that the maximal exponent in the propensities of the reactions producing S_i is equal to that in the propensities of the reactions consuming S_i. For example, to obtain a balance equation for species S_2, we compare the scaling exponents in propensities of reactions involving S_2 using (11a), and set the maximal exponents of production and consumption of S_2 equal. Similarly, using (11b), we set the maximal exponents in the production rates and the consumption rates of S_3 equal. Then, the balance equations for species S_2 and S_3 are

$$\max(\rho_3, \rho_4, \rho_5, \rho_6, \rho_7, \rho_8) = \max(\rho_2, \rho_9), \tag{12a}$$

$$\rho_2 = \max(\rho_3, \rho_5, \rho_6, \rho_7). \tag{12b}$$

If the maximal orders of magnitudes of production and consumption rates for S_2 are different from each other, the species number of S_2 should be large enough so that a difference between production and consumption of S_i is not noticeable. In other words if α_i's and β_k's do not satisfy (12a), α_2 should be at least as large as the scaling exponents located in all reaction terms in (11a) to prevent the limit becoming zero or blowing up to infinity. Similarly, in case (12b) is not satisfied, α_3 should be at least as large as the scaling exponents located in the reaction terms in (11b) to prevent the limit becoming zero or blowing up to infinity.

$$\alpha_2 \geq \gamma + \max(\rho_2, \rho_3, \rho_4, \rho_5, \rho_6, \rho_7, \rho_8, \rho_9),$$
$$\alpha_3 \geq \gamma + \max(\rho_2, \rho_3, \rho_5, \rho_6, \rho_7), \tag{13}$$

Solving (13) for γ, we get the following time-scale constraints:

$$\gamma \leq \alpha_2 - \max(\rho_2, \rho_3, \rho_4, \rho_5, \rho_6, \rho_7, \rho_8, \rho_9) \equiv u_2, \tag{14a}$$

$$\gamma \leq \alpha_3 - \max(\rho_2, \rho_3, \rho_5, \rho_6, \rho_7) \equiv u_3. \tag{14b}$$

Inequalities in (14) mean that if maximal production and consumption rates are not balanced either for S_2 or S_3, the chosen set of values for scaling exponents can be used to approximate the dynamics of the full network up to times of order N^{u_2} or N^{u_3}. For times later than those of order N^{u_2} or N^{u_3}, we need to choose another set of values for scaling exponents based on the balance equations. We call the balance equation and the time-scale constraint for each species as the *species balance condition*. If either (12a) or (14a) is satisfied, we say that the species balance condition for S_2 is satisfied.

Even though species balance conditions for S_2 and S_3 are satisfied, the limit of the normalized species numbers for S_2 or S_3 may become degenerate. Consider addition of species S_2 and S_3 as a single virtual species, and compare the collective rates of production and consumption of this species. Recall that S_{23} denotes addition of species S_2 and S_3. Since production of one species is canceled by consumption of the other species, maximal production rate of S_{23} may be different from that of S_2 or S_3. Suppose that the maximal collective rates of production or consumption of S_{23} are slower than the maximal production or consumption rates of S_2 and S_3. Also, suppose that the maximal collective rates of production and consumption of the complex have different orders of magnitude. Then, a limit of the normalized species number of S_{23} can be zero or infinity, even though the species balance conditions for S_2 and S_3 are satisfied. Therefore, we need an additional condition to obtain balance between collective production and consumption rates for S_{23}. To obtain a balance equation for S_{23}, we unnormalize (11a) and (11b) by multiplying N^{α_2} and N^{α_3}, respectively. Adding the unnormalized equations for species S_2 and S_3 and dividing it by $N^{\max(\alpha_2, \alpha_3)}$, we get

$$N^{-\max(\alpha_2,\alpha_3)} \left(N^{\alpha_2} Z_2^{N,\gamma}(t) + N^{\alpha_3} Z_3^{N,\gamma}(t) \right)$$
$$= N^{-\max(\alpha_2,\alpha_3)} \left(N^{\alpha_2} Z_2^{N,\gamma}(0) + N^{\alpha_3} Z_3^{N,\gamma}(0) \right)$$
$$+ N^{-\max(\alpha_2,\alpha_3)} R_4^t (N^{\gamma+\rho_4} \kappa_4 Z_1^{N,\gamma})$$
$$+ N^{-\max(\alpha_2,\alpha_3)} R_8^t (N^{\gamma+\rho_8} \kappa_8 Z_7^{N,\gamma})$$
$$- N^{-\max(\alpha_2,\alpha_3)} R_9^t (N^{\gamma+\rho_9} \kappa_9 Z_2^{N,\gamma} Z_6^{N,\gamma}). \tag{15}$$

Comparing the maximal scaling exponents of production and consumption of S_{23} in (15), a balance equation for S_{23} is given as

$$\max(\rho_4, \rho_8) = \rho_9. \tag{16}$$

In case (16) is not satisfied, the order of magnitude of the species number for S_{23} should be larger than those of collective production and consumption rates so that a difference between production and consumption is not noticeable. This gives

$$\max(\alpha_2, \alpha_3) \geq \gamma + \max(\rho_4, \rho_8, \rho_9). \tag{17}$$

Solving (17) for γ, we get

$$\gamma \leq \max(\alpha_2, \alpha_3) - \max(\rho_4, \rho_8, \rho_9) \equiv u_{23}. \tag{18}$$

Similarly to the time-scale constraint in the species balance condition, (18) implies that if maximal collective production and consumption rates for S_{23} are not balanced, our choice of values for scaling exponents are valid up to times of order $N^{u_{23}}$.

We call (16) and (18) the *collective species balance condition* for S_{23}, that is, either (16) or (18) must hold. The species balance conditions for all species and the collective species balance conditions for all positive linear combinations of species should be satisfied to obtain a nondegenerate limit of $Z_i^{N,\gamma}$ (Condition 3.2 in [9]). Condition 3.2 can be reduced by Lemma 3.4-3.8 and Remark 3.9 in [9]. A key idea is to find a minimum subset of linear combinations of species so that production of one species is canceled by consumption of the other species when we combine the species. In that case, maximal collective production rate of the linear combination of the species may be different from that of each species. Therefore, species balance conditions may not imply the collective species balance condition for the linear combination of the species. For example, a collective species balance condition for addition of S_2 and S_3 should be satisfied, since reactions producing S_2 or S_3 may not increase the species number of S_{23}. In Table 4, we choose linear combinations of species whose collective species balance conditions may not be satisfied by the species balance conditions. For other linear combinations of species, their collective species balance conditions are derived from the ones in Table 4. Satisfying all balance conditions in Table 4 guarantees satisfying balance conditions for all positive linear combination of species, and these conditions help to identify scaling exponents which give a nondegenerate limit of the normalized species numbers in the heat shock response model of *E. coli*. In most cases satisfying balance conditions gives nondegenerate limiting models in the times of interest, but we can still find counter examples as given in the last paragraph in the section for methods.

Based on species and collective species balance equations in Table 4, we choose appropriate values for α_i's and β_k's so that most of the balance equations are satisfied. If some of the balance equations are not satisfied, corresponding time-scale constraints give a range of γ where the chosen α_i's and β_k's are valid. The time-scale constraint, $\gamma \leq \gamma_0$, implies that the set of scaling exponents

α_i's and β_k's chosen is appropriate only up to time whose order of magnitude is equal to N^{γ_0}. For the times larger than $O(N^{\gamma_0})$, we need to choose a different set of values for the scaling exponents, α_i's. Assuming that reaction rate constants do not change in time and that the species numbers vary in time, we in general use one set of β_k's for all time scales and may use several sets of α_i's. A large change of the species numbers in time requires different α_i's in different time scales. For the heat shock model we identify three different time scales as we will see in the section of limiting models in three time scales, and α_1, α_2, α_3, α_8, and α_9 may depend on the time scale. α_4, α_5, α_6, and α_7 are the same for all time scales.

Before we determine scaling exponents for S_1, S_2, and S_3, we run one realization of stochastic simulation to find ranges of the species numbers in time. Using initial values for species S_1, S_2, and S_3, $X_1(0) = 10$ and $X_2(0) = X_3(0) = 1$ as given in Table 1 and using $N_0 = 100$, we set $X_1(t) \approx O(100) = O(N_0^{\alpha_1})$, $X_2(t) \approx O(1) = O(N_0^{\alpha_2})$, and $X_3(t) \approx O(1) = O(N_0^{\alpha_3})$ with $\alpha_1 = 1$ and $\alpha_2 = \alpha_3 = 0$ in the early stage of time period. Plugging in α_i's and β_k's in the balance equations for S_2, S_3, and S_{23}, equality holds in (12a) and (12b) but not in (16). Therefore, (18) gives

$$\gamma \leq \max(\alpha_2, \alpha_3) - \max(\rho_4, \rho_8, \rho_9)$$
$$= \max(0, 0) - \max(0, -2, -2) = 0.$$

Then, the first set of scaling exponents with $\alpha_1 = 1$ and $\alpha_2 = \alpha_3 = 0$ is valid only when $\gamma \leq 0$. Next, based on the fact that $X_2(t) \approx O(10)$ and $X_3(t) \approx O(10)$ in the medium stage of time period, we choose $\alpha_2 = \alpha_3 = 0$ for $\gamma > 0$. At this stage of time period, we set $X_1(t) = O(10) \approx O(N_0^{\alpha_1})$ with $\alpha_1 = 0$. Then, (12a) and (12b) are satisfied but not (16). The condition (18) gives $\gamma \leq 1$, and the second set of scaling exponents with $\alpha_1 = \alpha_2 = \alpha_3 = 0$ is valid when $\gamma \leq 1$. Finally, we set $\alpha_1 = 0$ and $\alpha_2 = \alpha_3 = 1$ for $\gamma > 1$ based on the fact that the numbers of molecules of S_2 and S_3 grow in time and are of order 100. Then, (12a), (12b), and (16) are all satisfied, and the third set of scaling exponents with $\alpha_1 = 0$ and $\alpha_2 = \alpha_3 = 1$ can be used for $\gamma > 1$.

The three sets of values for the scaling exponents chosen are given in the Additional file 1: Table S4. With chosen values for the scaling exponents, we check whether each balance equation is satisfied and give a time-scale constraint in the Additional file 1: Table S6 in case the balance equation is not satisfied. Different choices of α_i's and β_k's from the ones in the Additional file 1: Table S4 give different limiting models. As long as the chosen values for α_i's and β_k's satisfy balance conditions, the limiting model will describe nontrivial behavior of the species numbers which are nonzero and finite in the specific time of interest.

Table 4 Balance equations and time-scale constraints for each species and for each collective species chosen

	Balance equations	Time-scale constraints
S_1	$\rho_{13} = \rho_{14}$	$\gamma \le \alpha_1 - \max(\rho_{13}, \rho_{14})$
S_2	$\max(\rho_3, \rho_4, \rho_5, \rho_6, \rho_7, \rho_8) = \max(\rho_2, \rho_9)$	$\gamma \le \alpha_2 - \max(\rho_2, \rho_3, \rho_4, \rho_5, \rho_6, \rho_7, \rho_8, \rho_9)$
S_3	$\rho_2 = \max(\rho_3, \rho_5, \rho_6, \rho_7)$	$\gamma \le \alpha_3 - \max(\rho_2, \rho_3, \rho_5, \rho_6, \rho_7)$
S_4	$\rho_6 = \rho_{18}$	$\gamma \le \alpha_4 - \max(\rho_6, \rho_{18})$
S_5	$\rho_5 = \rho_{16}$	$\gamma \le \alpha_5 - \max(\rho_5, \rho_{16})$
S_6	$\max(\rho_7, \rho_8, \rho_{12}, \rho_{15}) = \max(\rho_9, \rho_{10}, \rho_{17})$	$\gamma \le \alpha_6 - \max(\rho_7, \rho_8, \rho_9, \rho_{10}, \rho_{12}, \rho_{15}, \rho_{17})$
S_7	$\rho_9 = \max(\rho_8, \rho_{15})$	$\gamma \le \alpha_7 - \max(\rho_8, \rho_9, \rho_{15})$
S_8	$\max(\rho_1, \rho_{12}) = \max(\rho_{10}, \rho_{11})$	$\gamma \le \alpha_8 - \max(\rho_1, \rho_{10}, \rho_{11}, \rho_{12})$
S_9	$\rho_{10} = \rho_{12}$	$\gamma \le \alpha_9 - \max(\rho_{10}, \rho_{12})$
$S_2 + S_3 + S_7$	$\rho_4 = \rho_{15}$	$\gamma \le \max(\alpha_2, \alpha_3, \alpha_7) - \max(\rho_4, \rho_{15})$
$S_2 + S_3$	$\max(\rho_4, \rho_8) = \rho_9$	$\gamma \le \max(\alpha_2, \alpha_3) - \max(\rho_4, \rho_8, \rho_9)$
$S_2 + S_7$	$\max(\rho_3, \rho_4, \rho_5, \rho_6, \rho_7) = \max(\rho_2, \rho_{15})$	$\gamma \le \max(\alpha_2, \alpha_7) - \max(\rho_2, \rho_3, \rho_4, \rho_5, \rho_6, \rho_7, \rho_{15})$
$S_6 + S_7 + S_9$	$\rho_7 = \rho_{17}$	$\gamma \le \max(\alpha_6, \alpha_7, \alpha_9) - \max(\rho_7, \rho_{17})$
$S_6 + S_7$	$\max(\rho_7, \rho_{12}) = \max(\rho_{10}, \rho_{17})$	$\gamma \le \max(\alpha_6, \alpha_7) - \max(\rho_7, \rho_{10}, \rho_{12}, \rho_{17})$
$S_6 + S_9$	$\max(\rho_7, \rho_8, \rho_{15}) = \max(\rho_9, \rho_{17})$	$\gamma \le \max(\alpha_6, \alpha_9) - \max(\rho_7, \rho_8, \rho_9, \rho_{15}, \rho_{17})$
$S_8 + S_9$	$\rho_1 = \rho_{11}$	$\gamma \le \max(\alpha_8, \alpha_9) - \max(\rho_1, \rho_{11})$

In each case, either the balance equation or the time-scale constraint must hold.

Limiting models in three time scales

In the heat shock response model of *E. coli*, we identify a time scale of interest using the chosen set of scaling exponents and derive a limiting model which approximates dynamics of the full chemical reaction network. Each limiting model involves a subset of species and reactions, and gives features of the full network during the time interval of interest.

To identify a time scale involving a limiting model with interesting dynamics (nondegenerate), we first need to determine a natural time scale of each species. Recall that a natural time scale of species S_i is the time period of order N^{γ_i} when $Z_i^{N,\gamma_i}(t)$ is of order 1. The natural time scale exponent γ_i for species S_i is rigorously determined by

$$\max_{k \in \Gamma_i^+ \cup \Gamma_i^-} (\gamma_i + \rho_k) = \alpha_i, \qquad (19)$$

where Γ_i^+ denotes the collection of reactions where the species number of S_i increases every time the reaction occurs. Similarly, Γ_i^- is the subset of reactions where the species number of S_i decreases every time the reaction occurs. In (19), the left-side term is the maximal order of magnitude of rates of reactions involving S_i and the right-side term is the order of magnitude of the species number for S_i. If times are earlier than those of order N^{γ_i} ($\gamma < \gamma_i$), fluctuations of species number of S_i due to the reactions involving S_i are not noticeable compared to magnitude of the species number of S_i. Then, the species number of S_i is approximated as its initial value. In the times of order N^{γ_i} ($\gamma = \gamma_i$), changes of species number of S_i due

to the reactions and the species number of S_i are similar in magnitude and behavior of the species number of S_i is described by its nondegenerate limit. If times are later than those of order N^{γ_i} ($\gamma > \gamma_i$), the species number of S_i fluctuates very rapidly due to the reactions involving S_i compared to the magnitude of the species number of S_i. Then, the averaged behavior of the species number of S_i is approximated by some function of other species numbers. Note that γ_i depends on α_i's and β_k's, and the time scale of the ith species may change if we use several sets of α_i's.

All values of α_i's and ρ_k's for three scalings which are used to derive limiting models are given in the Additional file 1: Table S4. The equations for normalized species numbers and the equation for $Z_{23}^{N,\gamma}$ which are used later in this section are given in the Additional file 1: Section 1 and Section 2, respectively. When we derive limiting models in three time scales, boundedness of the normalized species numbers is required. For first two time scales, we define stopping times so that the normalized species numbers are bounded up to those times. For the last time scale, we proved stochastic boundedness of some normalized species numbers in a finite time interval. For more details, see Additional file 1: Section 5.

Consider a model with the first set of scaling exponents including $\alpha_1 = 1$ and $\alpha_2 = \alpha_3 = 0$. Note that the first set of scaling exponents is valid when $\gamma \le 0$ based on the time scale constraints given in Table 4. Substituting $\alpha_2 = 0$ and ρ_k's for the first scaling to the equation for $Z_2^{N,\gamma}$ given in (11a), we have

$$Z_2^{N,\gamma}(t) = Z_2^{N,\gamma}(0) + R_3^t\left(N^\gamma\kappa_3 Z_3^{N,\gamma}\right) + R_4^t\left(N^\gamma\kappa_4 Z_1^{N,\gamma}\right)$$
$$+ R_5^t\left(N^{\gamma-1}\kappa_5 Z_3^{N,\gamma}\right) + R_6^t\left(N^{\gamma-1}\kappa_6 Z_3^{N,\gamma}\right)$$
$$+ R_7^t\left(N^{\gamma-1}\kappa_7 Z_3^{N,\gamma}\right) + R_8^t\left(N^{\gamma-2}\kappa_8 Z_7^{N,\gamma}\right)$$
$$- R_2^t\left(N^\gamma\kappa_2 Z_2^{N,\gamma}\right) - R_9^t\left(N^{\gamma-2}\kappa_9 Z_2^{N,\gamma} Z_6^{N,\gamma}\right).$$
$$(20)$$

When $\gamma = \gamma_2$, the maximal scaling exponent in the propensities of all reaction terms in (20) should be equal to the scaling exponent for the species number of S_2. Therefore, γ_2 satisfies

$$\max(\gamma_2, \gamma_2 - 1, \gamma_2 - 2) = 0 = \alpha_2, \quad (21)$$

and we get $\gamma_2 = 0$. Similarly, we get $\gamma_3 = \gamma_8 = 0$.

Next, we plug $\alpha_1 = 1$ and ρ_k's for the first scaling in the equation for $Z_1^{N,\gamma}$ and get

$$Z_1^{N,\gamma}(t) = Z_1^{N,\gamma}(0) + N^{-1}R_{13}^t\left(N^{\gamma-2}\kappa_{13}\right)$$
$$- N^{-1}R_{14}^t\left(N^{\gamma-1}\kappa_{14}Z_1^{N,\gamma}\right). \quad (22)$$

By comparing the maximal scaling exponent in the propensities of all reaction terms in (22) and the scaling exponent for the species number of S_1, γ_1 satisfies

$$\max(\gamma_1 - 2, \gamma_1 - 1) = 1 = \alpha_1, \quad (23)$$

and we get $\gamma_1 = 2$. Similarly, we get $\gamma_i > 0$ for $i = 4, 5, 6, 7, 9$. Among all natural time scale exponents of species, we choose the smallest one, $\gamma = 0$, and set $t \sim O(N^0) = O(1)$ as the first time scale we are interested in. Since $\gamma_1 > 0$, $Z_1^{N,0}(t) \to Z_1^0(0)$ as $N \to \infty$. Similarly, $Z_i^{N,0}(t) \to Z_i^0(0) = N_0^{-\alpha_i}X_i(0)$ for $i = 4, 5, 6, 7, 9$ as $N \to \infty$. To sum up, in this time scale with $\gamma = 0$, the species numbers of S_i's for $i = 1, 4, 5, 6, 7, 9$ change more slowly than other species numbers, and the species numbers with slow time scales are approximated as constant.

To derive the limiting equation for S_2, we set $\gamma = 0$ in (20). Since the 2nd, 3rd, and 4th reaction terms have propensities with $N^0 = 1$ and the species number of S_2 is of order 1, these reaction terms converge to nonzero limits in the limiting equation. On the other hand, the propensities of the 5th, 6th, 7th, 8th and 9th reaction terms are of order N^{-1} or N^{-2} which are smaller than the species number for S_2 of order 1. Therefore, these reaction terms converge to zero as $N \to \infty$ at least in the finite time interval. In the 2nd and 3rd reaction terms in (20), $Z_2^{N,0}(s) \to Z_2^0(s)$ and $Z_3^{N,0}(s) \to Z_3^0(s)$ as $N \to \infty$ since $\gamma_2 = \gamma_3 = 0$. Then, using $Z_1^{N,0}(s) \to Z_1^0(0)$ as $N \to \infty$, the limit of $Z_2^{N,0}$ satisfies

$$Z_2^0(t) = Z_2^0(0) + R_3^t\left(\kappa_3 Z_3^0\right) + R_4^t\left(\kappa_4 Z_1^0(0)\right) - R_2^t\left(\kappa_2 Z_2^0\right).$$

Similarly, we get a limiting model with Z_2^0, Z_3^0, and Z_8^0 for $\gamma = 0$ as given in (3).

Next, consider a model with the second set of scaling exponents including $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Note that the second set of scaling exponents is valid when $\gamma \leq 1$ based on the time scale constraints given in Table 4. To determine the natural time scale of S_6, substitute $\alpha_6 = 0$ and ρ_k's for the second scaling in the equation for $Z_6^{N,\gamma}$, and we have

$$Z_6^{N,\gamma}(t) = Z_6^{N,\gamma}(0) + R_7^t\left(N^{\gamma-1}\kappa_7 Z_3^{N,\gamma}\right)$$
$$+ R_8^t\left(N^{\gamma-2}\kappa_8 Z_7^{N,\gamma}\right) + R_{12}^t\left(N^{\gamma-1}\kappa_{12}Z_9^{N,\gamma}\right)$$
$$+ R_{15}^t\left(N^{\gamma-1}\kappa_{15}Z_4^{N,\gamma}Z_7^{N,\gamma}\right)$$
$$- R_9^t\left(N^{\gamma-2}\kappa_9 Z_2^{N,\gamma}Z_6^{N,\gamma}\right)$$
$$- R_{10}^t\left(N^{\gamma-1}\kappa_{10}Z_6^{N,\gamma}Z_8^{N,\gamma}\right)$$
$$- R_{17}^t\left(N^{\gamma-2}\kappa_{17}Z_6^{N,\gamma}\right).$$
$$(24)$$

Comparing the exponents inside and outside of the reaction terms in (24), γ_6 satisfies

$$\max(\gamma_6 - 1, \gamma_6 - 2) = 0 = \alpha_6, \quad (25)$$

and we get $\gamma_6 = 1$. Similarly, we get $\gamma_7 = \gamma_8 = 1$, $\gamma_i < 1$ for $i = 2, 3$, and $\gamma_i > 1$ for $i = 1, 4, 5, 9$. We already get the temporal behavior of species numbers of S_2, S_3, and S_8 through the limiting model when $\gamma = 0$. Thus, we set $t \sim O(N^1)$ as the second time scale we are interested in, and derive a limiting model for S_6, S_7, and S_8 when $\gamma = 1$. Note that species S_8 is involved in the limiting models for both $\gamma = 0$ and $\gamma = 1$, since we use different sets of scaling exponents in these models. For $i = 1, 4, 5, 9$ $Z_i^{N,1}(t) \to Z_i^1(0)$ as $N \to \infty$, since $\gamma_i > 1$. Thus, in the 12th and 15th reaction terms in (24), $Z_9^{N,1}(s) \to Z_9^1(0)$ and $Z_4^{N,1}(s) \to Z_4^1(0)$ as $N \to \infty$. Since the propensities of the 8th, 9th, and 17th reaction terms in (24) are of order $N^{\gamma-2} = N^{-1}$ for $\gamma = 1$ and the species number of S_6 is of order 1, these reaction terms go to zero as $N \to \infty$. In the 10th and 15th reaction terms in (24), $Z_6^{N,1}(s)$, $Z_7^{N,1}(s)$, and $Z_8^{N,1}(s)$ are asymptotically $O(1)$ and converge to $Z_6^1(s)$, $Z_7^1(s)$, and $Z_8^1(s)$ as $N \to \infty$ since $\gamma_6 = \gamma_7 = \gamma_8 = 1$.

Now, consider the asymptotic behavior of the 7th reaction term in (24) when $\gamma = 1$. Since $\gamma_3 < 1$, $Z_3^{N,1}(t)$ fluctuates very much, and there exists no functional limit as $N \to \infty$. However, $\int_0^t Z_3^{N,1}(s)\,ds$ still converges, which gives the averaged behavior of the normalized species number of S_3. To get the limit of $\int_0^t Z_3^{N,1}(s)\,ds$, we plug the second set of scaling exponents in the equation for $Z_3^{N,\gamma}$ and obtain

$$Z_3^{N,1}(t) = Z_3^{N,1}(0) + R_2^t\left(N\kappa_2 Z_2^{N,1}\right) - R_3^t\left(N\kappa_3 Z_3^{N,1}\right)$$
$$- R_5^t\left(\kappa_5 Z_3^{N,1}\right) - R_6^t\left(\kappa_6 Z_3^{N,1}\right) - R_7^t\left(\kappa_7 Z_3^{N,1}\right).$$
(26)

The law of large numbers of Poisson processes gives an asymptotic limit of the scaled reaction terms as

$$\lim_{N\to\infty}\sup_{u\le u_0}\left|\frac{Y_k(N^{\alpha_i}u)}{N^{\alpha_i}} - u\right| = 0, \qquad u_0 > 0 \quad (27)$$

where the Y_k's are unit Poisson processes and $\alpha_i > 0$. For example, the 2nd reaction term in (26) divided by N is approximated as

$$\frac{R_2^t(N\kappa_2 Z_2^{N,1})}{N} = \frac{Y_2(\int_0^t N\kappa_2 Z_2^{N,1}(s)\,ds)}{N} \approx \int_0^t \kappa_2 Z_2^{N,1}(s)\,ds.$$

Dividing (26) by N and using the law of large numbers for Poisson processes, we get

$$\int_0^t \left(\kappa_2 Z_2^{N,1}(s) - \kappa_3 Z_3^{N,1}(s)\right)\,ds \longrightarrow 0, \quad (28)$$

as $N \to \infty$.

We introduce an auxiliary variable to make the limiting model closed and define

$$Z_{23}^{N,\gamma}(t) \equiv Z_2^{N,\gamma}(t) + Z_3^{N,\gamma}(t).$$

Plugging $\alpha_2 = \alpha_3 = 0$ and ρ_k's in the second scaling in the equation for $Z_{23}^{N,\gamma}$, we get

$$Z_{23}^{N,\gamma}(t) = Z_{23}^{N,\gamma}(0) + R_4^t\left(N^{\gamma-1}\kappa_4 Z_1^{N,\gamma}\right) + R_8^t\left(N^{\gamma-2}\kappa_8 Z_7^{N,\gamma}\right)$$
$$- R_9^t\left(N^{\gamma-2}\kappa_9 Z_2^{N,\gamma} Z_6^{N,\gamma}\right).$$
(29)

Since $Z_{23}^{N,\gamma_{23}}(t) \sim O(1)$ where γ_{23} denotes a natural time scale exponent of S_{23}, we compare the scaling exponents of N in the reaction terms in (29) and the scaling exponent of N outside the reaction terms. Then γ_{23} satisfies

$$\max(\gamma_{23} - 1, \gamma_{23} - 2) = 0 = \max(\alpha_2, \alpha_3),$$

and we get $\gamma_{23} = 1$. Since these reaction terms have $N^{\gamma-2} = N^{-1}$ in their propensities when $\gamma = 1$, which is smaller than the species number for S_{23} of order 1, these reaction terms converge to zero as $N \to \infty$. Using $Z_1^{N,1}(s) \to Z_1^1(0)$, the limit of $Z_{23}^{N,1}$ satisfies

$$Z_{23}^1(t) = Z_{23}^1(0) + R_4^t\left(\kappa_4 Z_1^1(0)\right).$$

Adding and subtracting terms in (28) and dividing the equation by $-(\kappa_2 + \kappa_3)$, we get

$$\int_0^t \left(Z_3^{N,1}(s) - \frac{\kappa_2}{\kappa_2 + \kappa_3} Z_{23}^{N,1}(s)\right)\,ds \longrightarrow 0,$$

as $N \to \infty$, and this is used to obtain the limit of the 7th reaction term in (24). Letting $N \to \infty$, the limiting equation for $Z_6^{N,1}$ is given as

$$Z_6^1(t) = Z_6^1(0) + R_7^t\left(\frac{\kappa_2\kappa_7}{\kappa_2 + \kappa_3} Z_{23}^1\right) + R_{12}^t\left(\kappa_{12}Z_9^1(0)\right)$$
$$+ R_{15}^t\left(\kappa_{15}Z_4^1(0)Z_7^1\right) - R_{10}^t\left(\kappa_{10}Z_6^1 Z_8^1\right).$$
(30)

In (30), note that $R_{12}^t(\kappa_{12}Z_9^1(0)) = 0$ since $X_9(0) = 0$ as given in Table 1. Limiting equations for $Z_7^{N,1}$ and $Z_8^{N,1}$ can be derived similarly, and a limiting model with Z_{23}^1, Z_6^1, Z_7^1, and Z_8^1 for $\gamma = 1$ is given in (4).

Last, consider a model with the third scaling exponents with $\alpha_1 = 0$ and $\alpha_2 = \alpha_3 = 1$. To derive a limiting equation for $Z_{23}^{N,2}$, we plug ρ_k's and $\alpha_2 = \alpha_3 = 1$ for the third scaling in the equation for $Z_{23}^{N,\gamma}$ and get

$$Z_{23}^{N,2}(t) = Z_{23}^{N,2}(0) + N^{-1}\left[R_4^t\left(N\kappa_4 Z_1^{N,2}\right) + R_8^t\left(\kappa_8 Z_7^{N,2}\right)\right.$$
$$\left. - R_9^t\left(N\kappa_9 Z_2^{N,2} Z_6^{N,2}\right)\right].$$
(31)

In (31), the 8th reaction term is asymptotically zero, since the term is of order N^{-1}. Using the law of large numbers for Poisson processes in (27), the 4th and the 9th terms in (31) are asymptotically equal to

$$\int_0^t \left(\kappa_4 Z_1^{N,2}(s) - \kappa_9 Z_2^{N,2}(s) Z_6^{N,2}(s)\right)\,ds. \quad (32)$$

Since $\gamma_1 = 2$, $Z_1^{N,2}(s) \to Z_1^2(s)$ as $N \to \infty$. On the other hand, since $\gamma_2, \gamma_6 < 2$, both $Z_2^{N,2}(s)$ and $Z_6^{N,2}(s)$ in (32) fluctuate rapidly and we must identify the averaged limit. $Z_3^{N,2}$ is also averaged, since $\gamma_3 < 2$. We actually show convergence of the fast-fluctuating species numbers of S_2 and S_3 to some limits in the Additional file 1: Section 5.1. For any $\epsilon > 0$ and for any t such that $\epsilon < t \le \tau_\infty^2$,

$$Z_2^{N,2}(t) \longrightarrow \bar{Z}_2^2(t), \quad (33)$$

$$Z_3^{N,2}(t) \longrightarrow \bar{Z}_3^2(t), \quad (34)$$

uniformly as $N \to \infty$.

On the other hand, since $\gamma_6 < 2$, $\int_0^t Z_6^{N,2}(s)\,ds$ converges to a limit which gives averaged behavior of the normalized species number of S_6. Using the equation for $Z_6^{N,\gamma}$, we get

$$Z_6^{N,2}(t) = Z_6^{N,2}(0) + R_7^t(N^2\kappa_7 Z_3^{N,2}) + R_8^t(\kappa_8 Z_7^{N,2})$$
$$+ R_{12}^t(N^2\kappa_{12}Z_9^{N,2}) + R_{15}^t(N\kappa_{15}Z_4^{N,2}Z_7^{N,2})$$
$$- R_9^t(N\kappa_9 Z_2^{N,2}Z_6^{N,2}) - R_{10}^t(N^2\kappa_{10}Z_6^{N,2}Z_8^{N,2})$$
$$- R_{17}^t(\kappa_{17}Z_6^{N,2}).$$
(35)

Dividing (35) by N^2, using the law of large numbers for Poisson processes in (27), and using the stochastic boundedness of the propensities of the 8th, 9th, 15th, and 17th

reaction terms in the finite time interval shown in the Additional file 1: Section 5.1, we get

$$\int_0^t \left(\kappa_7 Z_3^{N,2}(s) + \kappa_{12} Z_9^{N,2}(s) - \kappa_{10} Z_6^{N,2}(s) Z_8^{N,2}(s) \right) ds \longrightarrow 0,$$

(36)

as $N \to \infty$. Therefore, a difference between the 10th and 12th reaction terms is approximated in terms of

$$\int_0^t \kappa_7 Z_3^{N,2}(s)\, ds,$$

(37)

which converges to $\int_0^t \kappa_7 \bar{Z}_3^2(s)\, ds$ from (34). Therefore, we get

$$\int_0^t \left(\kappa_{10} Z_6^{N,2}(s) Z_8^{N,2}(s) - \kappa_{12} Z_9^{N,2}(s) \right) ds \longrightarrow \int_0^t \kappa_7 \bar{Z}_3^2(s)\, ds,$$

(38)

as $N \to \infty$.

Now, from the equations for $Z_8^{N,\gamma}$ and $Z_9^{N,\gamma}$, we get

$$Z_8^{N,2}(t) = Z_8^{N,2}(0) + N^{-2} \Big[R_1^t \left(N^2 \kappa_1 \right) + R_{12}^t \left(N^2 \kappa_{12} Z_9^{N,2} \right)$$
$$- R_{10}^t \left(N^2 \kappa_{10} Z_6^{N,2} Z_8^{N,2} \right) - R_{11}^t \left(N^2 \kappa_{11} Z_8^{N,2} \right) \Big],$$

(39a)

$$Z_9^{N,2}(t) = Z_9^{N,2}(0) + N^{-2} \Big[R_{10}^t \left(N^2 \kappa_{10} Z_6^{N,2} Z_8^{N,2} \right)$$
$$- R_{12}^t \left(N^2 \kappa_{12} Z_9^{N,2} \right) \Big].$$

(39b)

Using the law of large numbers of Poisson processes in (27), the reaction terms in (39a) and (39b) are asymptotically equal to

$$N^{-2} \Big[R_1^t \left(N^2 \kappa_1 \right) + R_{12}^t \left(N^2 \kappa_{12} Z_9^{N,2} \right) - R_{10}^t \left(N^2 \kappa_{10} Z_6^{N,2} Z_8^{N,2} \right)$$
$$- R_{11}^t \left(N^2 \kappa_{11} Z_8^{N,2} \right) \Big]$$
$$\approx \int_0^t \left(\kappa_1 + \kappa_{12} Z_9^{N,2}(s) - \kappa_{10} Z_6^{N,2}(s) Z_8^{N,2}(s) \right.$$
$$\left. - \kappa_{11} Z_8^{N,2}(s) \right) ds,$$

(40a)

$$N^{-2} \Big[R_{10}^t (N^2 \kappa_{10} Z_6^{N,2} Z_8^{N,2}) - R_{12}^t (N^2 \kappa_{12} Z_9^{N,2}) \Big]$$
$$\approx \int_0^t \left(\kappa_{10} Z_6^{N,2}(s) Z_8^{N,2}(s) - \kappa_{12} Z_9^{N,2}(s) \right) ds.$$

(40b)

Using (40a), (40b), and (38), the limiting equations of (39a) and (39b) are given as

$$Z_8^2(t) = Z_8^2(0) + \int_0^t \left(\kappa_1 - \kappa_7 \bar{Z}_3^2(s) - \kappa_{11} Z_8^2(s) \right) ds,$$

$$Z_9^2(t) = Z_9^2(0) + \int_0^t \kappa_7 \bar{Z}_3^2(s)\, ds.$$

(41)

In (41), note that $Z_9^2(0) = 0$ since $X_9(0) = 0$ as given in Table 1.

Since $Z_8^{N,2}(0) > 0$ and balance conditions are satisfied, $Z_8^{N,2}(t) \neq 0$ in the finite time interval. Since $\gamma_8 = 2$,

$$\frac{1}{\kappa_{10} Z_8^{N,2}(t)} \longrightarrow \frac{1}{\kappa_{10} Z_8^2(t)}.$$

(42)

Using (38) and (42), $\int_0^t Z_6^{N,2}(s)\, ds$ is averaged as

$$\int_0^t Z_6^{N,2}(s)\, ds \longrightarrow \int_0^t \frac{\kappa_7 \bar{Z}_3^2(s) + \kappa_{12} Z_9^2(s)}{\kappa_{10} Z_8^2(s)}\, ds.$$

(43)

From (33) and (43), we get

$$\int_0^t \kappa_9 Z_2^{N,2}(s) Z_6^{N,2}(s)\, ds \longrightarrow \int_0^t \kappa_9 \bar{Z}_2^2(s)$$
$$\times \left(\frac{\kappa_7 \bar{Z}_3^2(s) + \kappa_{12} Z_9^2(s)}{\kappa_{10} Z_8^2(s)} \right) ds.$$

(44)

For more details used in (43) and (44), see Lemma 1.5 and Theorem 2.1 in [15]. Finally, we get the limiting equation of $Z_{23}^{N,2}$ as

$$Z_{23}^2(t) = Z_{23}^2(0) + \int_0^t \Big[\kappa_4 Z_1^2(s) - \kappa_9 \bar{Z}_2^2(s)$$
$$\times \left(\frac{\kappa_7 \bar{Z}_3^2(s) + \kappa_{12} Z_9^2(s)}{\kappa_{10} Z_8^2(s)} \right) \Big] ds.$$

Theorem 1. *For* $\gamma = 0$, $\{Z_2^{N,0}, Z_3^{N,0}, Z_8^{N,0}\}$ *converges to the solution of (3) for* $t \in [0, \tau_\infty^0)$. *For* $\gamma = 1$, $\{Z_{23}^{N,1}, Z_6^{N,1}, Z_7^{N,1}, Z_8^{N,1}\}$ *converges to the solution of (4) for* $t \in [0, \tau_\infty^1)$. *In (3),* Z_8^0 *is a discrete process, while* Z_8^1 *is a deterministic process in (4). For* $\gamma = 2$, $\{Z_1^{N,2}, Z_{23}^{N,2}, Z_4^{N,2}, Z_5^{N,2}, Z_8^{N,2}, Z_9^{N,2}\}$ *converges to the solution of (5) for* $t \in [0, \tau_\infty^2)$.

Conditional equilibrium distributions

In the previous section, we derived limiting models in three different time scales. Except for the subset of species in the limiting model, the remaining species are approximated as constants in the first time scale, since their natural time scale exponents (γ_i) are larger than $\gamma = 0$, i.e., species with $\gamma_i > \gamma = 0$ did not start to fluctuate at these times yet. In the second and third time scales, there are subsets of species whose natural time scale exponents are smaller than $\gamma = 1$ and 2, respectively. Normalized species numbers with $\gamma_i < \gamma$ fluctuate very rapidly at these times and their averaged behavior is approximated in terms of other variables which converge to a nondegenerate limit. For those species, the normalized species numbers do not converge to a limit in a functional sense, but still we can find a limit in a probabilistic sense (i.e.

convergence in distribution) and their distribution. Conditioned on the normalized species numbers which converge to a nondegenerate limit in the time scale of interest, we can find the *conditional equilibrium* (or the *local averaging*) distributions of species numbers whose natural time scale exponents are smaller than the time scale exponents of interests. Conditioning on the normalized species numbers which converge to a nondegenerate limit is similar to fixing slowly-moving variables and describing behavior of the fast-fluctuating variables in terms of slowly-moving variables treating them as constants. In the next remark, we give a conditional equilibrium distribution of the subset of species with natural time scale exponents smaller than $\gamma = 1$ and $\gamma = 2$.

Remark 2. *For $\gamma = 1$, for each $t > 0$, $\left(Z_2^{N,1}(t), Z_3^{N,1}(t)\right)$ converges in distribution to $\left(\hat{Z}_2^1(t), \hat{Z}_3^1(t)\right)$ such that $\left(\hat{Z}_2^1(t), \hat{Z}_3^1(t)\right)$ conditioned on $Z_{23}^1(t)$ has a binomial distribution with parameter*

$$\frac{\kappa_3}{\kappa_2 + \kappa_3}, \qquad \frac{\kappa_2}{\kappa_2 + \kappa_3},$$

respectively, that is,

$$P\left\{\hat{Z}_2^1(t) = z_2, \hat{Z}_3^1(t) = m - z_2 | Z_{23}^1(t) = m\right\}$$
$$= C(m, z_2) \left(\frac{\kappa_3}{\kappa_2 + \kappa_3}\right)^{z_2} \left(\frac{\kappa_2}{\kappa_2 + \kappa_3}\right)^{m-z_2}.$$

For $\gamma = 2$, for each $t > 0$, $\left(Z_6^{N,2}(t), Z_7^{N,2}(t)\right)$ converges in distribution to $\left(\hat{Z}_6^2(t), \hat{Z}_7^2(t)\right)$ where $\hat{Z}_6^2(t)$ and $\hat{Z}_7^2(t)$ are independent Poisson distributed random variables with parameters

$$\frac{\frac{\kappa_2 \kappa_7}{\kappa_2 + \kappa_3} Z_{23}^2(t) + \kappa_{12} Z_9^2(t)}{\kappa_{10} Z_8^2(t)},$$

and

$$\frac{\frac{\kappa_3 \kappa_9}{\kappa_2 + \kappa_3} Z_{23}^2(t)}{\kappa_{15} Z_4^2(t)} \cdot \frac{\frac{\kappa_2 \kappa_7}{\kappa_2 + \kappa_3} Z_{23}^2(t) + \kappa_{12} Z_9^2(t)}{\kappa_{10} Z_8^2(t)}.$$

The detailed method to compute conditional equilibrium distributions is given in Section 6 in [9].

Mean value of the random variable with a binomial distribution, $B(n, p)$, is equal to np. Therefore, for $\gamma = 1$, we treat $Z_{23}^1(t)$ as constant and get a limit of the averaged values for $Z_2^{N,1}(t)$ and $Z_3^{N,1}(t)$ as

$$\bar{Z}_2^1(t) = Z_{23}^1(t) \times \frac{\kappa_3}{\kappa_2 + \kappa_3},$$
$$\bar{Z}_3^1(t) = Z_{23}^1(t) \times \frac{\kappa_2}{\kappa_2 + \kappa_3}.$$

Mean value of the random variable with a Poisson distribution, $Pois(\lambda)$, is equal to λ, and we obtain a limit of the averaged values for $Z_6^{N,2}(t)$ and $Z_7^{N,2}(t)$ as the parameters given in Remark 2.

Simulation results

Recall that the normalized species numbers after a time change are defined as

$$Z_i^{N,\gamma}(t) = N^{-\alpha_i} X_i^N(tN^\gamma).$$

Using the limiting models in the three time scales given in (3)-(5), we approximate the species numbers in the full model by unnormalizing the species numbers and applying time change backward as

$$X_i(t) = X_i^{N_0}(t) \approx \lim_{N \to \infty} \left(\frac{N_0}{N}\right)^{\alpha_i} X_i^N \left(t\left(\frac{N}{N_0}\right)^\gamma\right)$$
$$= N_0^{\alpha_i} Z_i^\gamma(tN_0^{-\gamma}),$$

using a real value $N_0 = 100$ for the parameter. In Figures 2, 3, 4 and Figure 5(a)-(d), the panels located in the left column give mean and standard deviation from the mean of stochastic simulation for $X_i(t)$ and the panels located in the right column give mean and standard deviation from the mean of simulation for $N_0^{\alpha_i} Z_i^\gamma(tN_0^{-\gamma})$ using the limiting models. The mean and standard deviation of species numbers are computed from 3000 realizations of the sample path of the stochastic simulation.

In Figure 2, we compare the simulation for the full model and for the approximation using the limiting model in the first scaling. The first scaling ($\gamma = 0$) is for the times of order $N_0^0 = 1$ sec, and we look at the evolution of mean and standard deviation of the species numbers up to 100 sec. The full model and the limiting model for $\gamma = 0$ are stochastic, and the limiting model approximates the evolution of statistics of the species numbers quite precisely. As shown in Figure 2(f) $Z_8^0(t)$ overestimates $X_8(t)$, since the limiting model does not include reactions consuming S_8. Therefore, consumption of S_8 may not be captured well in the approximation.

In Figure 3, we compare the simulation for the full model and for the approximation using the limiting model in the second scaling. Since the second scaling ($\gamma = 1$) is for the times of order $N_0^1 = 100$ sec, we observe the evolution of the species numbers up to 1000 sec. In this time scale, the evolution of S_8 shown in Figure 3(h) is approximated by a deterministic variable. The evolution of the species number of S_8 in the full model given in Figure 3(g) is stochastic, but its standard deviation is very small. As in the previous time scale, $N_0 Z_8^1(tN_0^{-1})$ slightly overestimates $X_8(t)$, since the limiting model does not include any consumptions of S_8. The remaining three species, S_{23}, S_6, and S_7 are approximated by stochastic variables. The increasing species number of S_{23} in time and the rapid
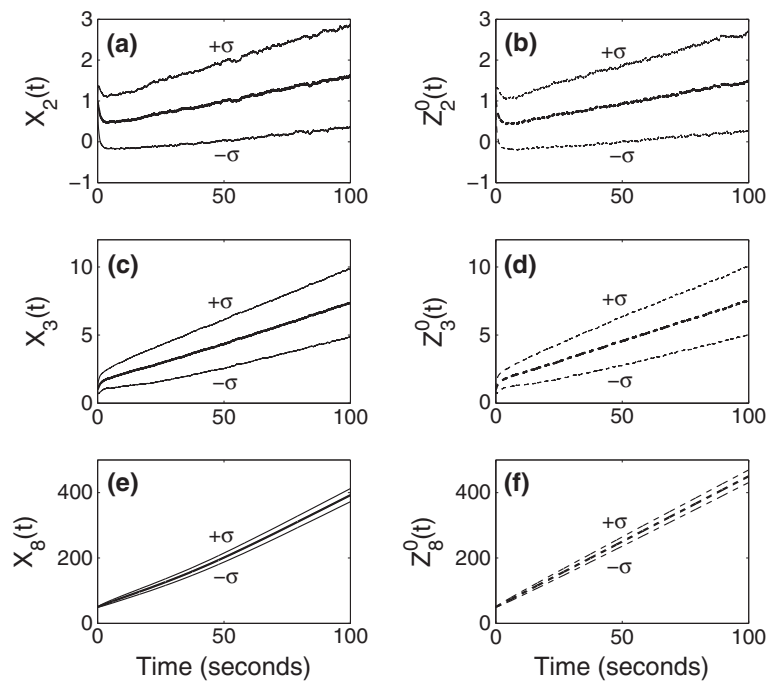
Figure 2 Simulation results when $\gamma = 0$. Simulation of the full model (left) and that of approximation using the limiting model (right) when the time is of order $N_0^0 (= 1)$.

decrease in species number of S_6 are well captured by the limiting model. The species numbers of S_7 are described by stochastic variables both in the full model and in the limiting model. The behavior of S_7 in two models is not exactly the same, and discrepancy of the mean species numbers of S_7 comes from the approximation of $X_4(t)$ in terms of its initial value. In the limiting model, S_7 is approximated as a stochastic process decreasing by 1 with the propensity proportional to $Z_4^1(0)$. However, $X_4(t)$ increases during the times in [0,1000] sec in the full model, and this difference gives slower decreasing rate of the mean number of S_7 in the limiting model than that in the full model.

In Figure 4 and Figure 5(a)-(d), we compare the simulation for the full model and for the approximation obtained from the limiting model in the third scaling. Since the third scaling ($\gamma = 2$) is for the times of order $N_0^2 = 10,000$ sec, we look at the simulation up to 20,000 sec. In this time scale, the limiting model is stochastic. The species number of S_1 in the limiting model is approximated by a stochastic discrete variable increasing and decreasing by 1, and the remaining species numbers in the limiting model satisfy stochastic equations driven by the stochastic discrete variable Z_1^2. As we have seen in the proof of Theorem 1 in the Additional file 1: Section 5.1, the processes for S_1 in the full model and in the limiting model are exactly the same. Therefore, we use a same series of random numbers, when we simulate the full

and limiting models. In Figure 4(b), the process for S_1 is random, but its standard deviation is very small. Therefore, in one realization of simulation of the limiting model, behavior of S_1 appears as constant. Since all the remaining variables in the limiting model are governed by the variable for S_1 and they satisfy the stochastic differential equations, evolution of one sample path of the species numbers for S_{23}, S_4, S_5, S_8, and S_9 in the limiting model looks like a solution of the system of ordinary differential equations.

In Figure 5, (e)-(h) are the species numbers for S_6 and S_7 in the full model and their averaged values in the limiting model. Note that the species numbers for S_6 and S_7 do not appear in the limiting model, since their values are approximated in terms of other species numbers. Therefore, the difference between mean species numbers for S_6 and S_7 in the full model and those in the approximation does not affect the error directly. For $\gamma = 2$, $Z_6^{N,2}$ and $Z_7^{N,2}$ are asymptotically averaged out by the variables in the limiting model as given in Remark 2. Since the averaged value for S_6 plays an important role in the evolution of Z_{23}^2 in the limiting model and since the averaged value for S_7 gives the conditional mean value for S_7 in the limiting model, we compare the species numbers of S_6 and S_7 in the full model and the approximated averaged values in the limiting model. In Figure 5(f) and (h), we plot the mean and standard deviation from the mean for
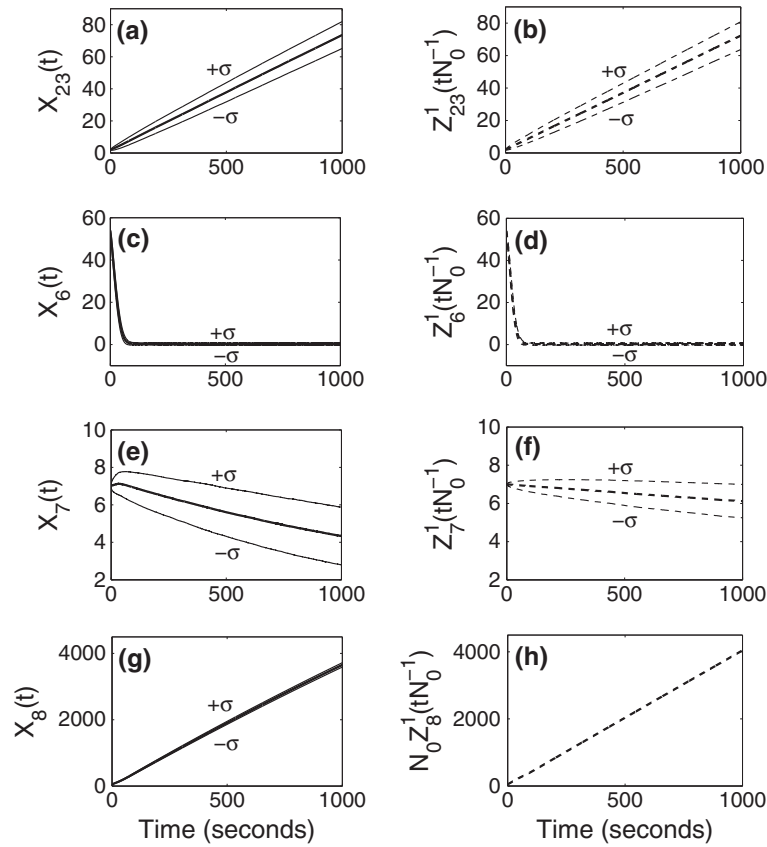
Figure 3 Simulation results when $\gamma = 1$. Simulation of the full model (left) and that of approximation using the limiting model (right) when the time is of order $N_0 (= 100)$.

$$\bar{Z}_6^2\left(tN_0^{-2}\right) = \frac{\frac{\kappa_2\kappa_7}{\kappa_2+\kappa_3}Z_{23}^2\left(tN_0^{-2}\right) + \kappa_{12}Z_9^2\left(tN_0^{-2}\right)}{\kappa_{10}Z_8^2\left(tN_0^{-2}\right)},$$

$$\bar{Z}_7^2\left(tN_0^{-2}\right) = \frac{\kappa_9\left(\frac{\kappa_3}{\kappa_2+\kappa_3}Z_{23}^2\left(tN_0^{-2}\right)\right)\left(\frac{\frac{\kappa_2\kappa_7}{\kappa_2+\kappa_3}Z_{23}^2\left(tN_0^{-2}\right)+\kappa_{12}Z_9^2\left(tN_0^{-2}\right)}{\kappa_{10}Z_8^2\left(tN_0^{-2}\right)}\right)}{\kappa_{15}Z_4^2\left(tN_0^{-2}\right)},$$

in time. They are stochastic variables determined by the ones in the limiting model with very small fluctuations. Since $\bar{Z}_6^2(tN_0^{-2})$ and $\bar{Z}_7^2(tN_0^{-2})$ describe averaged behavior of S_6 and S_7, $X_6(t)$ and $X_7(t)$ in Figure 5(e) and (g) have more fluctuations than the averaged species numbers in Figure 5(f) and (h).

In Figure 5(e)-(h) there is a discrepancy between the species numbers and their averaged values in the very early time, and the discrepancy comes from a disagreement in initial values of the species numbers in the full model and those of the averaged values in the limiting model. The integrated species numbers for S_6 and S_7 up to times of order 10,000 are supposed to be approximated by the integrated averaged values over the time

interval, and the initial difference is due to a boundary layer phenomenon.

Error estimates

In the previous sections, we scaled species numbers and derived their limit to approximate temporal behavior of the species numbers in the full network. Among three limiting models given in (3)-(5), the first two are systems with discrete variables (except for Z_8^1) which change by integer values. On the other hand, the last one is a hybrid system with both discrete and continuous variables. A discrete variable Z_1^2 increases or decreases by one and stochasticity of all other variables comes from how much Z_1^2 fluctuates. Since Z_1^2 rarely changes at the times of our interest, the rest of the variables in (5) behaves such as a solution of systems of ordinary differential equations. Our choice of the scaling parameter value, $N_0 = 100$, is not very large and it is possible that the limiting model does not contain enough fluctuations as much as the full network actually has due to our assumption that N_0 is replaced by a large parameter N.

In this section, we estimate an error between the normalized species numbers and their limit given in (5) at the times of 10,000 sec. Define
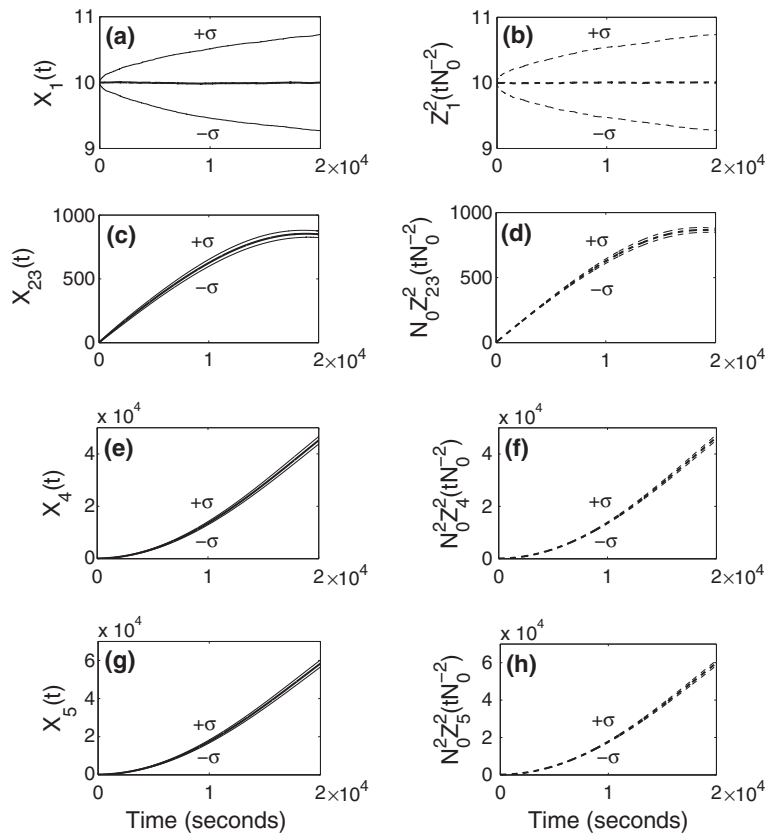
Figure 4 Simulation results when $\gamma = 2$. Simulation of the full model (left) and that of approximation using the limiting model (right) when the time is of order N_0^2 $(= 10000)$.

$$U^N(t) = N^{1/2}\left(Z_{23}^{N,2}(t) - Z_{23}^2(t), Z_4^{N,2}(t) - Z_4^2(t), Z_5^{N,2}(t)\right.$$
$$\left. - Z_5^2(t), Z_8^{N,2}(t) - Z_8^2(t), Z_9^{N,2}(t) - Z_9^2(t)\right)^T,$$

and denote $U(t) = (U_{23}(t), U_4(t), U_5(t), U_8(t), U_9(t))^T$ as a limit of $U^N(t)$ as N goes to infinity. Note that we do not consider an error between $Z_1^{N,2}(t)$ and $Z_1^2(t)$, since they are exactly the same processes. In the next remark, we show that $U^N(t)$ converges to $U(t)$ in the probabilistic sense and thus the error between $Z_i^{N,2}(t)$ and $Z_i^2(t)$ is approximately of order $N_0^{-1/2} = 0.1$. Since $U(t)$ gives an explicit form of the error, we have better approximation of $X_i(t)$ for $\gamma = 2$ as

$$X_i(t) \approx N_0^{\alpha_i}\left(Z_i^2(tN_0^{-2}) + N_0^{-1/2}U_i(tN_0^{-2})\right),$$

for $S_{23}, S_4, S_5, S_8,$ and S_9.

Remark 3. *For $\gamma = 2$, for each $t > 0$, $U^N(t)$ converges in distribution to $U(t)$ which is a solution of*

$$U(t) = U(0) + \int_0^t (1, 0, 0, 0, 0)^T \sqrt{\kappa_4 Z_1^2(s) + \kappa_9 \bar{Z}_2^2(s)\bar{Z}_6^2(s)}\, dW(s)$$

$$+ \int_0^t \begin{bmatrix} C_{23}(s)U_{23}(s) + C_8(s)U_8(s) + C_9(s)U_9(s) \\ \dfrac{\kappa_2\kappa_6}{\kappa_2 + \kappa_3}U_{23}(s) - \kappa_{18}U_4(s) \\ \dfrac{\kappa_2\kappa_5}{\kappa_2 + \kappa_3}U_{23}(s) - \kappa_{16}U_5(s) \\ -\dfrac{\kappa_2\kappa_7}{\kappa_2 + \kappa_3}U_{23}(s) - \kappa_{11}U_8(s) \\ \dfrac{\kappa_2\kappa_7}{\kappa_2 + \kappa_3}U_{23}(s) \end{bmatrix} ds,$$

where $W(t)$ is a standard Brownian motion and

$$C_{23}(s) = -\frac{\kappa_9}{\kappa_2 + \kappa_3}\left(\kappa_3\bar{Z}_6^2(s) + \frac{\kappa_2\kappa_7}{\kappa_{10}} \cdot \frac{\bar{Z}_2^2(s)}{Z_8^2(s)}\right),$$

$$C_8(s) = \kappa_9 \frac{\bar{Z}_2^2(s)\bar{Z}_6^2(s)}{Z_8^2(s)},$$

$$C_9(s) = -\frac{\kappa_9\kappa_{12}}{\kappa_{10}} \cdot \frac{\bar{Z}_2^2(s)}{Z_8^2(s)}.$$

The detailed method to compute an error using the central limit theorem is derived in [14].
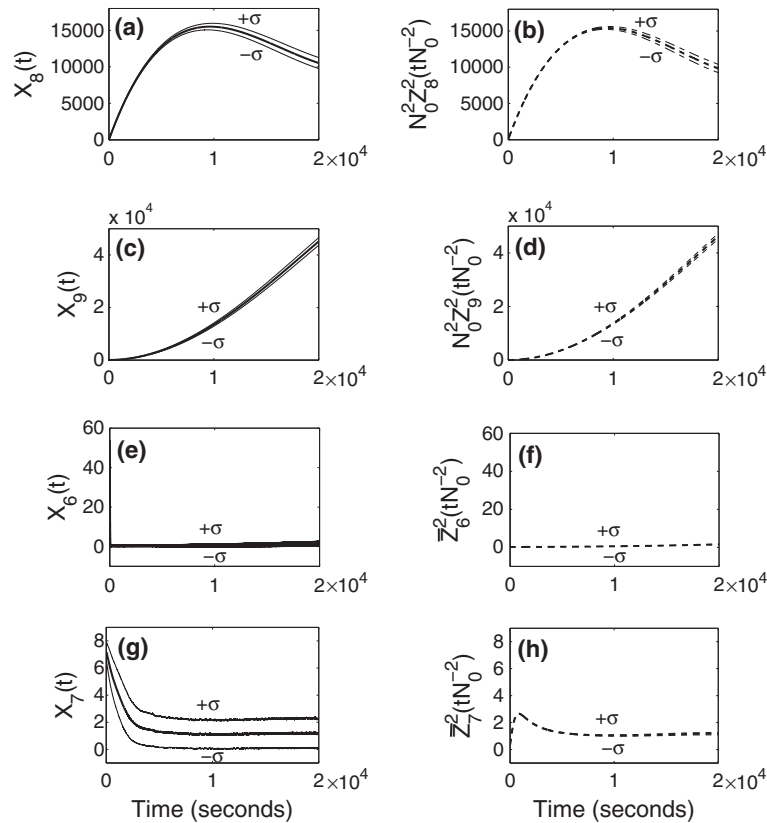
Figure 5 Simulation results when $\gamma = 2$ (continued). Simulation of the full model (left) and that of approximation using the limiting model (right) when the time is of order N_0^2. Figures (e), (f), (g), and (h) are simulation results for species 6 and 7. The graphs (f) and (h) give approximation of the averaged species numbers of S_6 and S_7.

Estimating order of magnitude of an error is an analogue of that in van Kampen's system size expansion [16]. A difference is that in the system size expansion, the system state representing the species numbers is scaled by the system size Ω and noise between the scaled process and its deterministic value is approximated as a random variable of order $\Omega^{-1/2}$. In our approach N is not a system size but a parameter for scaling, and species numbers are scaled by powers of N. Though the limiting model for $\gamma = 2$ is not deterministic, it is still possible to estimate an error analytically due to the fact that $Z_1^2(t)$ which produces stochasticity in the limiting model is an exact process equal to $Z_1^{N,2}(t)$. Another difference between our approach and van Kampen's system size expansion is that a subset of species numbers is averaged in terms of other species numbers which appear in the limiting model for $\gamma = 2$ due to the various scales involved.

Our estimates of the error is also different from diffusion approximations. In the diffusion approximations, the reaction terms centered by their propensities in the stochastic equations for discrete variables of species numbers are approximated in terms of time-changed Brownian motion. On the other hand, the noise term

in the error estimates is determined by both the centered reaction terms in the equations for discrete variables and a difference between the discrete variables for the normalized species number and their continuous limit.

To find the asymptotic order of magnitude of $Z_i^{N,2}(t) - Z_i^2(t)$, we show convergence of $r_N \left(Z_i^{N,2}(t) - Z_i^2(t) \right)$ to a nonzero finite limit for some r_N. Among the species S_{23}, S_4, S_5, S_8, and S_9, the species number of S_{23} is scaled with the smallest exponent, and thus noise in the limit of $r_N \left(Z_i^{N,2}(t) - Z_i^2(t) \right)$ is determined dominantly by the component $r_N \left(Z_{23}^{N,2}(t) - Z_{23}^2(t) \right)$. Since $Z_{23}^{N,2}(t)$ is the species number scaled by N, we expect that $r_N = N^{1/2}$ and the error between the scaled species numbers and their limit is of order $N_0^{-1/2}$. For a detailed approach to derive r_N and $U(t)$, see more about the central limit theorem in [14]. The fact that all components but the first one in the diffusion term in the equation for $U(t)$ are zero supports the idea that noise is dominantly determined by the error between $Z_{23}^{N,2}(t)$ and $Z_{23}^2(t)$. A sketch of the proof of Remark 3 is given in the Additional file 1: Section 6.

Conclusions

We considered a stochastic model for a well-stirred biochemical network with small numbers of molecules for some species. As the biochemical network consists of more species and reactions, network topology becomes more complex and it is harder to analyze. Therefore, how to reduce the biochemical network while preserving its important biochemical features is a very important issue.

In this paper, we applied the multiscale approximation method introduced by Ball et al. [8] and extended by Kang and Kurtz [9] to a heat shock response model of *E. coli* developed by Srivastava et al. [11]. Using the fact that the species numbers and the reaction rate constants in the model vary over several orders of magnitude, we scaled them using a scaling parameter with different exponents both of which contribute to determining the time scales of species. We derived balance conditions for each species and for a subset of linear combinations of species explicitly in this model, and chose appropriate values for the scaling exponents satisfying the balance conditions. Assuming that initial values of the species numbers are positive, satisfying the balance conditions is required to get a nondegenerate limiting model. We assumed that the reaction rate constants do not change in time, while we may use several sets of scaling exponents for the species numbers due to rapid changes in some species numbers in time. In this analysis, we chose three sets of scaling exponents, and they are used to derive limiting models in different time scales.

In each time scale we derived a limiting model, and used it to approximate the species numbers in the full network. In the limiting model, species numbers whose scaling exponents are larger than those of all rates of reactions involving the species are treated as constants, since changes of the species numbers due to the reactions are not noticeable at these times. When the scaling exponent of the species number is smaller than the scaling exponents of the rates of some productions and consumptions of the species and in case the scaling exponents for both kinds of reactions are equal, the scaled species number is averaged out and is approximated in terms of other variables. Therefore, the limiting model includes a subset of species and reactions and network topology in it becomes simpler. We derived the conditional equilibrium distributions of the fast-fluctuating species numbers and studied errors between the scaled species numbers and their limits in the third time scale.

Using the limiting models, we approximated the temporal evolution of species numbers in three time scales. By comparing stochastic simulation of the full model and approximations using the limiting models, we see that the main features of evolution of species numbers are well captured by the limiting models.

Additional file

Additional file 1: Supplementary material for "A multiscale approximation in a heat shock response model of E. coli." This is a supplementary material of the paper including calculations and tables.

References
1. Kærn M, Elston T, Blakem W, Collins J: **Stochasticity in gene expression: from theories to phenotypes.** *Nat Rev Genet* 2005, **6**(6):451–464.
2. Gillespie D: **A general method for numerically simulating the stochastic time evolution of coupled chemical reactions.** *J Comput Phys* 1976, **22**(4):403–434.
3. Gillespie D: **Exact stochastic simulation of coupled chemical reactions.** *J Phys Chem* 1977, **81**(25):2340–2361.
4. Rao C, Arkin A: **Stochastic chemical kinetics and the quasi-steady-state assumption: application to the Gillespie algorithm.** *J Chem Phys* 2003, **118**(11):4999–5010.
5. Haseltine E, Rawlings J: **Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics.** *J Chemi Phys* 2002, **117**(15):6959–6969.
6. Cao Y, Gillespie D, Petzold L: **Multiscale stochastic simulation algorithm with stochastic partial equilibrium assumption for chemically reacting systems.** *J Comput Phys* 2005, **206**(2): 395–411.
7. Pahle J: **Biochemical simulations: stochastic, approximate stochastic and hybrid approaches.** *Briefings Bioinf* 2009, **10**(1):53–64.
8. Ball K, Kurtz T, Popovic L, Rempala G: **Asymptotic analysis of multiscale approximations to reaction networks.** *Ann Appl Probability* 2006, **16**(4):1925–1961.
9. Kang HW, Kurtz T: **Separation of time-scales and model reduction for stochastic reaction networks.** 2012. arXiv preprint arXiv:1011.1672, to appear in Annals of Applied Probability.
10. Crudu A, Debussche A, Radulescu O: **Hybrid stochastic simplifications for multiscale gene networks.** *BMC Syst Biol* 2009, **3**:89.
11. Srivastava R, Peterson M, Bentley W: **Stochastic kinetic analysis of the Escherichia coli stress circuit using σ^{32}-targeted antisense.** *Biotechnol Bioeng* 2001, **75**(1):120–129.
12. Takahashi K, Kaizu K, Hu B, Tomita M: **A multi-algorithm, multi-timescale method for cell simulation.** *Bioinformatics* 2004, **20**(4):538–546.
13. Weinan E, Vanden-Eijnden E: **Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates.** *J Chem Phys* 2005, **123**(19):194107.

14. Kang HW, Popovic L, Kurtz T: **Central limit theorems and diffusion approximations for multiscale Markov chain models.** 2012. arXiv preprint arXiv:1208.3783, submitted.
15. Kurtz T: **Averaging for martingale problems and stochastic approximation.** *Appl Stochastic Anal* 1992:186–209.
16. Van Kampen NG: *Stochastic processes in physics and chemistry (North-Holland Personal Library)*: Elsevier; 2007.