A network partition method for solving large-scale complex nonlinear processes

Shucheng Pan, Jianhang Wang, Xiangyu Hu, and Nikolaus A. Adams

1 Department of Mechanical Engineering, Technical University of Munich, 85748 Garching, Germany

(Dated: January 22, 2018)

Abstract

A numerical framework based on network partition and operator splitting is developed to solve nonlinear differential equations of large-scale dynamic processes encountered in physics, chemistry and biology. Under the assumption that those dynamic processes can be characterized by sparse networks, we minimize the number of splitting for constructing subproblems by network partition. Then the numerical simulation of the original system is simplified by solving a small number of subproblems, with each containing uncorrelated elementary processes. In this way, numerical difficulties of conventional methods encountered in large-scale systems such as numerical instability, negative solutions, and convergence issue are avoided. In addition, parallel simulations for each subproblem can be achieved, which is beneficial for large-scale systems. Examples with complex underlying nonlinear processes, including chemical reactions and reaction-diffusion on networks, demonstrate that this method generates convergent solution in a efficient and robust way.
I. INTRODUCTION

Solving nonlinear dynamic processes which are ubiquitous in a broad range of physical, biological, and social systems is numerically challenging when the scale and complexity become large. The high accuracy, efficiency and robustness of classic numerical methods exhibited in small-scale simple nonlinear processes are deteriorated when applied to large complex systems, of which chemical reaction is a canonical example where multiple reacting species are governed by complex kinematics with a large number of reactions. The vastly disparate timescale of reactions leads to high stiffness of the system [30]. For instance, a typical chemical reaction mechanism in combustion consists of various timescales spanning from e.g. $10^{-4}$s to $10^{-12}$s [15]. This prohibits the use of explicit numerical methods such as Euler scheme and Runge-Kutta schemes, as the timestep required by stability condition is about $10^{-12}$ which is several order smaller than the timescale of turbulent mixing [15]. Thus in the combustion community, implicit methods such as ref. [4] are commonly employed to handle the reaction terms with large timesteps due to the numerical robustness. However those methods are only efficient for simple kinetic mechanisms as the operation complexity scales cubically with the the size of chemical kinetics [25]. For example, even for a combustion problem with simple methane chemistry (35 species and 217 reactions), more than 90% of the overall computational time is spent on chemistry calculations [33]. Thus numerical simulations of practical combustion problems with detailed chemical mechanism remain to be challenging due to the high stiffness and large number of species and reactions (in the order of $\mathcal{O}(10^3)$) [25], especially in three dimensions [10]. Using quasi-steady-state (QSS) approximation to handle fast reactions eliminate the stiffness of the system but this treatment the violates mass conservation properties and requires ad hoc determination of QSS species [25].

Another example is the dynamic processes, say reaction-diffusion, on a complex network [1 7 27] which have been widely used to represent behaviours in diseases spreading [2], protein-protein interaction [32], and regulation of gene expression [18]. Unlike the nonlinear processes under the homogeneous assumption of the substrate, which can be easily solved in regular lattice, the high heterogeneity and large-scale of complex networks present challenges for numerical solving the underlying nonlinear dynamics [27]. In this case, the implicit numerical methods are difficult to be applied due to the heterogeneous structures and thus
timestep is chosen to be small when the diffusion rate is large, leading to large computational cost.

Instead of solving the complicated system directly, operator splitting, such as Lie splitting scheme [22], decomposes the original problem into smaller subproblems based on different mathematical and physical properties, with each individual easily solved by dedicated methods. Compared to classic non-splitting numerical methods, the operator splitting method is more efficiency and easier to converge. It requires less amounts of memory and offers flexibility to select suitable discretized schemes as well [13]. If the subproblems are solved suitably, usually by using local implicit methods, numerically solving the original problems by the operator splitting becomes unconditionally stable. Many well-established numerical methods have been developed based on the operator splitting concept to solve different physical and mathematical problems [13]. One example is the projection method [6] which separately solves the velocity and the pressure field of the incompressible Navier-Stokes equations. Another one is the split Bregman method [14] which is developed to solve optimization problems in imaging processing, compressive sensing, and machine learning.

Canonical applications of splitting methods include the splitting reaction terms from convection terms for reacting flows [21] and splitting of reaction and diffusion terms for reaction-diffusion equations [8]. For a reacting system with a small number of subproblems (operators) $N$, the splitting can be applied to every elementary reactions and the results indicate that this treatment is unconditional stable and preserves conservation property [29]. However, in many complex system, $N$ is large, implying that directly applying existing schemes requires a large number of splitting [23, 29]. This leads to large splitting errors and difficulty to impose parallelization as the $N$ subproblems are solved sequentially. Many large-scale complex dynamic systems exhibit network structures whose nodes represent different spatial elements or physical terms. The topology of those networks, although very complex, is highly sparse in the sense that the number of interactions between different nodes in the network is small. For example, the species or elementary reactions are sparsely coupled in large chemical reacting systems [9]. This motivates the development of an efficient method that utilized sparsity of network structures to solve large-scale nonlinear processes with sparse network structures. The key idea is applying the network partition, which previously is used for e.g. detecting community structure [28] and serves to ensure only uncorrelated nodes being clustered into the same subset here, on numerical solving large-scale time-dependant
dynamic system. We expect the proposed method to be efficient, convergent, and numerical stable for solving complex nonlinear dynamics on a extremely large system. After dwelling on the method and its main features, a variety of numerical examples are tested to demonstrate the main features.

II. THE NETWORK PARTITION FOR A COMPLEX DYNAMIC SYSTEM

Given a large-scale complex system with time-dependent nonlinear dynamics, we first partition the network based on some chosen principals to achieve a small number of subproblems, each is a subset of decoupled processes that can be solved by classic analytical or numerical methods. This is carried out by three steps: (i) abstract the system with a network $\Gamma$ which usually is sparse and (ii) translate $\Gamma$ to a graph $G$ by a specific mapping $f : \Gamma \rightarrow G$ and (iii) partition $G$ and package the nodes to generate a subproblem. Use operator splitting methods, the subproblems are solved sequentially, that is, the solution of previous subproblem is the input of the next one. Here we use an illustrative example corresponding to a chemical reaction in Fig.1 to describe the details.

A. The network of a nonlinear complex process

As shown in Fig.1A, a chemical reaction system with 43 channels exhibits a network structure. In this network, each node (reaction) has multiple interactions with others by affecting the a number of reacting species. We can find a mapping $f : \Gamma \rightarrow G$ that convert the network to a graph $G = (V, E)$, where $V$ and $E$ are the set of nodes (elements) and the set of edges (pairs), respectively. For instance, one can use the definition in Fig.1B where multiple interactions between the same pair of reactions, $(x, y)$, collapse to one edge $e(x, y)$, i.e., the element $a(x, y)$ of the adjacency matrix $A(G)$ is 1 or 0, depending on whether the reactions $x$ ad $y$ evolve common species. The degree of each node $x$, $d(x) = \sum_{y \in V} a(x, y)$, is highly heterogeneous and the distribution of edges shows sparsity, i.e., the number of edges $n(E)$ is considerably smaller than of the complete graph.

Besides, the node in $G$ can be a group of coupled elementary processes which is solvable in the sense that it has analytical solutions or is easy to be numerically solved. For example, multiple reactions in Fig.1 interacted by common product species have analytical solutions
and thus can be considered as a single node in $G$, leading to a mapping different $f^*$. Or one define a group of reactions as a node if it is efficient to be solved by locally applying implicit time-integration methods. Then the network partition is performed to ensure there is no interaction between those small structures if they are grouped into the same subset.

B. Operator splitting and network partition methods

Large-scale systems are too difficult to solve directly, either due to high computational cost or numerical instabilities. Instead, we apply the operator splitting method which separates the original system into a number of small-scale subsystems, with each one being easy to be numerically solved, and thus offering convenience for solving large-scale nonlinear systems. For instance, the reaction system in Fig.1 can be solved by e.g. Lie splitting method, i.e., the $N$ elementary reactions are solved individually. In this way the results are positive and conservative without time-step constraints and omits costly matrix operations [29].

When $N$ is large, the large number of splitting generates significant splitting errors which prevent the use of large time-steps. Another issue is that the operator splitting method requires a sequential updating of all $N$ subproblems. This causality leads to difficulty for parallization which is usually required for solving large-scale problems. To reduce the number of splitting and alleviate the causality of operator splitting methods, we propose a numerical framework that utilizes the sparsity of the network structure. Usually in most of large-scale systems the directly interacting elements are sparse and the elements that are not direct interacted can be solved simultaneously without numerical difficulty. This inspires us to split the original system into $K$ ($K < N$) subsystems where their elements have no direct interaction and solves the dynamics of all elements inside a certain subsystem simultaneously. This leads to a network partition problem: given a graph $G(V, E)$ after mapping from a physical system, say chemical reaction, one cluster its nodes by a partition $S = \bigcup_{k=1}^K S_k$ that satisfies

$$S = \arg \min_s K_s \quad \text{subject to} \quad A_k = 0,$$

(1)

where $k$ indexes the subsets and $A_k$ is the adjacency matrix of the $k$-th subset. This partition can be achieved by a graph coloring algorithm. When the network is mapped to a planar graph, one have $K \leq 4$.

After partitioning the network, the elementary processes clustered into the same subset,
say $S_k$, generate a subproblem which can be solved in parallel, as the causality updating constraint is avoided in $S_k$. Consider $N_k$ elementary irrelevant reactions belong to the subset $S_k$. The $N_k$ processes are solved simultaneously and the updated involved species are chosen to be input of $S_{k+1}$. The entire reacting system is numerically solved by Lie splitting scheme, $X^{n+1} = E_1(\Delta t) \circ E_2(\Delta t) \circ E_3(\Delta t)$, if $K = 3$, where $X \in \mathbb{R}^M$ is the concentration vector and $E_k$ is the time-integration of the subset $S_k$. Note that the network partition is a semi-constructive procedure and $E$ can be prescribed, say the backward Euler scheme. In this paper, we use analytical solution according the type of every elementary reaction.

The splitting error arising from decoupling treatment of the original couple system can be further reduced by some existing strategies [35]. And the order of accuracy, limited to 1st order due to Lie splitting, can be increased by higher order operator splitting schemes [8, 36, 37]. Consider the widely used Strang splitting scheme [36], we can arrange the decoupled subproblems symmetrically and solve them by $X^{n+1} = E_1(\Delta t/2) \circ E_2(\Delta t/2) \circ E_3(\Delta t) \circ E_2(\Delta t/2) \circ E_1(\Delta t/2)$. The adaptive time-stepping techniques can be applied to improve the accuracy based on local errors estimation. Here we use a similar way with ref. [23] to control the local errors. We reduce the time-step if the relative error between two solutions of different level are larger than a tolerance $\varepsilon \Delta t^k$, where $k$ is the order of splitting method and $\varepsilon$ is a small constant. Consider the base time-step is $\Delta t$ and we perform $n_0^0$ full numerical evaluations at the coarsest level. If the refinement level is $\ell$, the time-step can be reduced to $\Delta t/(2^\ell n_0^0)$.

And there exists a large number of partition strategies for large-scale system. Indeed, the optimal partition is the one that minimizes the splitting error, say $\frac{\Delta t^2}{2} [M_1, M_2]X^0$ for $K = 2$, where $[,]$ is the commutator and $M$ is the operator for each subproblem. However this strategy requires applying optimization on-the-fly, which is costly for large $N$. Alternatively, we can find a partition that the most relevant nodes are clustered into the same subset. One can relate the network with a Markov chain and define a metric called “diffusion map” which measures the correlation of different nodes [20]. For a reacting system, a weight matrix that measures the pairwise interaction strength is defined as $w(x, y) = \sum_z \max(\alpha(x)|\nu(x, z) + \mu(x, z)|, \alpha(y)|\nu(y, z) + \mu(y, z)|)$, which assembles the connectivity measure in some reaction mechanism reduction methods [24], where $z$ labels all common species of reactions $x$ and $y$, $\alpha$ is the reaction rate coefficient, $\nu(x, z)$ is the stoichiometric coefficient of reactant $z$ in reaction $x$, and $\mu(x, z)$ is the stoichiometric coefficient of product $z$ in reaction $x$. Then the transition
matrix of the corresponding Markov chain is determined by \( p(x, y) = w(x, y)/\sum_{z \in G} w(x, z) \). And let \( \lambda_l, \psi_l \) and \( \phi_l \) be the eigenvalue, normalized right and left eigenvectors of \( P \) with \( 0 \leq l \leq N - 1 \). The diffusion map \( \Psi_t(x) = [\lambda_1^t \psi_1(x), \lambda_2^t \psi_2(x), \ldots, \lambda_q^t \psi_q(x)]^T \) is introduced in ref. [20] so that one can determine the connectivity of different nodes by measuring their distance in diffusion coordinates, \( D^2_t(x, y) = \| \Psi_t(x) - \Psi_t(y) \|^2 \). Then nodes are clustered into subsets depending on their diffusion distance to the geometric centroids of subsets \( c(S_k) = \sum_{x \in S_k} \phi_0(x) \Psi_t(x) \), where \( \tilde{\phi}_0(S_k) = \sum_{x \in S_k} \phi_0(x) \). This leads to a new partition that minimizing the distortion \( \sum_k \sum_{x \in S_k} \phi_0(x) \| \Psi_t(x) - c(S_k) \|^2 \). For our case, we follow the clustering algorithm in [20] to modify our initial partition, \( S^{(0)} = \bigcup_{k=1}^{K} S_k^{(0)} \) which is generated by graph coloring algorithm. Then the partition becomes

\[
S_k^{(p)} = \left\{ x | k = \arg \min_{l \in [1, K]} \| \Psi_t(x) - c(S_k^{(p-1)}) \|^2 \right\} \quad \text{subject to} \quad A_k = 0
\]

which is solved by constrained k-means algorithm [39] with \( p \) indexing the iteration step. Then the most relevant nodes are grouped into the same subset while maintaining the constraint \( A_k = 0 \). Although this is not the optimal partition, it outperforms the initial partition generated by coloring algorithm in our test cases.

C. Main features

The current method has the following advantages when applied to large-scale complex systems:

1. This method is unconditionally stable, provided the time-integration \( E \) for each subproblem is stable. Thus the chosen time-step can be very large, which offers efficiency for computations of large-scale systems.

2. Negative concentration generation is avoided for systems with large stiffness.

3. The number of splitting, \( K \), is small for large-scale complex systems due to the sparsity feature, indicating a reduced splitting error.

4. The method are well suited for parallel simulations as the elements of every subproblem have no direct correlation.
III. NUMERICAL EXAMPLES

We start with relatively small problems to analyze the convergence and demonstrate the robustness of the current method. Afterwards, we apply the method to solve large-scale systems to show its efficiency.

A. Chemical reactions

Given a well-stirred chemical system with $M$ chemical species which interact through $N$ reactions. If the stochastic effects in this system are neglected and the concentration vector $\mathbf{X} = \{X_1, X_2, \cdots, X_M\}^T$ is assumed to vary continuously in time, the time evolution of the system can be described by the reaction rate equation,

$$\frac{dX_i}{dt} = \sum_{j=1}^{N} (\nu_{ji} - \mu_{ji}) C_j(t),$$

for $1 \leq i \leq M$, where $\mu_{ji}$ and $\nu_{ji}$ are stoichiometric coefficients of reactant $i$ and product $i$ involved in the reaction $j$, respectively. $C_j(t)$ is the rate of reaction $j$ and usually determined by the law of mass action, $C_j(t) = \alpha_j \prod_j X_i(t)^{\mu_{ji}}$, where $\alpha_j$ is the reaction rate coefficient and the index $l$ labels all involved species in the reaction $j$. Usually the number of involved species and the reactions are enormous, which presents challenges when numerically solve the underlying kinematics. We will test our method through different types of chemical reactions, including the constant rate-coefficient reaction, temperature dependent reaction and stochastic reaction.

1. Constant rate-coefficient reaction

We consider the p53-SMAR1 regulatory biochemical network [26] whose mechanism, as detailed in Table II, involves 18 species (proteins and complexes) and 35 elementary reactions. The system is initialized with $\mathbf{X} = \mathbf{0}$ and the constant reaction-rate coefficients are listed in Table II. Although this case is small, its reaction rates of elementary reaction cross a broad range of magnitudes, $O(10^{-5}) \sim O(1)$, indicating high stiffness of this system. We use adaptive time-step mentioned above, where the largest time-step is $\Delta t$ and the refinement level is $\ell = 3$, i.e., the finest time-step is $\Delta t/8$ if the local errors are larger than.
TABLE I: Convergence results of the apoptosis-regulation reaction network [26] with $n_s = 2$ and $\ell = 3$ at $t = 100$ hours by solving the reaction rate equation [3]. The partition, $S^{(0)}$, which is obtained by graph coloring algorithm and listed in Table [IV] is used. The rate of convergence order is listed in parenthesis. The error tolerance is $\epsilon = 0.01\Delta t$.

|          | $\Delta t = 8.0s$ | 4.0s   | 2.0s   | 1.0s   |
|----------|-------------------|--------|--------|--------|
| p53$_p$-A | 83.60             | 47.40(0.8) | 24.80(0.9) | 12.80(1.0) |
| p53$_p$-p300 | 2.420             | 1.284(0.9) | 0.658(1.0) | 0.334(1.0) |
| p53       | 0.920             | 0.485(0.9) | 0.245(1.0) | 0.123(1.0) |
| p300      | 0.222             | 0.114(1.0) | 0.058(1.0) | 0.029(1.0) |
| HDAC1     | 0.119             | 0.062(0.9) | 0.031(1.0) | 0.016(1.0) |
| Mdm2      | 0.077             | 0.038(0.9) | 0.019(1.0) | 0.009(1.0) |
| p53$_p$   | 1.8e-5            | 1.8e-5(0.0) | 1.2e-5(0.6) | 6.5e-6(0.9) |
| SMAR1     | 1.3e-4            | 6.6e-5(1.0) | 3.3e-5(1.0) | 1.7e-5(1.0) |

The ability to use large time-steps demonstrates the robustness of our method. As shown in Fig.2, the time history of the species indicate that the positive $X$ is preserved for a very large time-step. However, for the 2nd-order Runge-Kutta scheme, numerical instability is observed even we use a small time-step $\Delta t = 1$. Negative concentration of species occur due to spurious solutions on the level of the truncation error, as shown in Fig.2D. We compare the result of our method with the original Lie splitting method. The results of the Lie splitting method show larger errors compared to those of the three solutions obtained by our method. This demonstrates that splitting errors are significantly reduced by applying the network partition, as one motivation of developing the present method. Then we test different partition strategies listed in Table [IV] and compare the results with the reference solution ($\Delta t = 1$). As shown in Fig.2, the results of the optimized partition $S^{(p)}$ in Eq.2 show better agreement with the reference than those of graph coloring based partition $S^{(0)}$. 


in Eq.1. If we consider the mapping $f^*$ above, $K$ is reduced to 9, see $S^{(0),*}$ in Table IV. And the numerical results are also improved due to a smaller number of splitting.

2. Temperature dependent reaction

This type of chemical reactions are widely encountered in computer simulations of combustion. The reaction rates are determined by the Arrhenius law, $\alpha^f_j = A T^B \exp(-E_j/T)$ for forward reactions (odd value $j$) and $\alpha^b_j = \alpha^f_{j-1}/\alpha^p_j$ for backward reactions (even value $j$), where

$$\alpha^p_j = \left(\frac{p_0}{RT}\right)^\sum M_i(\nu_{ji} - \mu_{ji}) \exp\left(-\sum M_i W_i \nu_{ji} - \mu_{ji} \frac{h_i}{T} - s_i\right).$$

with $p_0 = 1$ atm. $R$ is the universal gas constant, $T$ is the temperature of the system, and $W_i$ is the molecular weight of the species $i$, see Table XI. The specific internal enthalpy $h_i$ and entropy $s_i$ of the species $i$ are calculated by the equation in Supporting Information.

We select the hydrogen-oxygen reaction with 9 species and 23 reversible reactions (forward and backward). The mechanism and the partition $K = 34$ are listed in Table V and Table VI, respectively. The initial pressure of the $H_2$-air mixture is 1 atm and the molar ratio is $2 : 1 : 3.76$ for $H_2 : O_2 : N_2$. Nitrogen is inert and thus treated as a diluent.

With an initial temperature of $1200K$, the results calculated by constant time-step ($\ell = 1$) $\Delta t = 1.28\mu s$ and $\Delta t = 0.01\mu s$ are shown in Fig.3A. The time evolution of species concentration indicates that numerical instabilities and negative solution are prevented in our method. Before the ignition, the results of very large time-step agree with those of the small time-step. The errors of temperature at $t = 200\mu s$ are measured and the convergence rate shown in Fig.3B is first order, as expected. Then the initial temperature is reduced from $1200K$ to $950K$ to verify the computed ignition delay time which is measured by the maximum $dT/dt$. As shown in Fig.3B, our numerical results are in good agreement with the experimental data [31], even with a large time-step. Using 2nd-order Runge-Kutta scheme with $\Delta t = 0.01\mu s$ produce numerical instability and negative concentration of species. Note that the Lie splitting method can produce similar results as this small system is almost fully coupled, leading to similar number of splitting for network partition ($K = 34$) and the Lie splitting method ($K = 46$).
3. Stochastic reaction

When the molecular populations are relatively small, the dynamic behaviour of the reacting system described by the deterministic differential equation (3) becomes inaccurate. In this case, the stochastic reaction kinetics should be considered [17] and \( X \in \mathbb{N}^M \) becomes the molecular number vector. Here we show how to extend our method to stochastic reaction simulation where analytical solutions are more difficult to find than in deterministic problems above. As mentioned above, our method can prescribe time-integration schemes for each subproblem. Then some existing accelerated approximated stochastic methods, such as the \( \tau \)-leaping method [12], can be employed to simulate a well-stirred system with low molecular number. After partition, the \( \tau \)-leaping formula,

\[
X^k(t + \tau) = X^k(t) + \sum_{l=1}^{N_k} (\nu_l - \mu_l) \mathcal{P}_l \left( a_l(X^k(t)) \tau \right),
\]

is applied for the subset \( S_k \), where \( \tau \) is the leap time and \( X^k \) is the molecular number vector of all species belongs to \( S_k \). The propensity function \( a_l(X^k(t)) \) of the reaction \( l \) is determined by

\[
a_l(X^k(t)) = \alpha_l \prod_i \frac{X_i(t)!}{\mu_{li}! (X_i(t) - \mu_{li})!},
\]

where \( i \) is the index of all reactants for the reaction \( l \). \( \mathcal{P}_l \left( a_l(X^k(t)) \tau \right) \) is a Poisson random variable with mean and variance being \( a_l(X^k(t)) \tau \).

For reactants with small molecular number, the unbounded \( \mathcal{P}_l \) of the \( \tau \)-leaping method may lead to negative solutions [5, 38] which are easily avoided in our method by simply bounding the copy number of reactants as the reaction channels are uncorrelated for reactants in every subset. This treatment is simpler than existing strategies, e.g. the hybrid of \( \tau \)-leaping method and Gillespie’s stochastic simulation algorithm (SSA) [11] in ref.[5], and does not affect the computational efficiency of the \( \tau \)-leaping method.

We consider the LacZ/LacY reaction model [19] with 22 reactions, as shown in Table VII to verify the accuracy present method. This case, inspite its small size (\( N = 22 \)), show distinct sparse network features and can be which are clustered into \( K = 5 \) subsets. Initially, the population of PLac is 1 while others are 0. The numbers of RNAP and Ribosome are updated by \( \mathcal{N}(35V, 3.5^2) \) and \( \mathcal{N}(350V, 35^2) \), where \( V = 1.0 + t/t_g \) is the cell volume and \( t_g = 2100s \) is the cell generation time [19]. And the propensities of the reactions (1, 8, 9, 20)
in Table VII are rescaled by $V$. First, we simulate this system until $t = 300s$ by SSA. Then using this result as the input, 10,000 simulations are performed in the time interval $[300s, 330s]$ by our method and SSA. Mean and standard deviations of molecular numbers are computed. As shown in Fig.4, the predicted trajectories of our results with $\tau = 0.00625s$ agree well with of the SSA results. And no negative molecular number is observed although the number of RbsLacY is very low ($\simeq 1$) where original $\tau$-leaping method easily produces negative solutions [38].

4. Large-scale reactions

In this section we extend the application to large-scale reacting systems to test the efficiency of our method. The type of each elementary reaction is randomly selected from all types in the LacZ/LacY reaction model and initially we set $X = 0$. The reaction rate coefficient is given by $U(10^{-3}, 1)$ and the deterministic equation (3) are solved by our method with the law of mass action. We increase the size of the reaction system from $N = 100$ to $N = 10^6$ and set $M = N$, as shown in Fig.5A.

The errors measured at $t = 100$ show first-order convergence rate in Fig.5B if we use Lie scheme after partition. As expected, a 2nd-order convergence rate is observed if we use Strang scheme. Then we parallize the algorithm and compute the speedup which is the ratio of CPU time for the serial simulation and for the parallel simulations performed on a 12-core desktop. When $N$ is small, the speedup is low as the computational cost of every subset is insufficient. It approaches the expected speedup value (12) when the scale is increasing, indicating the parallelization property of our method and thus substantially save the computational time for large-scale computations.

B. Reaction-diffusion on complex networks

Reaction-diffusion process is the underlying mechanism of many pattern formulations in biological systems. Its behaviour on cellular networks can be used to model early stage morphogenesis [31]. Recently, reaction-diffusion (RD) processes on random networks with size up to $10^3$ nodes show significant difference from the class behaviour [27]. In this section we demonstrate our method can be used to efficiently solve RD processes on large-scale
networks.

Consider a RD system defined on a complex network with $N^*$ nodes and $M$ different species. On each node, the local reactions change the concentrations of every species which are diffusively transported to connected nodes. The dynamic behaviour is described by

$$\frac{dX_{in}}{dt} = \sum_{j=1}^{N} (\nu_{ji} - \mu_{ji})C_j(t) + D_i \sum_{m=1}^{N^*} L_{nm}X_{im},$$

where $1 \leq i \leq M$, $1 \leq n \leq N^*$, $D_i$ is the diffusion rate of the species $i$ and $L_{nm} = A_{nm} - d_n \delta_{nm}$ is the network Laplacian matrix.

The Lie operator splitting method can be used for highly dissipative system in particle simulations of fluid mechanics whose diffusion term is handled by decoupling all fluxes for each particle and updating the solution in a pairwise manner [23]. Here we consider the flux $F_{nm}^i = X_{in} - X_{im}$ as the element and cluster it by the same way for chemical reactions, i.e., any two fluxes of each subset do not flow into or out of the same node. For the species $i$ and the subset $S_k$, the concentrations of node $n$ and $m$ are updated by $X_{in} = X_{in} - D_i \Delta t F_{nm}^i$ and $X_{im} = X_{in} - D_i \Delta t F_{nm}^i$, respectively. Thus positivity of $X$ can be ensured by limiting the flux $F_{nm}^i = \max \left( -F_{im}/(D_i \Delta t), \min \left( F_{nm}^i, F_{in}/(D_i \Delta t) \right) \right)$.

We first test pure diffusion process on a small scale-free network [3] with $N^* = 200$ nodes and a mean degree $d = 12$. The diffusion rate $D = 2.04$ and initially sate is $X = 10 + 10U(-0.01, 0.01)$. As shown in Fig.6A, errors are measured when the system achieves the equilibrium state and exhibit first-order convergence. And our method shows better accuracy than the Lie splitting method. If we increase $D$, the numerical instabiles are observed if we use the 2nd-order Runge-Kutta scheme with large time-steps. This issue is addressed by our method, indicating that our method, like that in ref.[23], can handle highly dissipative system efficiently. Then the Brusselator reaction model [16],

$$\frac{dX_{1n}}{dt} = 1 - 3.9X_{1n} + (X_{1n})^2X_{2n} - X_{1n} + X_{3n}, \quad \frac{dX_{2n}}{dt} = 2.9X_{1n} - (X_{1n})^2X_{2n}, \quad \frac{dX_{3n}}{dt} = X_{1n} - X_{3n},$$

is added for every node. The corresponding steady state is $\bar{X} = \{1, 2.9, 1\}$ and will be destabilized under specific diffusion rate, leading to the Turing instability. A small perturbation 0.01$\bar{X}$ is imposed on $\bar{X}$ initially and the diffusion rates are set as $D_1 = D_2 = 7 \times 10^{-3}$ and $D_3 = 0.161$. The convergence results in Fig.6B exhibit first order rate for $L_1$ and $L_\infty$ norms.
Then we study a large-scale RD system defined on a scale-free network with $10^4$ nodes with $\bar{d} = 10$. The p53-SMAR1 regulatory reaction model is considered for the reaction term. The initial condition of $X$ and the diffusion rate are listed in Table XXX. The 49986 fluxes of this system are clustered into 569 subsets and the partition of the chemical reaction is $S^{(0)}$ in Table IV. The initial condition of $X$ and diffusion rates are listed in Table I. When the diffusion rate is low (0.1% of the rates in Table), the distribution of p53 is highly heterogeneous (Fig.7A) and will be homogenized by large diffusion rate (Fig.7B). The convergence analysis is shown in Fig.7C which indicates our method is 1st-order accurate in this case and more accurate than Lie splitting method.

IV. CONCLUSION

We present a numerical method based on network partition and operator splitting to solve large-scale complex processes. It overcomes numerical difficulties of conventional methods for large-scale problems, including convergence issue, memory overload, conservation, robustness and efficiency. Our method tries to exploit the network structure of large-scale system and utilize the sparsity of it to partition the system into smaller subproblems which are easy for numerical computations. In this way, large computational costs due to stiffness, high dissipation, and large number of involved species for chemical-reaction and reaction-diffusion processes are significantly reduced. This method is easy to be parallelized and shows good convergence properties. A range of applications demonstrate the flexibility, modularity, robustness, and versatility of our methods, indicating that it is suitable for solving complex problems involving a large number of elementary processes.

MATERIALS AND METHODS

The source code of the numerical method in this paper has been uploaded to https://gitlab.com and the mechanism is detailed in Supporting Information.
ACKNOWLEDGMENT

This research is supported by China Scholarship Council (No. 201306290030), National Natural Science Foundation of China (No. 11628206), and Deutsche Forschungsgemeinschaft (HU 1527/6-1).

[1] M. Asllani, J. D. Challenger, F. S. Pavone, L. Sacconi, and D. Fanelli. The theory of pattern formation on directed networks. Nat Commun, 5:4517, 2014.
[2] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani. Multiscale mobility networks and the spatial spreading of infectious diseases. Proc Natl Acad Sci USA, 106(51):21484–21489, 2009.
[3] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. science, 286(5439):509–512, 1999.
[4] P. N. Brown, G. D. Byrne, and A. C. Hindmarsh. Vode: A variable-coefficient ODE solver. SIAM J Sci Comput, 10(5):1038–1051, 1989.
[5] Y. Cao, D. T. Gillespie, and L. R. Petzold. Avoiding negative populations in explicit Poisson tau-leaping. J Chem Phys, 123(5):054104, 2005.
[6] A. J. Chorin. A numerical method for solving incompressible viscous flow problems. J Comput Phys, 135(2):118–125, 1997.
[7] V. Colizza, R. Pastor-Satorras, and A. Vespignani. Reaction–diffusion processes and metapopulation models in heterogeneous networks. Nat Phys, 3(4):276–282, 2007.
[8] S. Descombes. Convergence of a splitting method of high order for reaction-diffusion systems. Math Comput, 70(236):1481–1501, 2001.
[9] P. Deuflhard and U. Nowak. Efficient numerical simulation and identification of large chemical reaction systems. Ber Bunsenges Phys Chem, 90(11):940–946, 1986.
[10] F. Diegelmann, S. Hickel, and N. A. Adams. Three-dimensional reacting shock–bubble interaction. Combust Flame, 181:300–314, 2017.
[11] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. J Phys Chem, 81(25):2340–2361, 1977.
[12] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J Chem Phys*, 115(4):1716–1733, 2001.

[13] R. Glowinski, S. J. Osher, and W. Yin. *Splitting Methods in Communication, Imaging, Science, and Engineering*. Springer, 2017.

[14] T. Goldstein and S. Osher. The split bregman method for L1-regularized problems. *SIAM J Imaging Sci*, 2(2):323–343, 2009.

[15] X. Gou, W. Sun, Z. Chen, and Y. Ju. A dynamic multi-timescale method for combustion modeling with detailed and reduced chemical kinetic mechanisms. *Combust Flame*, 157(6):1111–1121, 2010.

[16] S. Hata, H. Nakao, and A. S. Mikhailov. Dispersal-induced destabilization of metapopulations and oscillatory Turing patterns in ecological networks. *Sci Rep*, 4:3585, 2014.

[17] D. J. Higham. Modeling and simulating chemical reactions. *SIAM Rev*, 50(2):347–368, 2008.

[18] G. Karlebach and R. Shamir. Modelling and analysis of gene regulatory networks. *Nat Rev Mol Cell Biol*, 9(10):770–780, 2008.

[19] A. M. Kierzek. STOCKS: STOChastic Kinetic Simulations of biochemical systems with Gillespie algorithm. *Bioinformatics*, 18(3):470–481, 2002.

[20] S. Lafon and A. B. Lee. Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization. *IEEE Trans Pattern Anal Mach Intell*, 28(9):1393–1403, 2006.

[21] R. J. LeVeque and H. C. Yee. A study of numerical methods for hyperbolic conservation laws with stiff source terms. *J Comput Phys*, 86(1):187–210, 1990.

[22] S. Lie and F. Engel. Theorie der Transformationsgruppen. 1888.

[23] S. Litvinov, M. Ellero, X. Hu, and N. Adams. A splitting scheme for highly dissipative smoothed particle dynamics. *J Comput Phys*, 229(15):5457–5464, 2010.

[24] T. Lu and C. K. Law. A directed relation graph method for mechanism reduction. *Proc Combust Inst*, 30(1):1333–1341, 2005.

[25] T. Lu and C. K. Law. Toward accommodating realistic fuel chemistry in large-scale computations. *Prog Energy Combust Sci*, 35(2):192–215, 2009.

[26] M. Z. Malik, M. J. Alam, R. Ishrat, S. M. Agarwal, and R. B. Singh. Control of apoptosis by SMAR 1. *Mol Biosyst*, 13(2):350–362, 2017.
[27] H. Nakao and A. S. Mikhailov. Turing patterns in network-organized activator-inhibitor systems. Nat Phys, 6(7):544–550, 2010.

[28] M. E. Newman. Communities, modules and large-scale structure in networks. Nat Phys, 8(1):25, 2012.

[29] K. Nguyen, A. Caboussat, and D. Dabdub. Mass conservative, positive definite integrator for atmospheric chemical dynamics. Atmos Environ, 43(40):6287–6295, 2009.

[30] E. S. Oran and J. P. Boris. Numerical simulation of reactive flow. Cambridge University Press, 2005.

[31] H. G. Othmer and L. Scriven. Instability and dynamic pattern in cellular networks. J Theor Biol, 32(3):507–537, 1971.

[32] J.-C. Rain, L. Selig, H. De Reuse, V. Battaglia, C. Reverdy, S. Simon, G. Lenzen, F. Petel, J. Wojcik, V. Schächter, et al. The protein–protein interaction map of helicobacter pylori. Nature, 409(6817):211–215, 2001.

[33] D. A. Schwer, P. Lu, and W. H. Green. An adaptive chemistry approach to modeling complex kinetics in reacting flows. Combust Flame, 133(4):451–465, 2003.

[34] M. Slack and A. Grillo. Investigation of hydrogen-air ignition sensitized by nitric oxide and by nitrogen dioxide. 1977.

[35] R. L. Speth, W. H. Green, S. MacNamara, and G. Strang. Balanced splitting and rebalanced splitting. SIAM J Numer Anal, 51(6):3084–3105, 2013.

[36] G. Strang. On the construction and comparison of difference schemes. SIAM J Numer Anal, 5(3):506–517, 1968.

[37] M. Thalhammer. High-order exponential operator splitting methods for time-dependent Schrödinger equations. SIAM J Numer Anal, 46(4):2022–2038, 2008.

[38] T. Tian and K. Burrage. Binomial leap methods for simulating stochastic chemical kinetics. J Chem Phys, 121(21):10356–10364, 2004.

[39] K. Wagstaff, C. Cardie, S. Rogers, S. Schrödl, et al. Constrained k-means clustering with background knowledge. In Proceedings of the Eighteenth International Conference on Machine Learning, volume 1, pages 577–584, 2001.
FIG. 1: An illustrative example of network partition method for chemical reactions. A reaction system involving 42 channels can be represented by a network $\Gamma$ in (A) and is mapped to a planar graph $G$ (B), where we consider a primitive mapping $f : \Gamma \rightarrow G$ that an edge $e(x, y)$ exists when two reactions $x$ and $y$ has common species. After partition $G$ by graph coloring algorithm with 4 colors (C), we package the reactions with the same color, which are evolved simultaneously without any interaction.
FIG. 2: Time evolution of the p53-SMAR1 regulatory reaction by the Lie and network partition based splitting methods. The parameters for adaptive time-step are $n_s^0 = 2$, $\ell = 3$, and $\epsilon = 0.01\Delta t$. The reference solution is obtained by a small time-step $\Delta t = 1s$ while the colored lines indicate solutions with large time-step $\Delta t = 32s$ and the dashed lines are the solutions of the Lie splitting method. Three different partition strategies, $S^{(0)}$, $S^{(p)}$, and $S^{*,(0)}$, as listed in Table IV are employed. And the insert of Mdm2_p300 and Mdm2_HDAC1 are the results of a 2nd-order Runge-Kutta method.
FIG. 3: Results of the hydrogen-oxygen reaction. The time history of $X$ is plotted in (A) where a small and large time-step, $\Delta t = 0.08\mu s$ (dashed lines) and $\Delta t = 0.01\mu s$ (solid lines), are chosen. No numerical instabilities or negative concentration are observed. The temperature at $t = 200\mu s$ is measured and its error convergence is shown in (B) with $\Delta t$ increasing from $0.01\mu s$ to $0.08\mu s$. The comparison of the ignition delay time between low resolution numerical results ($\Delta t = 0.08\mu s$), high resolution numerical results ($\Delta t = 0.01\mu s$), and experimental data [34] is given in (B).
FIG. 4: Time evolution of means ($\chi$) and standard deviations ($\sigma$) of the species concentration in the LacZ/LacY reaction. The upper, middle, and lower trajectories (lines and symbols) corresponds to the mean ($\chi$) and the bounds $\chi \pm \sigma$, respectively. Numerical results generated by the SSA are considered as the reference solutions and represented by the black solid lines. Starting from the initial condition generated by SSA at $t = 300s$, we perform numerical simulations until $t = 330s$ by the Lie splitting method (symbols $\triangle$) and our network-partition method (symbols $\Delta$). A constant leap time $\tau = 0.00625$ is used and the ensemble average is performed by 10,000 simulations to obtain the mean and standard deviation values. The molecular numbers of RNAP and Ribosome are updated by $\mathcal{N}(35\mathcal{V}, 3.5^2)$ and $\mathcal{N}(350\mathcal{V}, 35^2)$, where $\mathcal{V} = 1.0 + t/t_g$ is the cell volume and $t_g = 2100s$ is the cell generation time. Negative solutions are not observed in any of the 10,000 simulations.
FIG. 5: Numerical results of large-scale randomly generated reaction systems by the Lie splitting and the network-partition based methods. The elementary reactions are randomly chosen from Table VII and reaction rate coefficient is $\alpha = \mathcal{U}(10^{-3}, 1)$. The network structures of reaction systems with $N = 100, 1000$ and $10,000$ and $M = N$ reacting species are shown in (A). The elements belonging to the same subset are represented by the same color, exhibiting clear network structures. The convergence of errors generated by our method are measured at $t = 300s$ and plotted in (B) with $K = 4$ (i) and $K = 64$ (ii and iii), where $K$ is the number of subsets and “id” is the index of reacting species. Combining the Strang splitting scheme with our method achieves second-order convergence rate in (B, iii). (C) The speedup of parallel simulations, defined by the ratio of CPU time between the serial simulation and the parallel simulation performed on a 12-core (Intel Xeon CPU E5-2609 v4, 1.70GHz) desktop, approaches the expected value (12) when the scale of the reaction is increasing from $N = 100$ to $N = 10^6$. 
FIG. 6: Numerical results of reaction-diffusion processes on a small scale-free network with \( N^* = 200 \) nodes and a mean degree \( \bar{d} = 12 \). The error convergence results of pure diffusion are sketched in (A). The nodes are sorted in increasing order of their degree, i.e., larger index value ("id") indicates larger node degree. The results of our method (symbols ◦) are compared to those of the Lie splitting method (symbols ∆). The convergence of \( L_1 \) and \( L_∞ \) errors of a 3-speceis reaction-diffusion are sketched in (B).

FIG. 7: Numerical results of large-scale RD processes. Spatial distribution of p53 for low diffusion rate and high diffusion rate are shown in (A) and (B), respectively. The convergence results (◦) are shown in (C) and compared to the results (∆) of Lie operator splitting.
SUPPORTING INFORMATION (SI)

A. The mechanism of chemical reactions and the partition results

We provide the reaction mechanism for the cases used in the main text. First the p53-SMAR1 regulatory model that has 35 elementary reactions is listed in Table II. Its network is partitioned by \( S^{(0)} \), \( S^{(p)} \) and \( S^{*(0)} \), as shown in Table IV. The hydrogen-oxygen reaction with 9 species and 23 reversible reactions is detailed in Table V and the corresponding partition result is shown in Table VI. The LacZ/LacY reaction model contains 22 reactions and 23 species, as shown in Table VII. The SSA results at \( t = 300s \) are used for initial condition in Table VIII. The 22 elementary reactions are partitioned into 5 subsets in Table IX.

B. Numerical details

The thermochemical properties of the involved species can be obtained by empirical equations for fitting thermodynamic functions as

\[
\frac{c_p}{R} = \sum_{n=-2}^{4} a_{n+3} T^n,
\]

\[
\frac{h}{RT} = -\frac{a_1}{T^2} + a_2 \frac{\ln T}{T} + \sum_{n=0}^{4} \frac{a_{n+3} T^n}{n+1} + \frac{b_1}{T},
\]

\[
\frac{s}{R} = -\frac{a_1}{2T^2} - \frac{2}{T} + a_3 \ln T + \sum_{n=1}^{4} \frac{a_{n+3} T^n}{n} + b_2,
\]

over a wide temperature range, where coefficients \( a_1 \) to \( a_7 \) and \( b_1,2 \) are tabulated in Table XI. For each species, the first row applies for a temperature range from 1000K to 6000 K and second row is used for temperature below 1000K. The temperature \( T \) of the mixture is updated by iteratively solving the thermodynamics relation

\[
h = e + \frac{p}{\rho} = e + \sum_{m=1}^{M} y_m R_m \rho T,
\]

considering the internal energy \( e \) is constant during the adiabatic process of \( \Delta t \).
TABLE II: Reaction mechanism for p53-SMAR1 regulatory biochemical network (part I).

| Id | Elementary Reaction | Reaction Rate Coefficient |
|----|---------------------|---------------------------|
| 1  | p53 + Mdm2_p300 → ∅ | 8.25e-4                   |
| 2  | Mdm2_mRNA → Mdm2_mRNA + Mdm2 | 4.95e-4 |
| 3  | p53 → p53 + Mdm2_mRNA | 1e-4                      |
| 4  | Mdm2_mRNA → ∅       | 1e-4                      |
| 5  | Mdm2 → ∅            | 4.33e-4                   |
| 6  | ∅ → p53             | 0.078                     |
| 7  | p53_Mdm2 → Mdm2     | 8.25e-4                   |
| 8  | p53 + Mdm2 → p53_Mdm2 | 11.55e-4                |
| 9  | p53_Mdm2 → p53 + Mdm2 | 11.55e-6                 |
| 10 | ATM_I → ATM_A       | 1e-4                      |
| 11 | ATM_A → ATM_I       | 5e-4                      |
| 12 | p53 + ATM_A → ATM_A + p53_p | 5e-4 |
| 13 | p53_p → p53         | 0.5                       |
| 14 | p300 → ∅            | 1e-4                      |
| 15 | p53_p + p300 → p53_p_p300 | 1e-4 |
| 16 | p53_p_p300 → p53_A   | 1e-4                      |
| 17 | p53_A + Mdm2_SMAR1_HDAC1 → p53 | 1e-5 |
TABLE III: Reaction mechanism for p53-SMAR1 regulatory biochemical network (part II).

| Id | elementary reaction | reaction rate coefficient |
|----|---------------------|--------------------------|
| 17 | p53_A + Mdm2_SMAR1_HDAC1 → p53 | 1e-5 |
| 18 | Mdm2 + SMAR1 → Mdm2_SMAR1 | 2e-4 |
| 19 | p53_Mdm2 + p300 → p53_Mdm2_p300 | 5e-4 |
| 20 | Mdm2 + p300 → Mdm2_p300 | 5e-4 |
| 21 | p53_Mdm2_p300 → Mdm2 + p53_p_p300 | 1e-4 |
| 22 | HDAC1 → ∅ | 1e-4 |
| 23 | ∅ → p300 | 0.1 |
| 24 | ∅ → HDAC1 | 2e-4 |
| 25 | HDAC1 + Mdm2_SMAR1 → Mdm2_SMAR1_HDAC1 | 1e-4 |
| 26 | ∅ → SMAR1 | 0.08 |
| 27 | SMAR1 → ∅ | 1e-4 |
| 28 | Mdm2_SMAR1 → ∅ | 2e-4 |
| 29 | p53 + SMAR1 → p53_p | 1e-4 |
| 30 | p53_Mdm2 + SMAR1 → p53_Mdm2_SMAR1 | 1e-3 |
| 31 | p53_Mdm2_SMAR1 → p53_p + Mdm2_SMAR1 | 1e-3 |
| 32 | Mdm2 + HDAC1 → Mdm2_HDAC1 | 2e-3 |
| 33 | Mdm2_HDAC1 + p53_A → p53 | 5 |
| 34 | p53_p_p300 + SMAR1 → p53 + SMAR1 | 1e-4 |
| 35 | p300 + SMAR1 → SMAR1 | 0.5 |
| subset $S_1$  | elements of $S^{(0)}$ | elements of $S^{(p)}$ | elements of $S^{*,(0)}$ |
|--------------|------------------------|------------------------|------------------------|
| $S_1$        | $\{1, 2, 10, 14, 16, 22, 26, 28\}$ | $\{1, 24, 35\}$ | $\{1, 2, 10, 14, 16, 22, 26, 28\}$ |
| $S_2$        | $\{3, 5, 11, 15, 24, 27\}$ | $\{3, 5, 16\}$ | $\{3, 5, 11, 15, 24, 27\}$ |
| $S_3$        | $\{4, 6, 7, 23\}$ | $\{6, 31, 32\}$ | $\{4, 6, 7, 23, 25\}$ |
| $S_4$        | $\{8, 31, 35\}$ | $\{4, 8, 22\}$ | $\{8, 31, 35\}$ |
| $S_5$        | $\{9\}$ | $\{9, 15, 27, 28\}$ | $\{9, 17, 21, 34\}$ |
| $S_6$        | $\{12, 18, 19\}$ | $\{12, 20, 30\}$ | $\{12, 18, 19\}$ |
| $S_7$        | $\{13, 20, 30\}$ | $\{7, 10, 13, 26\}$ | $\{13, 20, 30\}$ |
| $S_8$        | $\{17, 21\}$ | $\{17, 18, 23\}$ | $\{29, 32\}$ |
| $S_9$        | $\{29, 32\}$ | $\{19, 25, 29\}$ | $\{33\}$ |
| $S_{10}$     | $\{33\}$ | $\{11, 21, 33\}$ | $\emptyset$ |
| $S_{11}$     | $\{34\}$ | $\{2, 14, 34\}$ | $\emptyset$ |
| Id | elementary reaction | $A$     | $B$     | $E_a$  |
|----|---------------------|---------|---------|--------|
| 1,2| $H + O_2 \iff OH + O$| $1.91e+14$| 0.0     | 16.44  |
| 3,4| $H_2 + O \iff H + OH_2$| $5.08e+04$| 2.67    | 6.292  |
| 5,6| $H_2 + OH \iff H + H_2O$| $2.16e+08$| 1.51    | 3.43   |
| 7,8| $O + H_2O \iff OH + OH$| $2.97e+06$| 2.02    | 13.4   |
| 9,10*| $H_2 + M \iff H + H + M$| $4.57e+19$| -1.4    | 105.1  |
| 11,12*| $O + O + M \iff O_2 + M$| $6.17e+15$| -0.5    | 0.0    |
| 13,14*| $H + O + M \iff OH + M$| $4.72e+18$| -1.0    | 0.0    |
| 15,16**| $H + OH + M \iff H_2O + M$| $4.50e+22$| -2.0    | 0.0    |
| 17,18***| $H + O_2 + M \iff HO_2 + M$| $3.48e+16$| -0.41   | -1.12  |
| 19,20***| $H + O_2 \iff HO_2$| $1.48e+12$| 0.60    | 0.0    |
| 21,22| $H + HO_2 \iff H_2 + O_2$| $1.66e+13$| 0.0     | 0.82   |
| 23,24| $H + HO_2 \iff OH + OH$| $7.08e+13$| 0.0     | 0.3    |
| 25,26| $HO_2 + O \iff OH + O_2$| $3.25e+13$| 0.0     | 0.0    |
| 27,28| $OH + HO_2 \iff H_2O + O_2$| $2.89e+13$| 0.0     | -0.5   |
| 29,30| $HO_2 + HO_2 \iff H_2O_2 + O_2$| $4.20e+14$| 0.0     | 11.98  |
| 31,32| $HO_2 + HO_2 \iff H_2O_2 + O_2$| $1.30e+11$| 0.0     | -1.629 |
| 33,34*| $H_2O_2 + M \iff OH + OH + M$| $1.27e+17$| 0.0     | 45.5   |
| 35,36*| $H_2O_2 \iff OH + OH$| $2.95e+14$| 0.0     | 48.4   |
| 37,38| $H_2O_2 + H \iff H_2O + OH$| $2.41e+13$| 0.0     | 3.97   |
| 39,40| $H_2O_2 + H \iff H_2 + HO_2$| $6.03e+13$| 0.0     | 7.95   |
| 41,42| $H_2O_2 + O \iff OH + HO_2$| $9.55e+06$| 2.0     | 3.97   |
| 43,44| $H_2O_2 + OH \iff H_2O + HO_2$| $1.00e+12$| 0.0     | 0.0    |
| 45,46| $H_2O_2 + OH \iff H_2O + HO_2$| $5.80e+14$| 0.0     | 9.56   |

*Third-body collision coefficients (default value is 1.0) in reactions with M: * $H_2O = 12.0$, $H_2 = 2.5$; ** $H_2O = 12.0$, $H_2 = 0.73$; ***$H_2O = 14.0$, $H_2 = 1.3$. 

TABLE V: H2-air mixture reaction mechanism for combustion.
TABLE VI: Partition for the hydrogen-oxygen reaction.

| $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_5$ | $S_6$ | $S_7$ |
|-------|-------|-------|-------|-------|-------|-------|
| elements {1} | {2} | {3, 29} | {4, 30} | {5, 11} | {6, 12} | {7, 9, 31} |
| $S_8$ | $S_9$ | $S_{10}$ | $S_{11}$ | $S_{12}$ | $S_{13}$ | $S_{14}$ |
| elements {8, 32} | {13} | {14} | {15} | {16} | {17, 33} | {18, 34} |
| $S_{15}$ | $S_{16}$ | $S_{17}$ | $S_{18}$ | $S_{19}$ | $S_{20}$ | $S_{21}$ |
| elements {19, 35} | {20, 36} | {21} | {22} | {23} | {24} | {25} |
| $S_{22}$ | $S_{23}$ | $S_{24}$ | $S_{25}$ | $S_{26}$ | $S_{27}$ | $S_{28}$ |
| elements {26} | {27} | {28} | {37} | {38} | {39} | {40} |
| $S_{29}$ | $S_{30}$ | $S_{31}$ | $S_{32}$ | $S_{33}$ | $S_{34}$ |
| elements {41} | {42} | {43} | {44} | {45} | {46} |
| id | elementary reaction                                                                 | reaction rate coefficient |
|----|-------------------------------------------------------------------------------------|---------------------------|
| 1  | PLac + RNAP → PLacRNAP                                                              | 0.17                      |
| 2  | PLacRNAP → PLac + RNAP                                                              | 10                        |
| 3  | PLacRNAP → TrLacZ1                                                                  | 1                         |
| 4  | TrLacZ1 → RbsLacZ + PLac + TrLacZ2                                                 | 1                         |
| 5  | TrLacZ2 → TrLacY2                                                                   | 0.015                     |
| 6  | TrLacY1 → RbsLacY + TrLacY2                                                         | 1                         |
| 7  | TrLacY2 → RNAP                                                                      | 0.36                      |
| 8  | Ribosome + RbsLacZ → RbsribosomeLacZ                                                | 0.17                      |
| 9  | Ribosome + RbsLacY → RbsribosomeLacY                                                | 0.17                      |
| 10 | RbsribosomeLacZ → Ribosome + RbsLacZ                                                | 0.45                      |
| 11 | RbsribosomeLacY → Ribosome + RbsLacY                                                | 0.45                      |
| 12 | RbsribosomeLacZ → TrRbsLacZ + RbsLacZ                                              | 0.4                       |
| 13 | RbsribosomeLacY → TrRbsLacY + RbsLacY                                              | 0.4                       |
| 14 | TrRbsLacZ → LacZ                                                                    | 0.015                     |
| 15 | TrRbsLacZ → LacY                                                                    | 0.036                     |
| 16 | LacZ → dgrLacZ                                                                      | 6.42e-5                   |
| 17 | LacY → dgrLacY                                                                      | 6.42e-5                   |
| 18 | RbsLacZ → dgrLacY                                                                   | 0.3                       |
| 19 | RbsLacZ → dgrRbsLacY                                                                | 0.3                       |
| 20 | LacZ + lactose → LacZlactose                                                         | 9.52e-5                   |
| 21 | LacZlactose → product + LacZ                                                        | 431                       |
| 22 | LacY → lactose + LacY                                                                | 14                        |
TABLE VIII: Initial population of reacting species of the LacZ/LacY reaction.

| id | reacting species               | number of molecules |
|----|--------------------------------|---------------------|
| 1  | PLac                           | 0                   |
| 2  | RNAP                           | 40                  |
| 3  | PLacRNAP                       | 1                   |
| 4  | TrLacZ1                        | 0                   |
| 5  | TrLacZ2                        | 15                  |
| 6  | TrLacY1                        | 0                   |
| 7  | TrLacY2                        | 1                   |
| 8  | RbsLacZ                        | 0                   |
| 9  | RbsLacY                        | 1                   |
| 10 | Ribosome                       | 471                 |
| 11 | RbsRibosomeLacZ                | 38                  |
| 12 | RbsRibosomeLacY                | 35                  |
| 13 | TrRbsLacZ                      | 883                 |
| 14 | TrRbsLacY                      | 332                 |
| 15 | LacZ                           | 1880                |
| 16 | LacY                           | 1608                |
| 17 | dgrLacZ                        | 15                  |
| 18 | dgrLacY                        | 12                  |
| 19 | dgrRbsLacZ                     | 37                  |
| 20 | dgrRbsLacY                     | 24                  |
| 21 | lactose                        | 141918              |
| 22 | LacZlactose                    | 48                  |
| 23 | product                        | 1673873             |
TABLE IX: Partition for the LacZ/LacY reaction.

| subset elements | elements          |
|-----------------|-------------------|
| $S_1$           | {1, 5, 8, 13, 14, 17} |
| $S_2$           | {2, 6, 10, 15, 16}  |
| $S_3$           | {3, 7, 9, 12, 20}   |
| $S_4$           | {4, 11, 21, 22}     |
| $S_5$           | {18, 19}           |

TABLE X: Initial population and diffusion rate of a large-scale reaction-diffusion system.

| id | reacting species | number of molecules | diffusion rate |
|----|------------------|---------------------|----------------|
| 1  | p53              | 38.3355             | 0.5648         |
| 2  | Mdm2             | 1.81548             | 0.3727         |
| 3  | Mdm2_mRNA        | 38.1101             | 0.3465         |
| 4  | p53_Mdm2         | 15.0388             | 0.9453         |
| 5  | ATM_I            | 0                   | 0.4528         |
| 6  | ATM_A            | 0                   | 0.0801         |
| 7  | p53P             | 0.000191227         | 0.1599         |
| 8  | p300             | 8.99129             | 0.007136       |
| 9  | p53p_p300        | 671.386             | 0.0483         |
| 10 | p53_A            | 18465.5             | 0.6786         |
| 11 | HDAC1            | 2.66895             | 5.491          |
| 12 | Mdm2_SMAR1_HDAC1 | 0.000243163         | 0.0795         |
| 13 | p53_Mdm2_p300    | 675.173             | 0.1973         |
| 14 | Mdm2_p300        | 0.257587            | 0.7926         |
| 15 | SMAR1            | 0.00516598          | 0.08357        |
| 16 | Mdm2_SMAR1       | 0.170117            | 0.6910         |
| 17 | p53_Mdm2_SMAR1   | 0.0777557           | 0.1425         |
| 18 | Mdm2_HDAC1       | 0                   | 0.1439         |
**TABLE XI**: Thermochemical coefficients of the species in the H2-air mixture reaction (part I).

| Species | $a_1$       | $a_2$       | $a_3$       | $a_4$       | $a_5$       |
|---------|-------------|-------------|-------------|-------------|-------------|
| H       | 0.000000000 | 0.000000000 | 2.500000000 | 0.000000000 | 0.000000000 |
|         | 6.078774250e1 | -1.819354417e1 | 2.500211817 | -1.226512864e-7 | 3.732876330e-11 |
| O       | -7.953611300e3 | 1.607177787e2 | 1.966226438 | 1.013670310e-3 | -1.110415423e-6 |
|         | 2.619020262e5 | -7.298722030e2 | 3.317177270 | -4.281334360e-4 | 1.036104594e-7 |
| H2O     | -3.947960830e4 | 5.755731020e2 | 9.317826530e-1 | 7.222712860e-3 | -7.342557370e-6 |
|         | 1.034972096e6 | -2.412698562e3 | 4.646110780 | 2.291998307e-3 | -6.836830480e-7 |
| OH      | -1.998589900e3 | 9.300136160e1 | 3.050854229 | 1.529529288e-3 | -3.157890998e-6 |
|         | 1.017393379e6 | -2.509957276e3 | 5.116547860 | 1.305299930e-4 | -8.284322260e-8 |
| O2      | -3.425563420e4 | 4.847000970e2 | 1.119010961 | 4.293889240e-3 | -6.836300520e-7 |
|         | -1.037939022e6 | 2.344830282e3 | 1.819732036 | 1.267847582e-3 | -2.188067988e-7 |
| H2      | 4.078323210e4 | -8.009186040e2 | 8.214702010 | -1.269714457e-2 | 1.753605076e-5 |
|         | 5.608128010e5 | -8.371504740e2 | 2.975364532 | 1.252249124e-3 | -3.74071690e-7 |
| H2O2    | -9.279533580e4 | 1.564748385e3 | -5.976460140 | 3.270744520e-2 | -3.932193260e-5 |
|         | 1.489428027e6 | -5.170821780e3 | 1.128204970e1 | -8.042397790e-5 | -1.818383769e-8 |
| HO2     | -7.598882540e4 | 1.329383918e3 | -4.677388240 | 2.508308202e-2 | -3.006551588e-5 |
|         | -1.810669724e6 | 4.963192030e3 | -1.039498992 | 4.560148530e-3 | -1.061859447e-6 |
| N2      | 2.210371497e4 | -3.818461820e2 | 6.08273836 | -8.530914410e-3 | 1.384646189e-5 |
|         | 5.877124060e5 | -2.239249073e3 | 6.06694922 | -6.139685500e-4 | 1.491806679e-7 |
TABLE XII: Thermochemical coefficients of the species in the H2-air mixture reaction (part II).

| species | \(a_6\) | \(a_7\) | \(b_1\) | \(b_2\) |
|---------|---------|---------|---------|---------|
| H       | 0.000000000 | 0.000000000 | 2.547370801e4 | -4.466828530e-1 |
|         | -5.687744560e-15 | 3.410210197e-19 | 2.547486398e+4 | -4.481917770e-1 |
| O       | 6.517507500e-10 | -1.584779251e-13 | 2.840362437e+4 | 8.404241820 |
|         | -9.438304330e-12 | 2.725038297e-16 | 3.392428060e+4 | -6.679585350e-1 |
| H2O     | 4.955043490e-09 | -1.336933246e-12 | -3.303974310e+4 | 1.724205775e+1 |
|         | 9.426468930e-11 | -4.822380530e-15 | -1.384286509e+4 | -7.978148510 |
| OH      | 3.315446180e-9 | -1.138762683e-12 | 2.991214235e+3 | 4.674110790 |
|         | 2.006475941e-11 | -1.556993656e-15 | 2.019640206e+4 | -1.101282337e1 |
| O2      | -2.023372700e-9 | 1.039040018e-12 | -3.391454870e+3 | 1.849699470e1 |
|         | 2.053719572e-11 | -8.193467050e-16 | -1.689010929e+4 | 1.738716506e1 |
| H       | -1.202860270e-8 | 3.368093490e-12 | 2.682484665e+3 | -3.043788844e1 |
|         | 5.936625200e-11 | -3.606994100e+4 | 5.339824410e+3 | -2.202774769 |
| H2O2    | 2.509255235e-8 | -6.465045290e-12 | -2.494004728e+4 | 5.877174180e1 |
|         | 6.947265590e-12 | -4.827831900e-16 | 1.418251038e+4 | -4.650855660e1 |
| HO2     | 1.895600056e-8 | -4.828567390e-12 | -5.873350960e+3 | 5.193602140e1 |
|         | 1.144567878e-10 | -4.763064160e-15 | -3.200817190e+4 | 4.066850920e1 |
| N2      | -9.625793620e-9 | 2.519705809e-12 | 7.108460860e+2 | -1.076003744e1 |
|         | -1.923105485e-11 | 1.061954386e-15 | 1.283210415e+4 | -1.586640027e1 |