

## ANALYSIS OF THE MULTIPLANE METHOD FOR STOCHASTIC SIMULATIONS OF REACTION NETWORKS WITH FLUCTUATIONS\*

BARUCH BARZEL<sup>†</sup>, OFER BIHAM<sup>†</sup>, AND RAZ KUPFERMAN<sup>‡</sup>

**Abstract.** Chemical reaction networks are commonly modeled by rate equations, which are systems of ordinary differential equations describing the evolution of species concentrations. Such models break down at low concentrations, where stochastic effects become dominant. Instead, one has to solve the master equation that governs the multidimensional probability distribution of particle populations. For large networks such an approach is often computationally prohibitive due to the exponential dependence of the number of states on the number of components. The multiplane method is a dimension reduction technique that exploits the structure of the network to derive approximate dynamics for the marginal distributions of pairs of coreacting species. This method was introduced in [A. Lipshtat and O. Biham, *Phys. Rev. Lett.*, 93 (2004), 170601] as an uncontrolled approximation for specific examples in the context of interstellar chemical reactions. In this paper we formalize the method and prove that it is asymptotically exact in the two extreme limits of small and large population sizes. Our analysis concentrates on steady-state conditions, although numerical simulations indicate that the method is equally well applicable to time-dependent solutions. This analysis partially explains the surprisingly high accuracy of the method.

**Key words.** reaction networks, master equation, dimension reduction, multiplane method, asymptotic expansion

**AMS subject classifications.** 60J22, 41A60

**DOI.** 10.1137/070685245

**1. Introduction.** Reaction networks are a modeling framework used in many branches of science, such as chemical reactions, molecular genetics, and population dynamics. A reaction network can be represented as a directed hypergraph, where the nodes represent the various species and the hyperedges represent the reactions. A directed hyperedge connects a set of source nodes—the *reactants*—to a set of target nodes—the *products*. Each reaction has a number associated with it—a rate constant. The semantics of such networks vary between applications. Most commonly, reaction networks are interpreted as *rate equations*, which are systems of ordinary differential equations that govern the evolution of species concentrations. Rate equations are accurate in situations where the population sizes of the various species are large, and stochastic effects are therefore negligible. Rate equations, however, become inappropriate in situations where the typical population size of one or more species is of order one or less. In such situations, reaction networks are usually interpreted as *continuous-time Markov processes*, where transitions between states occur at random times, and in every such event a single reaction occurs.

While sample paths of continuous-time Markov processes can be easily simulated on a computer, the inference of their statistical properties (e.g., the prediction of mean populations or mean reaction rates) requires the solution of a *master equation*, which

---

\*Received by the editors March 14, 2007; accepted for publication (in revised form) July 11, 2007; published electronically December 13, 2007.

<http://www.siam.org/journals/mms/6-3/68524.html>

<sup>†</sup>Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel (baruch@phys.huji.ac.il, biham@phys.huji.ac.il). These authors were partially supported by the Israel Science Foundation and the Adler Foundation for Space Research.

<sup>‡</sup>Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel (raz@math.huji.ac.il). This author was partially supported by the Israel Science Foundation.

is an infinite-dimensional linear differential system that governs the evolution of the joint probability density of species populations. Even if this system is truncated into a finite-dimensional one, its dimension can remain very large, especially when the number of reacting species is large; the number of states is exponential in the number of species. This “explosion of states” over the years has motivated the development of numerous analytical and computational methods to approximate the solution of the master equation by the solution of a lower-dimensional equation. In certain cases, e.g., when an explicit separation of scales is identified, the lower-dimensional approximation can be justified rigorously and error bounds can be provided. In most situations, however, the approximation is “uncontrolled” and justified a posteriori by numerical experiments; see [10] for a review of dimension reduction techniques in similar and other contexts.

It should be noted that in many cases the master equation is not used directly due to its complexity. Instead, stochastic realizations are performed following the Gillespie algorithm [9]. This approach is used, for example, by Charnley [6] in the context of interstellar reactions. Other attempts to reduce the complexity of the system is the use of moment equations. There the issue is the derivation of suitable closure relations. For example, in order to close the equations at the level of second moments, it is necessary to express the third moment in terms of the first two moments. Examples for such closure relations can be found in [18, 11].

In [17, 1] an approximation method was proposed and applied to chemical networks associated with interstellar gas and dust. In interstellar clouds, certain chemical reactions occur on the surface of microscopic dust grains [13]. When the grains are very small, the mean populations of some reactive species are often less than one atom/molecule per grain, which renders the use of rate equations inappropriate [7, 5]. On larger grains, mean populations are intermediate in size, often not large enough to justify the use of the rate equations, and yet sufficiently large to render the truncated master equation of too high dimension to be solved by direct methods [20]. The approximation method developed in [17] relies on the premise that species that do not directly interact are in some sense conditionally independent so that lower-dimensional master equations can be derived for marginal distributions of subsets of interacting species. This approximation, which was named the *multiplane method*, was applied to a number of model systems. In [17, 1] the multiplane method reduces the multivariate master equation into a set of coupled equations for the marginal distributions of population sizes within maximal cliques of interacting species. In the present paper, the method is simplified even further, resulting in a set of equations for the marginal distributions of only pairs of coreacting species. This new formulation yields an approximation method that is more efficient and does not depend on the structure of the reaction network (e.g., does not require the network to be sparse).

The multiplane method was tested for a number of specific networks and was found to provide approximations of surprisingly high accuracy over a large range of parameters comprising both small and large grains. The main goal of the present paper is to elucidate the reasons for this success. Our main finding is that the multiplane method is asymptotically exact both in the limit of very small grains and in the limit of very large grains. More precisely, let  $r_{ij}$  denote the mean rate of a reaction that takes as reactants species  $X_i$  and  $X_j$ , and let  $\hat{r}_{ij}$  denote the mean reaction rate predicted by the multiplane method; both exact and approximate rates depend on the size  $s$  of the grain. We find that the *relative error* of the multiplane method has the following two-sided asymptotic behavior:

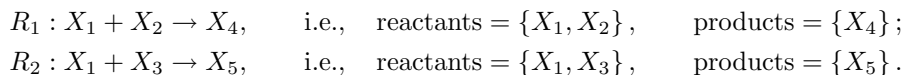
$$(1.1) \quad \frac{\hat{r}_{ij} - r_{ij}}{r_{ij}} = \begin{cases} O(s), & s \ll 1, \\ O(1/s), & s \gg 1. \end{cases}$$

It should be noted that each of these limits admits a relatively simple solution—rate equations are valid in the large grain limit, whereas a low-dimensional truncated master equation is valid in the small grain limit. The virtue of the multiplane method is that it “interpolates” between these two limiting regimes in which it is exact. We believe that it is this property that makes it relatively accurate even in intermediate regimes, where no analytical justification can be provided. Our asymptotic analysis concentrates on steady-state conditions; numerical evidence indicates that the method applies equally well to time-dependent situations.

The structure of this paper is as follows: In section 2 we define chemical networks and formulate the corresponding master equation. In section 3 we describe the multiplane method as a general algorithm. In section 4 we show in detail how to apply the multiplane method to a particular low-dimensional system and assess its accuracy by a numerical solution of both exact and approximate equations. In sections 5 and 6 we analyze the multiplane method in the limits of small and large grain size, respectively. In particular, we estimate the error for the predicted mean reaction rates and derive the two-sided asymptotic behavior (1.1). A numerical test supporting our analysis is reported in section 7. A discussion concludes this paper in section 8.

**2. Chemical reaction networks.** Consider a chemical network of  $M$  species,  $\mathbf{X} = \{X_i\}_{i=1}^M$ , and  $K$  reactions,  $\mathbf{R} = \{R_j\}_{j=1}^K$ . Every reaction  $R_j$  involves a set of reactants and a set of products. A reaction network can be represented graphically as a directed hypergraph. The nodes in the graph are the species  $\mathbf{X}$ ; there is one node per species. The hyperedges in the graph represent the reactions. The  $j$ th hyperedge connects a set of source nodes—the reactants—to a set of target nodes—the products. This chemical network is supplemented by influxes (e.g., due to the adsorption of molecules onto a surface), which are graphically represented as inward arrows, and by effluxes (e.g., due to the desorption of molecules from a surface), which are graphically represented as outward arrows.

*Example.* An example of such a network is depicted in Figure 2.1. The network involves five species, denoted by  $\mathbf{X} = \{X_1, \dots, X_5\}$ . All species have outgoing fluxes, whereas only the species  $X_1, X_2, X_3$  have incoming fluxes. There are two reactions:



Each reaction  $R_j$  has a rate constant  $a_j$  associated with it; the reaction rate is the product of the rate constant and the population sizes of its reactants (i.e., reaction rates depend combinatorially on the abundance of reactants). In addition, every species has an influx rate  $f_i$  and an efflux/desorption rate constant  $w_i$ ; the efflux of a species is the product of this constant and its population size. This chemical network is assumed to be governed by a *Markov birth-and-death process*, that is, a continuous-time Markov process in which single reactions occur at exponentially distributed random times with rates dictated by the state of the system. Such random events may either be the addition of a single molecule, the removal of a single molecule, or a chemical reaction (see, e.g., Gardiner [8]).

General formulations tend to require tedious notation. Here we will restrict ourselves to binary reactions: every reaction involves two reactants (which may be of the

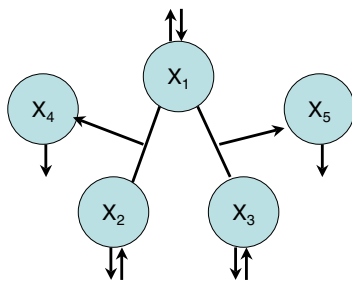


FIG. 2.1. A hypergraph representing a model chemical network. The network involves five species,  $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_4$ , and  $X_5$ , and two reactions,  $X_1 + X_2 \rightarrow X_4$  and  $X_1 + X_3 \rightarrow X_5$ .

same species) and one product. The list of reactions  $\mathbf{R}$  can therefore be represented as a list of triplets:

$$(i, j, k) \in \mathbf{R} \quad \text{means that reaction} \quad X_i + X_j \rightarrow X_k \quad \text{occurs.}$$

We will denote the rate constant of the reaction  $(i, j, k)$  by  $a_{ij}$ ; the reasons for the omission of the index  $k$  are that the reactants uniquely define the product and that the reaction rate depends only on the abundance of reactant molecules.

Networks of binary reactions can be endowed with another graph structure. The set of nodes remains the set of species,  $\mathbf{X}$ . An undirected edge  $E_{ij}$  connects the nodes  $X_i$  and  $X_j$  if these species coreact, i.e., if there exists a product  $X_k$  such that  $(i, j, k) \in \mathbf{R}$ . For reactions  $(i, i, k)$  involving two reacting molecules of the same species, no edge is drawn. We denote the set of edges by  $\mathcal{E}$  and will henceforth call the resulting undirected graph,  $(\mathbf{X}, \mathcal{E})$ , the *reaction network graph*.

*Example.* For the network of Figure 2.1 there are only two edges:  $\mathcal{E} = \{E_{12}, E_{13}\}$ .

Let the population size of the  $i$ th species be denoted by  $N_i$ . The vector of population sizes  $\mathbf{N} = (N_1, \dots, N_M)$  is a stochastic (Markov) process in continuous time. Furthermore, let  $P(\mathbf{n}) = P(n_1, \dots, n_M)$ ,  $n_j \geq 0$ , be the multivariate probability distribution of the population sizes, generally a function of time.

We introduce increment and decrement operators: for functions  $g = g(\mathbf{n})$  we define

$$(S_j^\pm g)(n_1, \dots, n_M) = g(n_1, \dots, n_j \pm 1, \dots, n_M).$$

If  $n_j = 0$ , then the decrement operator yields zero. With a slight abuse of notation, we denote, for example, by  $S_j^+(n_j g)$  the function

$$\mathbf{n} \mapsto (n_j + 1) g(n_1, \dots, n_j + 1, \dots, n_M).$$

Using this abbreviated notation, the *master equation* for the joint distribution  $P(\mathbf{n})$  takes the form

$$(2.1) \quad \frac{dP}{dt} = \sum_{X_i \in \mathbf{X}} f_i (S_i^- - I)P + \sum_{X_i \in \mathbf{X}} w_i (S_i^+ - I)(n_i P) + \sum_{(i,j,k) \in \mathbf{R}} a_{ij} (S_i^+ S_j^+ S_k^- - I)(n_i n_j P).$$

The first term on the right-hand side of (2.1) represents the influx of molecules, the second term represents desorption, and the third term represents the chemical

reactions. Note that  $P(\mathbf{n})$  grows if before the reaction each reactant had a population size greater by one, and each product had a population size smaller by one. For reactions  $(i, i, k)$  where the two reactants are of the same species, the reaction term has to be modified into

$$a_{ii}(S_i^+ S_i^+ S_k^- - I)(n_i(n_i - 1)P).$$

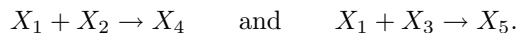
The master equation (2.1) is a linear evolution equation on the infinite-dimensional discrete space,  $\mathbb{N}^M$ . The stationary distribution of this process is the solution of (2.1) with the left-hand side equal to zero; averages with respect to the stationary distribution are denoted by  $\langle \cdot \rangle$ . The rate of the  $(i, j, k)$  reaction at steady state is given by

$$r_{ij} = a_{ij} \langle N_i N_j \rangle.$$

If  $i = j$ , the reaction rate at steady state is

$$r_{ii} = a_{ii} \langle N_i(N_i - 1) \rangle.$$

*Example.* Consider again the network shown in Figure 2.1. In this case we have  $M = 5$  species and  $K = 2$  reactions, given by



The master equation (2.1) for  $P = P(n_1, \dots, n_5)$  is

$$\begin{aligned} \frac{dP}{dt} = & \sum_{i=1}^3 f_i (S_i^- - I)P + \sum_{i=1}^5 w_i (S_i^+ - I)(n_i P) \\ & + a_{12}(S_1^+ S_2^+ S_4^- - I)(n_1 n_2 P) + a_{13}(S_1^+ S_3^+ S_5^- - I)(n_1 n_3 P). \end{aligned}$$

Since the species  $X_4$  and  $X_5$  are only end products, they can be eliminated from the model by summing up the master equation over all possible values of  $n_4$  and  $n_5$  (taking a partial trace). The reduced master equation for  $P = P(n_1, n_2, n_3)$  is

$$(2.2) \quad \begin{aligned} \frac{dP}{dt} = & \sum_{i=1}^3 f_i (S_i^- - I)P + \sum_{i=1}^3 w_i (S_i^+ - I)(n_i P) \\ & + a_{12}(S_1^+ S_2^+ - I)(n_1 n_2 P) + a_{13}(S_1^+ S_3^+ - I)(n_1 n_3 P). \end{aligned}$$

We are interested in chemical reactions that occur in adsorption sites on surfaces. An important parameter is the area of the surface,  $s$  (in interstellar chemistry  $s$  is the surface area of the dust grain). Small grains imply weaker influxes of particles and faster kinetics, as two atoms that diffuse on the grain surface have a higher collision rate. Specifically, the influxes  $f_i$  are directly proportional to  $s$ , whereas the reaction rates  $a_{ij}$  are inversely proportional to  $s$ . The rate of desorption,  $w_i$ , is, on the other hand, independent of the grain size.

Except for very simple situations [12, 3, 4], the master equation cannot be solved analytically, and numerical solutions are needed. The infinite-dimensional master equation is then truncated into a finite-dimensional system. To get meaningful results, the truncation has to be performed at values of  $n_i$  large compared to typical population sizes. For large networks, a numerical solution can still be impractical, as the system size is exponential in the number of species. This is precisely the motivation for approximation methods such as the one described in the next section.

**3. The multiplane method.** Consider a reaction network of the type discussed in the previous section, where every reaction  $(i, j, k) \in \mathbf{R}$  involves two reactants,  $X_i$  and  $X_j$ , and one product,  $X_k$ . Such a network is represented by a reaction network graph,  $(\mathbf{X}, \mathcal{E})$ , where  $E_{ij} \in \mathcal{E}$  if  $X_i$  and  $X_j$  coreact.  $|\mathcal{E}|$  denotes the number of edges in the graph; note that it can differ from the total number of reactions,  $|\mathbf{R}|$ , since reactions where the two reactants are of the same species are not included in  $\mathcal{E}$ . For every  $E_{ij} \in \mathcal{E}$  we denote by  $P_{ij}$  the marginal distribution:

$$P_{ij}(n_1, n_2) = \text{Prob}(N_i = n_1, N_j = n_2).$$

The *multiplane method* replaces the master equation (2.1) by  $|\mathcal{E}|$  coupled master equations whose solutions approximate the marginal distributions  $P_{ij}$ ,  $E_{ij} \in \mathcal{E}$ , associated with all pairs of coreacting species. The rationale behind the method is that the multivariate distribution of population sizes is captured, to a large extent, by the pairwise marginal distributions of coreacting species. The method is intended only to approximate the marginal distributions of pairs of coreacting species; i.e., it is not expected to yield accurate predictions for, say, a moment  $\langle N_i N_j \rangle$  for  $E_{ij} \notin \mathcal{E}$ .

We denote by  $\mathbf{C}_{ij} = \{X_i, X_j\}$ ,  $E_{ij} \in \mathcal{E}$ , the pairs of reacting species and by  $Q_{ij} = Q_{ij}(n_i, n_j)$  the distribution that approximate the marginal distribution  $P_{ij}$  of the population sizes in  $\mathbf{C}_{ij}$ . Note that if a species  $X_i$  reacts with more than one species,  $X_j, X_k, \dots$ , then the approximate distributions  $Q_{ij}, Q_{ik}, \dots$  induce marginal distributions on the population size of  $X_i$ , but these marginals are not a priori required to coincide. Mathematically, this means that we have generated several distinct replicas of  $X_i$ .

For every  $E_{ij} \in \mathcal{E}$ , the approximate master equation for  $Q_{ij}$  is derived as follows: We start with the master equation (2.1) and take its trace over all species  $X_k \notin \mathbf{C}_{ij}$ . When taking a partial trace over the reaction terms, we need to distinguish between various cases, depending on whether one or two of the species involved are in  $\mathbf{C}_{ij}$ , and within each case to distinguish between the reactants and the product. The only case that cancels upon summation is where all three species are not in  $\mathbf{C}_{ij}$ . The resulting equation for  $P_{ij}$  is

$$\begin{aligned} (3.1) \quad \frac{dP_{ij}}{dt} &= f_i(S_i^- - I)P_{ij} + f_j(S_j^- - I)P_{ij} + w_i(S_i^+ - I)(n_i P_{ij}) + w_j(S_j^+ - I)(n_j P_{ij}) \\ &+ a_{ij}(S_i^+ S_j^+ - I)(n_i n_j P_{ij}) \\ &+ \sum_{(i, \ell, m) \in \mathbf{R}} a_{i\ell}(S_i^+ - I)(n_i \langle N_\ell \rangle_{n_i, n_j} P_{ij}) + \sum_{(j, \ell, m) \in \mathbf{R}} a_{j\ell}(S_j^+ - I)(n_j \langle N_\ell \rangle_{n_i, n_j} P_{ij}) \\ &+ \sum_{(i, \ell, j) \in \mathbf{R}} a_{i\ell}(S_i^+ S_j^- - I)(n_i \langle N_\ell \rangle_{n_i, n_j} P_{ij}) + \sum_{(j, \ell, i) \in \mathbf{R}} a_{j\ell}(S_j^+ S_i^- - I)(n_j \langle N_\ell \rangle_{n_i, n_j} P_{ij}) \\ &+ \sum_{(k, \ell, i) \in \mathbf{R}} a_{k\ell}(S_i^- - I)(\langle N_k N_\ell \rangle_{n_i, n_j} P_{ij}) + \sum_{(k, \ell, j) \in \mathbf{R}} a_{k\ell}(S_j^- - I)(\langle N_k N_\ell \rangle_{n_i, n_j} P_{ij}), \end{aligned}$$

where the various summations over  $k, \ell, m$  are over species not in  $\mathbf{C}_{ij}$ . The last four lines represent reaction terms: the first line is for the reaction involving  $X_i$  and  $X_j$  as reactants; the second line represents reactions where either  $X_i$  or  $X_j$  is a reactant and neither is a product; the third line represents reactions where  $X_i$  is a reactant and  $X_j$  a product, and vice versa; the fourth line represents reactions where either  $X_i$  or  $X_j$  is a product and neither is a reactant.

The  $|\mathcal{E}|$  master equations (3.1) do not form a closed system because they contain expressions such as

$$\langle N_\ell \rangle_{n_i, n_j} \quad \text{and} \quad \langle N_k N_\ell \rangle_{n_i, n_j},$$

which are moments of population sizes of species not in  $\mathbf{C}_{ij}$ , conditional on the state of the population sizes in  $\mathbf{C}_{ij}$ . That is,

$$\langle N_\ell \rangle_{n_i, n_j} = \frac{\sum_{n_m, X_m \notin \mathbf{C}_{ij}} n_\ell P(\mathbf{n})}{\sum_{n_m, X_m \notin \mathbf{C}_{ij}} P(\mathbf{n})} \quad \text{and} \quad \langle N_k N_\ell \rangle_{n_i, n_j} = \frac{\sum_{n_m, X_m \notin \mathbf{C}_{ij}} n_k n_\ell P(\mathbf{n})}{\sum_{n_m, X_m \notin \mathbf{C}_{ij}} P(\mathbf{n})},$$

which requires the knowledge of the full multivariate distribution  $P$ .

Note that in the expressions  $n_k \langle N_\ell \rangle_{n_i, n_j}$ , the species  $X_k \in \mathbf{C}_{ij}$  and  $X_\ell \notin \mathbf{C}_{ij}$  coreact, which means that  $E_{k\ell} \in \mathcal{E}$ . The multiplane method consists of the following approximation:

$$(3.2) \quad n_k \langle N_\ell \rangle_{n_i, n_j} \approx n_k \langle N_\ell \rangle_{k\ell, n_k} \quad \text{and} \quad \langle N_k N_\ell \rangle_{n_i, n_j} \approx \langle N_k N_\ell \rangle_{k\ell},$$

where averages  $\langle \cdot \rangle_{k\ell}$  are with respect to the marginal distribution  $P_{k\ell}$ , and conditional averages  $\langle \cdot \rangle_{k\ell, n_k}$  are averages with respect to  $P_{k\ell}$  conditioned on  $N_k = n_k$ . We thus obtain a closed system of equations for the  $|\mathcal{E}|$  marginal distributions. The approximate marginal distributions  $Q_{ij}$  satisfy the equation

$$(3.3) \quad \begin{aligned} \frac{dQ_{ij}}{dt} = & f_i (S_i^- - I) Q_{ij} + f_j (S_j^- - I) Q_{ij} + w_i (S_i^+ - I) (n_i Q_{ij}) + w_j (S_j^+ - I) (n_j Q_{ij}) \\ & + a_{ij} (S_i^+ S_j^+ - I) (n_i n_j Q_{ij}) \\ & + \sum_{(i, \ell, m) \in \mathbf{R}} a_{i\ell} (S_i^+ - I) (n_i \langle N_\ell \rangle_{i\ell, n_i} Q_{ij}) + \sum_{(j, \ell, m) \in \mathbf{R}} a_{j\ell} (S_j^+ - I) (n_j \langle N_\ell \rangle_{j\ell, n_j} Q_{ij}) \\ & + \sum_{(i, \ell, j) \in \mathbf{R}} a_{i\ell} (S_i^+ S_j^- - I) (n_i \langle N_\ell \rangle_{i\ell, n_i} Q_{ij}) + \sum_{(j, \ell, i) \in \mathbf{R}} a_{j\ell} (S_j^+ S_i^- - I) (n_j \langle N_\ell \rangle_{j\ell, n_j} Q_{ij}) \\ & + \sum_{(k, \ell, i) \in \mathbf{R}} a_{k\ell} (S_i^- - I) (\langle N_k N_\ell \rangle_{k\ell} Q_{ij}) + \sum_{(k, \ell, j) \in \mathbf{R}} a_{k\ell} (S_j^- - I) (\langle N_k N_\ell \rangle_{k\ell} Q_{ij}), \end{aligned}$$

where

$$(3.4) \quad \langle N_\ell \rangle_{k\ell, n_k} = \frac{\sum_{n_\ell} n_\ell Q_{k\ell}(n_k, n_\ell)}{\sum_{n_\ell} Q_{k\ell}(n_k, n_\ell)} \quad \text{and} \quad \langle N_k N_\ell \rangle_{k\ell} = \sum_{n_k, n_\ell} n_k n_\ell Q_{k\ell}(n_k, n_\ell).$$

This system of  $|\mathcal{E}|$  bivariate master equations is coupled through the conditional moments (3.4). Note that unlike the original master equation, the multiplane equations (3.3) are nonlinear.

Consider a situation where there are  $M$  species and the typical population size of each species is  $\langle N \rangle$ . In such a case, the truncated master equation is a linear system of size  $\langle N \rangle^M$ . The number of binary reactions  $|\mathcal{E}|$  is at most  $M(M - 1)/2$ . The multiplane method yields a nonlinear system of dimension

$$|\mathcal{E}| \langle N \rangle^2,$$

which for large  $M$  is significantly smaller than  $\langle N \rangle^M$ , even in the extreme case of  $|\mathcal{E}| = M(M - 1)/2$ .

*Example.* For the network in Figure 2.1 with the end products  $X_4, X_5$  eliminated,  $\mathcal{E} = \{E_{12}, E_{13}\}$ ,  $Q_{12}(n_1, n_2)$ , and  $Q_{13}(n_1, n_3)$  satisfy

$$\begin{aligned}
 \frac{dQ_{12}}{dt} &= f_1(S_1^- - I)Q_{12} + f_2(S_2^- - I)Q_{12} \\
 &\quad + w_1(S_1^+ - I)(n_1Q_{12}) + w_2(S_2^+ - I)(n_2Q_{12}) \\
 &\quad + a_{12}(S_1^+S_2^+ - I)(n_1n_2Q_{12}) + a_{13}(S_1^+ - I)(n_1\langle N_3 \rangle_{13,n_1}Q_{12}), \\
 \frac{dQ_{13}}{dt} &= f_1(S_1^- - I)Q_{13} + f_3(S_3^- - I)Q_{13} \\
 &\quad + w_1(S_1^+ - I)(n_1Q_{13}) + w_3(S_3^+ - I)(n_3Q_{13}) \\
 &\quad + a_{13}(S_1^+S_3^+ - I)(n_1n_3Q_{13}) + a_{12}(S_1^+ - I)(n_1\langle N_2 \rangle_{12,n_1}Q_{13}),
 \end{aligned}
 \tag{3.5}$$

where

$$\langle N_3 \rangle_{13,n_1} = \frac{\sum_{n_3} n_3 Q_{13}(n_1, n_3)}{\sum_{n_3} Q_{13}(n_1, n_3)} \quad \text{and} \quad \langle N_2 \rangle_{12,n_1} = \frac{\sum_{n_2} n_2 Q_{12}(n_1, n_2)}{\sum_{n_2} Q_{12}(n_1, n_2)}.$$

We expect this approximation to yield accurate estimates for moments such as

$$\langle N_1 \rangle, \langle N_2 \rangle, \langle N_3 \rangle, \langle N_1N_2 \rangle, \langle N_1N_3 \rangle$$

but not for, say,  $\langle N_2N_3 \rangle$ .

**4. A numerical example.** To get better insight into the multiplane method, consider once again the network of Figure 2.1 with the end products  $X_4$  and  $X_5$  eliminated. The infinite-dimensional master equation can be turned into a finite-dimensional one by truncating the distribution at a certain population size. To obtain a particularly low-dimensional system, we truncate the distribution at  $n = 1$ ; i.e., every one of the three species exists in either zero or one copy. Thus, the only allowed states are  $(0, 0, 0)$ ,  $(1, 0, 0)$ ,  $(0, 1, 0)$ ,  $(0, 0, 1)$ ,  $(1, 1, 0)$ ,  $(1, 0, 1)$ ,  $(0, 1, 1)$ , and  $(1, 1, 1)$ . This eight-state Markov process is depicted in Figure 4.1.

Ordering the states by their binary representation, the master equation takes the form

$$\frac{dP}{dt} = \mathbb{R}P,
 \tag{4.1}$$

where the *rate matrix* is

$$\mathbb{R} = \begin{pmatrix} \star & w_3 & w_2 & 0 & w_1 & a_{13} & a_{12} & 0 \\ f_3 & \star & 0 & w_2 & 0 & w_1 & 0 & a_{12} \\ f_2 & 0 & \star & w_3 & 0 & 0 & w_1 & a_{13} \\ 0 & f_2 & f_3 & \star & 0 & 0 & 0 & w_1 \\ f_1 & 0 & 0 & 0 & \star & w_3 & w_2 & 0 \\ 0 & f_1 & 0 & 0 & f_3 & \star & 0 & w_2 \\ 0 & 0 & f_1 & 0 & f_2 & 0 & \star & w_3 \\ 0 & 0 & 0 & f_1 & 0 & f_2 & f_3 & \star \end{pmatrix}.$$

The diagonal entries of  $\mathbb{R}$ , which we represent by  $\star$ , make each column sum up to zero. Taking an arbitrary choice of parameters,

$$\begin{aligned}
 f_1 &= 1, & f_2 &= 2, & f_3 &= 3, \\
 w_1 &= 4, & w_2 &= 5, & w_3 &= 6, \\
 a_{12} &= 7, & a_{13} &= 8,
 \end{aligned}$$



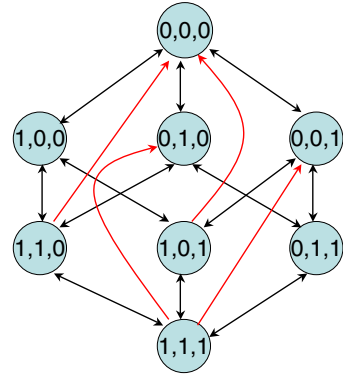


FIG. 4.1. Graphical representation of the Markov process obtained by truncating the population size of each reactive species at  $N = 1$ .

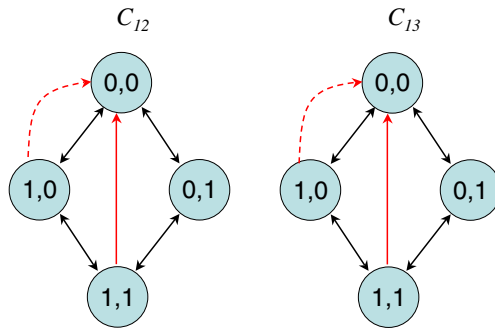


FIG. 4.2. Graphical representation of the two reduced Markov processes obtained by applying the multiplane method to the Markov process of Figure 4.1.

(4.1) has the following stationary distribution:

$$P \begin{pmatrix} 0, 0, 0 \\ 0, 0, 1 \\ 0, 1, 0 \\ 0, 1, 1 \\ 1, 0, 0 \\ 1, 0, 1 \\ 1, 1, 0 \\ 1, 1, 1 \end{pmatrix} = \begin{pmatrix} 0.4385 \\ 0.2032 \\ 0.1615 \\ 0.0762 \\ 0.0739 \\ 0.0227 \\ 0.0181 \\ 0.0059 \end{pmatrix} .$$

In contrast, the multiplane method introduces the two Markov processes depicted in Figure 4.2. The network on the left represents  $C_{12} = \{X_1, X_2\}$ , and the network on the right represents  $C_{13} = \{X_1, X_3\}$ . As above, solid edges represent reactions; the dashed edges represent reactions whose rates involve average population sizes with respect to the other network.

The two reduced master equations,

$$\frac{dQ_{12}}{dt} = \mathbb{R}_{12}Q_{12} \quad \text{and} \quad \frac{dQ_{13}}{dt} = \mathbb{R}_{13}Q_{13},$$

have rate matrices

$$\mathbb{R}_{12} = \begin{pmatrix} \star & w_2 & w_1 + a_{13}\langle N_3 \rangle_{13,1} & a_{12} \\ f_2 & \star & 0 & w_1 \\ f_1 & 0 & \star & w_2 \\ 0 & f_1 & f_2 & \star \end{pmatrix}$$

and

$$\mathbb{R}_{13} = \begin{pmatrix} \star & w_3 & w_1 + a_{12}\langle N_2 \rangle_{12,1} & a_{13} \\ f_3 & \star & 0 & w_1 \\ f_1 & 0 & \star & w_3 \\ 0 & f_1 & f_3 & \star \end{pmatrix},$$

where

$$\langle N_2 \rangle_{12,1} = \frac{Q_{12}(1,1)}{Q_{12}(1,0) + Q_{12}(1,1)},$$

$$\langle N_3 \rangle_{13,1} = \frac{Q_{13}(1,1)}{Q_{13}(1,0) + Q_{13}(1,1)}.$$

This nonlinear system can be solved, for example, iteratively. The stationary distributions are

$$Q_{12} \begin{pmatrix} 0,0 \\ 0,1 \\ 1,0 \\ 1,1 \end{pmatrix} = \begin{pmatrix} 0.6438 \\ 0.2324 \\ 0.0972 \\ 0.0267 \end{pmatrix} \quad \text{and} \quad Q_{13} \begin{pmatrix} 0,0 \\ 0,1 \\ 1,0 \\ 1,1 \end{pmatrix} = \begin{pmatrix} 0.6016 \\ 0.2754 \\ 0.0924 \\ 0.0307 \end{pmatrix}.$$

Note, in particular, that the marginal distribution of  $X_1$  is not the same in the two subsystems. The exact marginal distributions of  $\mathbf{C}_{12}$  and  $\mathbf{C}_{13}$  are

$$P_{12} \begin{pmatrix} 0,0 \\ 0,1 \\ 1,0 \\ 1,1 \end{pmatrix} = \begin{pmatrix} 0.6416 \\ 0.2377 \\ 0.0967 \\ 0.0240 \end{pmatrix} \quad \text{and} \quad P_{13} \begin{pmatrix} 0,0 \\ 0,1 \\ 1,0 \\ 1,1 \end{pmatrix} = \begin{pmatrix} 0.6000 \\ 0.2793 \\ 0.0921 \\ 0.0286 \end{pmatrix}.$$

The agreement between the two procedures is surprisingly good. The Kullback–Leibler (KL) divergence between the approximate and exact marginal distributions is

$$KL[Q_{12}|P_{12}] = 2.08 \times 10^{-4} \quad \text{and} \quad KL[Q_{13}|P_{13}] = 1.06 \times 10^{-4}.$$

The total variation (TV) distance, on the other hand, is of order 0.01, which is excellent, given that all the parameters are  $O(1)$ .

As will be shown below, this truncated Markov process provides an excellent approximation in the limit of small grain size. We set the grain size to be  $s = \epsilon$  and take  $f_i \mapsto \epsilon f_i$  and  $a_{ij} \mapsto a_{ij}/\epsilon$  (the reaction rates are proportional to the mean sweeping time of the grain's surface by a diffusing atom). In such situations, the mean population sizes are of the order  $O(\epsilon)$ . Numerical tests show that the TV distances between  $P_{12}, P_{13}$  and  $Q_{12}, Q_{13}$  scale like  $O(\epsilon^2)$ . The predicted reaction rates show a *relative error* that is  $O(\epsilon)$  (see Figure 4.3). The reason for this asymptotic convergence will be clarified in the next section.

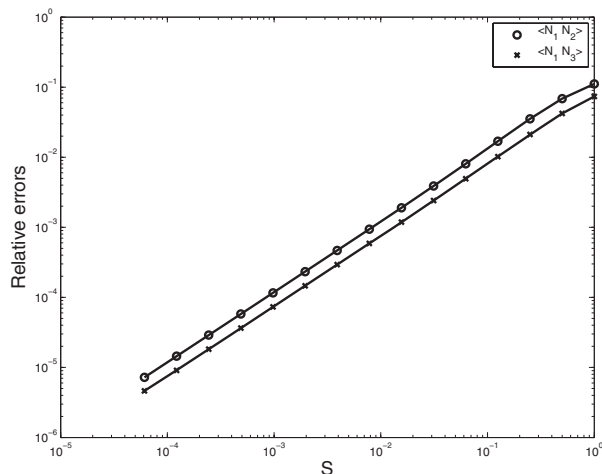


FIG. 4.3. Numerical results for the model problem of Figure 4.1: comparison between the full master equation and the multiplane method. Relative errors in the prediction of the correlations  $\langle N_1 N_2 \rangle$  (circles) and  $\langle N_1 N_3 \rangle$  (crosses) versus the grain size  $s$ . The relative errors decay proportionally to the grain size.

**5. Asymptotic analysis for small grains.** We now turn to analyze the accuracy of the multiplane method in two limits. In this section we treat the limit of small grain size, whereas the limit of large grain size is treated in the next section. The limit of small grains is imposed by setting  $f_i \mapsto \epsilon f_i$  and  $a_{ij} \mapsto a_{ij}/\epsilon$ , where  $s = \epsilon \ll 1$  is the size of the grain. That is, influxes are very weak and reactions are very fast; hence, as will be shown, mean population sizes are  $O(\epsilon)$ .

The numerical tests in the previous section suggest that the multiplane method yields approximations with relative errors that are  $O(\epsilon)$ . Note that we are interested only in relative errors since the  $s \rightarrow 0$  limit is trivial—the limiting distribution is concentrated at the empty state—hence all absolute errors tend to zero.

At steady state the master equation takes the form

$$(5.1) \quad 0 = \epsilon \sum_{X_i \in \mathbf{X}} f_i (S_i^- - I) P + \sum_{X_i \in \mathbf{X}} w_i (S_i^+ - I) (n_i P) + \frac{1}{\epsilon} \sum_{(i,j,k) \in \mathbf{R}} a_{ij} (S_i^+ S_j^+ S_k^- - I) (n_i n_j P),$$

whereas the multiplane model consists of the coupled system

$$(5.2) \quad \begin{aligned} 0 = & \epsilon f_i (S_i^- - I) Q_{ij} + \epsilon f_j (S_j^- - I) Q_{ij} + w_i (S_i^+ - I) (n_i Q_{ij}) + w_j (S_j^+ - I) (n_j Q_{ij}) \\ & + \frac{a_{ij}}{\epsilon} (S_i^+ S_j^+ - I) (n_i n_j Q_{ij}) \\ & + \sum_{(i,\ell,m) \in \mathbf{R}} \frac{a_{i\ell}}{\epsilon} (S_i^+ - I) (n_i \langle N_\ell \rangle_{i\ell, n_i} Q_{ij}) + \sum_{(j,\ell,m) \in \mathbf{R}} \frac{a_{j\ell}}{\epsilon} (S_j^+ - I) (n_j \langle N_\ell \rangle_{j\ell, n_j} Q_{ij}) \\ & + \sum_{(i,\ell,j) \in \mathbf{R}} \frac{a_{i\ell}}{\epsilon} (S_i^+ S_j^- - I) (n_i \langle N_\ell \rangle_{i\ell, n_i} Q_{ij}) + \sum_{(j,\ell,i) \in \mathbf{R}} \frac{a_{j\ell}}{\epsilon} (S_j^+ S_i^- - I) (n_j \langle N_\ell \rangle_{j\ell, n_j} Q_{ij}) \\ & + \sum_{(k,\ell,i) \in \mathbf{R}} \frac{a_{k\ell}}{\epsilon} (S_i^- - I) (\langle N_k N_\ell \rangle_{k\ell} Q_{ij}) + \sum_{(k,\ell,j) \in \mathbf{R}} \frac{a_{k\ell}}{\epsilon} (S_j^- - I) (\langle N_k N_\ell \rangle_{k\ell} Q_{ij}) \end{aligned}$$

for all  $E_{ij} \in \mathcal{E}$ , where

$$\langle N_\ell \rangle_{k\ell, n_k} = \frac{\sum_{n_\ell} n_\ell Q_{k\ell}(n_k, n_\ell)}{\sum_{n_\ell} Q_{k\ell}(n_k, n_\ell)} \quad \text{and} \quad \langle N_k N_\ell \rangle_{k\ell} = \sum_{n_k, n_\ell} n_k n_\ell Q_{k\ell}(n_k, n_\ell).$$

**5.1. Expansion of the full master equation.** To investigate the asymptotic behavior of the solution to (5.1) as  $\epsilon \rightarrow 0$ , we formally expand the stationary solution in powers of  $\epsilon$ :

$$P(\mathbf{n}) = P^{(0)}(\mathbf{n}) + \epsilon P^{(1)}(\mathbf{n}) + \epsilon^2 P^{(2)}(\mathbf{n}) + \dots$$

We then substitute this expansion into (5.1) and equate terms of the same power of  $\epsilon$ . The result is a hierarchy of equations:

$$\begin{aligned} 0 &= \sum_{(i,j,k) \in \mathbf{R}} a_{ij} (S_i^+ S_j^+ S_k^- - I) (n_i n_j P^{(0)}), \\ 0 &= \sum_{X_i \in \mathbf{X}} w_i (S_i^+ - I) (n_i P^{(0)}) + \sum_{(i,j,k) \in \mathbf{R}} a_{ij} (S_i^+ S_j^+ S_k^- - I) (n_i n_j P^{(1)}), \\ 0 &= \sum_{X_i \in \mathbf{X}} f_i (S_i^- - I) P^{(0)} + \sum_{X_i \in \mathbf{X}} w_i (S_i^+ - I) (n_i P^{(1)}) \\ &\quad + \sum_{(i,j,k) \in \mathbf{R}} a_{ij} (S_i^+ S_j^+ S_k^- - I) (n_i n_j P^{(2)}), \end{aligned}$$

and so on.

The first set of equations in the hierarchy consists only of reaction terms. For every state  $\mathbf{n}_0$  we sum up this equation for all  $\mathbf{n}$  below it (lexicographically), yielding, for all  $(i, j, k) \in \mathbf{R}$ ,

$$S_i^+ S_j^+ S_k^- (n_i n_j P^{(0)}) = 0,$$

or

$$P^{(0)}(\mathbf{n}) = 0 \quad \text{if there exist } n_i, n_j > 0 \text{ such that } E_{ij} \in \mathcal{E}.$$

The second set of equations in the hierarchy consists of both desorption and reaction terms. Here again, partial summation up to every  $\mathbf{n}$  yields, for every  $X_i$ ,

$$S_i^+ (n_i P^{(0)}) = 0,$$

i.e.,

$$P^{(0)}(\mathbf{n}) = 0 \quad \text{for all } \mathbf{n} \neq \mathbf{0}.$$

Moreover, for all  $(i, j, k) \in \mathbf{R}$ ,

$$S_i^+ S_j^+ S_k^- (n_i n_j P^{(1)}) = 0$$

or

$$P^{(1)}(\mathbf{n}) = 0 \quad \text{if there exist } n_i, n_j > 0 \text{ such that } E_{ij} \in \mathcal{E}.$$

Thus, the only entry of  $P^{(0)}$  that does not vanish is  $P^{(0)}(\mathbf{0})$ , which by normalization is equal to one:

$$(5.3) \quad P^{(0)}(\mathbf{n}) = \begin{cases} 1 & \text{for } \mathbf{n} = \mathbf{0}, \\ 0 & \text{for } \mathbf{n} \neq \mathbf{0}. \end{cases}$$

We proceed with the third set of equations in the hierarchy, which now includes influxes as well. It is easy to see that

$$\begin{aligned} P^{(1)}(\mathbf{n}) &= 0 && \text{for } |\mathbf{n}| > 1, \\ P^{(2)}(\mathbf{n}) &= 0 && \text{if there exist } n_i, n_j > 0 \text{ such that } E_{ij} \in \mathcal{E}, \end{aligned}$$

where  $|\mathbf{n}| = n_1 + \dots + n_M$  is the total population size. The fact that the probability of finding a pair of reacting molecules is at most  $O(\epsilon^3)$  is not surprising. Without reactions, the probability of any population size of two or more is  $O(\epsilon^2)$ . Since the reaction rates scale like  $1/\epsilon$ , it follows that the probability of finding a pair of reacting species is (at least) by a factor of  $\epsilon$  lower than in the absence of reactions.

The only nonvanishing equations at this level in the hierarchy yield

$$(5.4) \quad P^{(1)}(\mathbf{n}) = \begin{cases} \frac{f_i}{w_i} & \text{for } |\mathbf{n}| = 1, n_i = 1, \\ -\sum_{j=1}^M \frac{f_j}{w_j} & \text{for } \mathbf{n} = \mathbf{0}, \\ 0 & \text{for } |\mathbf{n}| > 1. \end{cases}$$

We then take the fourth set of equations in the hierarchy, namely,

$$0 = \sum_{X_i \in \mathbf{X}} f_i (S_i^- - I) P^{(1)} + \sum_{X_i \in \mathbf{X}} w_i (S_i^+ - I) (n_i P^{(2)}) + \sum_{(i,j,k) \in \mathbf{R}} a_{ij} (S_i^+ S_j^+ S_k^- - I) (n_i n_j P^{(3)}).$$

By now, it is clear that the distribution for a given  $\mathbf{n}$  scales at most like  $\epsilon$  to the power of  $|\mathbf{n}|$ . The entries of  $P^{(2)}$  for two-particle states are given by

$$(5.5) \quad P^{(2)}(\mathbf{n}) = \begin{cases} \frac{1}{w_i + w_j} \left( \frac{f_i^2}{w_i} + \frac{f_j^2}{w_j} \right) & \text{for } |\mathbf{n}| = 2, n_i = n_j = 1, E_{ij} \notin \mathcal{E}, \\ \frac{f_i^2}{2w_i^2} & \text{for } |\mathbf{n}| = 2, n_i = 2, a_{ii} \notin \mathcal{E}, \\ 0 & \text{for } |\mathbf{n}| = 2, n_i = n_j = 1, E_{ij} \in \mathcal{E}, \end{cases}$$

whereas the entries of  $P^{(2)}$  for the one-particle states  $|\mathbf{n}| = 1, n_i = 1$  are given by

$$P^{(2)}(\mathbf{n}) = -\frac{f_i}{w_i} \left( \sum_{X_j \in \mathbf{X}} \frac{f_j}{w_j} \right) - \frac{f_i}{w_i^2} \left( \sum_{X_j \in \mathbf{X}} f_j \right) + \frac{f_i^2}{w_i^2} + \sum_{j, E_{ij} \notin \mathcal{E}} \frac{w_j/w_i}{w_i + w_j} \left( \frac{f_i^2}{w_i} + \frac{f_j^2}{w_j} \right).$$

By normalization, we may then extract  $P^{(2)}(\mathbf{0})$ .

Finally, given a reaction  $(i, j, k) \in \mathbf{R}$ , we set  $|\mathbf{n}| = 1, n_k = 1$  to find

$$(5.6) \quad P^{(3)}(\mathbf{n}) = \frac{f_i f_j}{a_{ij}} \left( \frac{1}{w_i} + \frac{1}{w_j} \right) \quad \text{for } |\mathbf{n}| = 2, n_i = n_j = 1, E_{ij} \in \mathcal{E}.$$

Thus, (5.3), (5.4), and (5.5) give to leading order the distributions of zero-, one-, and two-particle states, whereas (5.6) gives the leading order term for the two-interacting-particle states. This expansion in powers of  $\epsilon$  can, in principle, be continued further. At this point we are already in measure to evaluate the mean reaction rates:

$$(5.7) \quad r_{ij} = f_i f_j \left( \frac{1}{w_i} + \frac{1}{w_j} \right) \epsilon^3 + O(\epsilon^4).$$

Note, in particular, that the influxes contribute multiplicatively to the reactions rates, whereas the desorption rates act “in parallel.” Note also the “asymptotic independence” between any two reactions; the probability of the two-particle state of two reactive species is asymptotically unaffected by all other reactions.

**5.2. Expansion of the multiplane method.** A similar asymptotic expansion can be performed for the multiplane equations (5.2). It is easy to see that to leading order the results are unchanged, meaning that relative errors between the two procedures are at most  $O(\epsilon)$ . Specifically, we assume the following asymptotic behavior:

$$(5.8) \quad \langle N_\ell \rangle_{k\ell, n_k} = O(\epsilon^2) \quad \text{and} \quad \langle N_k N_\ell \rangle_{k\ell} = O(\epsilon^3),$$

which implies that the equation for  $Q_{ij}$  is a closed equation up to relative errors of order  $\epsilon$ :

$$(5.9) \quad 0 = \epsilon f_i (1 + O(\epsilon)) (S_i^- - I) Q_{ij} + \epsilon f_j (S_j^- - I) Q_{ij} \\ + w_i (1 + O(\epsilon)) (S_i^+ - I) (n_i Q_{ij}) + w_j (S_j^+ - I) (n_j Q_{ij}) + \frac{a_{ij}}{\epsilon} (S_i^+ S_j^+ - I) (n_i n_j Q_{ij}).$$

Expanding  $Q_{ij}(\mathbf{n})$  in power of  $\epsilon$ ,

$$Q_{ij}(\mathbf{n}) = Q_{ij}^{(0)}(\mathbf{n}) + \epsilon Q_{ij}^{(1)}(\mathbf{n}) + \epsilon^2 Q_{ij}^{(2)}(\mathbf{n}) + \dots,$$

we obtain, after an analysis similar but much simpler than in the previous section, the following solution:

$$\begin{aligned} Q_{ij}(0, 0) &= 1 + O(\epsilon), \\ Q_{ij}(1, 0) &= \frac{f_i}{w_i} \epsilon + O(\epsilon^2), \\ Q_{ij}(0, 1) &= \frac{f_j}{w_j} \epsilon + O(\epsilon^2), \\ Q_{ij}(2, 0) &= \frac{f_i^2}{2w_i^2} \epsilon^2 + O(\epsilon^3), \\ Q_{ij}(0, 2) &= \frac{f_j^2}{2w_j^2} \epsilon^2 + O(\epsilon^3), \\ Q_{ij}(1, 1) &= \frac{f_i f_j}{a_{ij}} \left( \frac{1}{w_i} + \frac{1}{w_j} \right) \epsilon^3 + O(\epsilon^4). \end{aligned}$$

One may now verify a posteriori that the ansatz (5.8) is consistent with this solution.

If we denote by  $\hat{r}_{ij}$  the stationary reaction rate predicted by the multiplane method, then

$$(5.10) \quad \hat{r}_{ij} = f_i f_j \left( \frac{1}{w_i} + \frac{1}{w_j} \right) \epsilon^3 + O(\epsilon^4),$$

which compared with (5.7) gives one side of (1.1).

**6. Asymptotic analysis for large grains.** In the limit of large grains we set  $f_i \mapsto f_i/\epsilon$  and  $a_{ij} \mapsto \epsilon a_{ij}$ , where the grain size is  $s = 1/\epsilon$ ,  $\epsilon \ll 1$ . In this limit, influxes are very strong and reactions are very slow, and, as a result, mean population sizes are of order  $O(1/\epsilon)$ . Our goal is to show that, like in the small grain limit, the multiplane method yields approximations with *relative errors* that decay with  $\epsilon$ .

For large population size, one performs an asymptotic expansion known as *Van Kampen's  $\Omega$ -expansion* [21, 8]. Specifically, we define for each species  $X_i$  a rescaled population size  $x_i = \epsilon n_i$ , which is an order one quantity, viewed as a continuous parameter. We then define  $\rho(\mathbf{x}) = P(\mathbf{x}/\epsilon)$ . Our analysis relies on a formal asymptotic expansion. Limit theorems making this analysis rigorous may be found in Kurtz [15, 16].

**6.1. Expansion of the full master equation.** Substituting the scaled parameters  $f_i, a_{ij}$  and the definition of the distribution  $\rho(\mathbf{x})$ , the master equation (2.2) takes the form

$$\epsilon \frac{\partial \rho}{\partial t} = \sum_{X_i \in \mathbf{X}} f_i (S_i^- - I)\rho + \sum_{X_i \in \mathbf{X}} w_i (S_i^+ - I)(x_i \rho) + \sum_{(i,j,k) \in \mathbf{R}} a_{ij} (S_i^+ S_j^+ S_k^- - I)(x_i x_j \rho),$$

where the increment and decrement operators add or subtract in quanta of  $\epsilon$ . Assuming that  $\rho(\mathbf{x})$  converges as  $\epsilon \rightarrow 0$  to a twice differentiable function, we expand:

$$S_i^\pm = I \pm \epsilon \frac{\partial}{\partial x_i} + \frac{\epsilon^2}{2} \frac{\partial^2}{\partial x_i^2} + O(\epsilon^3),$$

which leads to the following asymptotic expansion of the master equation:

$$\begin{aligned} (6.1) \quad \frac{\partial \rho}{\partial t} &= - \sum_{X_i \in \mathbf{X}} f_i \frac{\partial \rho}{\partial x_i} + \sum_{X_i \in \mathbf{X}} w_i \frac{\partial(x_i \rho)}{\partial x_i} + \sum_{(i,j,k) \in \mathbf{R}} a_{ij} \left( \frac{\partial}{\partial x_i} + \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_k} \right) (x_i x_j \rho) \\ &+ \frac{\epsilon}{2} \sum_{X_i \in \mathbf{X}} f_i \frac{\partial^2 \rho}{\partial x_i^2} + \frac{\epsilon}{2} \sum_{X_i \in \mathbf{X}} w_i \frac{\partial^2(x_i \rho)}{\partial x_i^2} + \frac{\epsilon}{2} \sum_{(i,j,k) \in \mathbf{R}} a_{ij} \left( \frac{\partial}{\partial x_i} + \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_k} \right)^2 (x_i x_j \rho) \\ &+ O(\epsilon^2). \end{aligned}$$

To  $O(\epsilon^2)$  terms, (6.1) is a *transport-diffusion* equation. The  $O(1)$  terms, which involve only first derivatives, can be identified as the *Liouville equation* for the deterministic system

$$(6.2) \quad \frac{d\mathbf{x}}{dt} = \mathbf{g}(\mathbf{x}),$$

where

$$g_i(\mathbf{x}) = f_i - w_i x_i - \sum_{(i,k,\ell) \in \mathbf{R}} a_{ik} x_i x_k + \sum_{(k,\ell,i) \in \mathbf{R}} a_{k\ell} x_k x_\ell.$$

Equations (6.2) are the *rate equations*, which constitute the deterministic large population size limit.

The  $O(\epsilon)$  terms, which involve second derivatives, transform the equation into a parabolic *Fokker-Planck equation* that corresponds to the Itô *stochastic differential system*

$$(6.3) \quad d\mathbf{x} = \mathbf{g}(\mathbf{x}) dt + \sqrt{\epsilon} \Gamma(\mathbf{x}) d\mathbf{W},$$

where  $\mathbf{W}(t)$  is a vector of  $M$  independent Brownian motions, and  $\Gamma(\mathbf{x})$  is the  $M$ -by- $M$  diffusion matrix, whose square has entries

$$\Gamma_{ij}^2(\mathbf{x}) = \left( f_i + w_i x_i + \sum_{(i,k,\ell) \in \mathbf{R}} a_{ik} x_i x_k + \sum_{(k,\ell,i) \in \mathbf{R}} a_{k\ell} x_k x_\ell \right) \delta_{ij} + \sum_{(i,j,k) \in \mathbf{R}} a_{ij} x_i x_j + \sum_{(i,k,j) \in \mathbf{R}} a_{ik} x_i x_k.$$

(The square root of a matrix is defined up to an orthogonal rotation; we may choose without loss of generality the unique symmetric positive-definite square root.) This stochastic system is the *diffusion approximation* to the master equation in the limit of large population size.

At this stage, the Fokker–Planck equation (6.1) can be solved perturbatively, using a WKB expansion [2]. To leading order, the rescaled population size  $\mathbf{x}$  tends, up to an  $1/\sqrt{\epsilon}$  stochastic correction, to the solution of the rate equation (6.2). Switching back to the original discrete variables,  $\mathbf{n}$ , the stationary reaction rates predicted by the rate equations deviate by a relative error of  $O(\epsilon)$  from the exact stationary reaction rate. Namely, if  $\tilde{r}_{ij}$  are the reaction rates predicted by the rate equations, then

$$(6.4) \quad \frac{\tilde{r}_{ij} - r_{ij}}{r_{ij}} = O(\epsilon).$$

**6.2. Expansion of the multiplane method.** A similar expansion can be performed for the multiplane method. First, we replace  $Q_{ij}(\mathbf{n})$  by  $\sigma_{ij}(\epsilon \mathbf{n})$ :

$$(6.5) \quad \begin{aligned} \epsilon \frac{d\sigma_{ij}}{dt} = & f_i (S_i^- - I) \sigma_{ij} + f_j (S_j^- - I) \sigma_{ij} + w_i (S_i^+ - I) (x_i \sigma_{ij}) + w_j (S_j^+ - I) (x_j \sigma_{ij}) \\ & + a_{ij} (S_i^+ S_j^+ - I) (x_i x_j \sigma_{ij}) \\ & + \sum_{(i,\ell,m) \in \mathbf{R}} a_{i\ell} (S_i^+ - I) (x_i \langle X_\ell \rangle_{i\ell, x_i} \sigma_{ij}) + \sum_{(j,\ell,m) \in \mathbf{R}} a_{j\ell} (S_j^+ - I) (x_j \langle X_\ell \rangle_{j\ell, x_j} \sigma_{ij}) \\ & + \sum_{(i,\ell,j) \in \mathbf{R}} a_{i\ell} (S_i^+ S_j^- - I) (x_i \langle X_\ell \rangle_{i\ell, x_i} \sigma_{ij}) + \sum_{(j,\ell,i) \in \mathbf{R}} a_{j\ell} (S_j^+ S_i^- - I) (x_j \langle X_\ell \rangle_{j\ell, x_j} \sigma_{ij}) \\ & + \sum_{(k,\ell,i) \in \mathbf{R}} a_{k\ell} (S_i^- - I) (\langle X_k X_\ell \rangle_{k\ell} \sigma_{ij}) + \sum_{(k,\ell,j) \in \mathbf{R}} a_{k\ell} (S_j^- - I) (\langle X_k X_\ell \rangle_{k\ell} \sigma_{ij}), \end{aligned}$$

where

$$\langle X_\ell \rangle_{k\ell, x_k} = \frac{\sum_{x_\ell} x_\ell \sigma_{k\ell}(x_k, x_\ell)}{\sum_{x_\ell} \sigma_{k\ell}(x_k, x_\ell)} \quad \text{and} \quad \langle X_k X_\ell \rangle_{k\ell} = \sum_{x_k, x_\ell} x_k x_\ell \sigma_{k\ell}(x_k, x_\ell).$$

(Recall that the summation over  $k, \ell, m$  is over all species not in  $\mathbf{C}_{ij}$ .)

We proceed with the expansion as in the previous subsection, except that the functions  $\langle X_\ell \rangle_{k\ell, x_k}$  and  $\langle X_k X_\ell \rangle_{k\ell}$  are a priori unknown. The function  $\sigma_{ij}$  satisfies to order  $O(\epsilon)$ :

$$\begin{aligned} \frac{d\sigma_{ij}}{dt} = & -f_i \frac{\partial}{\partial x_i} \sigma_{ij} - f_j \frac{\partial}{\partial x_j} \sigma_{ij} + w_i \frac{\partial}{\partial x_i} (x_i \sigma_{ij}) + w_j \frac{\partial}{\partial x_j} (x_j \sigma_{ij}) \\ & + a_{ij} \left( \frac{\partial}{\partial x_i} + \frac{\partial}{\partial x_j} \right) (x_i x_j \sigma_{ij}) \end{aligned}$$



$$\begin{aligned}
 &+ \sum_{(i,\ell,m) \in \mathbf{R}} a_{i\ell} \frac{\partial}{\partial x_i} (x_i \langle X_\ell \rangle_{i\ell, x_i} \sigma_{ij}) + \sum_{(j,\ell,m) \in \mathbf{R}} a_{j\ell} \frac{\partial}{\partial x_j} (x_j \langle X_\ell \rangle_{j\ell, x_j} \sigma_{ij}) \\
 &+ \sum_{(i,\ell,j) \in \mathbf{R}} a_{i\ell} \left( \frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_j} \right) (x_i \langle X_\ell \rangle_{i\ell, x_i} \sigma_{ij}) \\
 &+ \sum_{(j,\ell,i) \in \mathbf{R}} a_{j\ell} \left( \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_i} \right) (x_j \langle X_\ell \rangle_{j\ell, x_j} \sigma_{ij}) \\
 &- \sum_{(k,\ell,i) \in \mathbf{R}} a_{k\ell} \frac{\partial}{\partial x_i} (\langle X_k X_\ell \rangle_{k\ell} \sigma_{ij}) - \sum_{(k,\ell,j) \in \mathbf{R}} a_{k\ell} \frac{\partial}{\partial x_j} (\langle X_k X_\ell \rangle_{k\ell} \sigma_{ij}).
 \end{aligned}$$

The next order terms are diffusive.

The equation for  $\sigma_{ij}$  is the Fokker–Planck equation for the Itô stochastic differential system

$$(6.6) \quad d\mathbf{x}_{ij} = \mathbf{g}_{ij}(\mathbf{x}_{ij}) dt + \sqrt{\epsilon} \Gamma_{ij}(\mathbf{x}_{ij}) d\mathbf{W}_{ij},$$

where  $\mathbf{x}_{ij} = (x_i, x_j)$  is a stochastic process representing the populations sizes in  $\mathbf{C}_{ij}$ . The drift function  $\mathbf{g}_{ij}$  is given by

$$\begin{aligned}
 [\mathbf{g}_{ij}]_i(\mathbf{x}_{ij}) &= f_i - w_i x_i - a_{ij} x_i x_j - \sum_{(i,\ell,m) \in \mathbf{R}} a_{i\ell} x_i \langle X_\ell \rangle_{i\ell, x_i} \\
 &- \sum_{(i,\ell,j) \in \mathbf{R}} a_{i\ell} x_j \langle X_\ell \rangle_{i\ell, x_i} + \sum_{(j,\ell,i) \in \mathbf{R}} a_{j\ell} x_j \langle X_\ell \rangle_{j\ell, x_i} + \sum_{(k,\ell,i) \in \mathbf{R}} a_{k\ell} \langle X_k X_\ell \rangle_{k\ell}.
 \end{aligned}$$

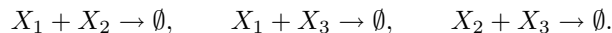
An analogous expression holds for  $[\mathbf{g}_{ij}]_j(\mathbf{x}_{ij})$ .

The set of coupled Fokker–Planck equations can also be solved perturbatively, using a WKB expansion. To leading order, the copies of the same  $x_i$  coincide, and the variables  $\{\mathbf{x}_{ij}\}$ , with their common components identified, satisfy the same rate equations (6.2) as  $\mathbf{x}$  in the previous section. Thus, the reaction rates  $\hat{r}_{ij}$  predicted by the multiplane approximation deviate by a relative error of  $O(\epsilon)$  from the prediction of the rate equations,

$$\frac{\hat{r}_{ij} - \tilde{r}_{ij}}{\tilde{r}_{ij}} = O(1/s), \quad s \gg 1,$$

and hence from the exact stationary reaction rate,  $r_{ij}$ .

**7. A numerical test.** We now test of the accuracy of the multiplane method for a particular example. We consider a network of three reactive species that form a “clique”—every pair of species coreacts. The products of the three reactions are not reactive; hence they can be eliminated. Thus, the network consists of the following reactions:



The multiplane method calculates the approximate bivariate distributions  $Q_{12}$ ,  $Q_{13}$ , and  $Q_{23}$ . We used the following parameters:

$$\begin{aligned}
 f_1 &= s, & f_2 &= 1.5s, & f_3 &= 2s, \\
 w_1 &= 2.5, & w_2 &= 3, & w_3 &= 3.5, \\
 a_{12} &= 1/s, & a_{13} &= 2/s, & a_{23} &= 3/s,
 \end{aligned}$$

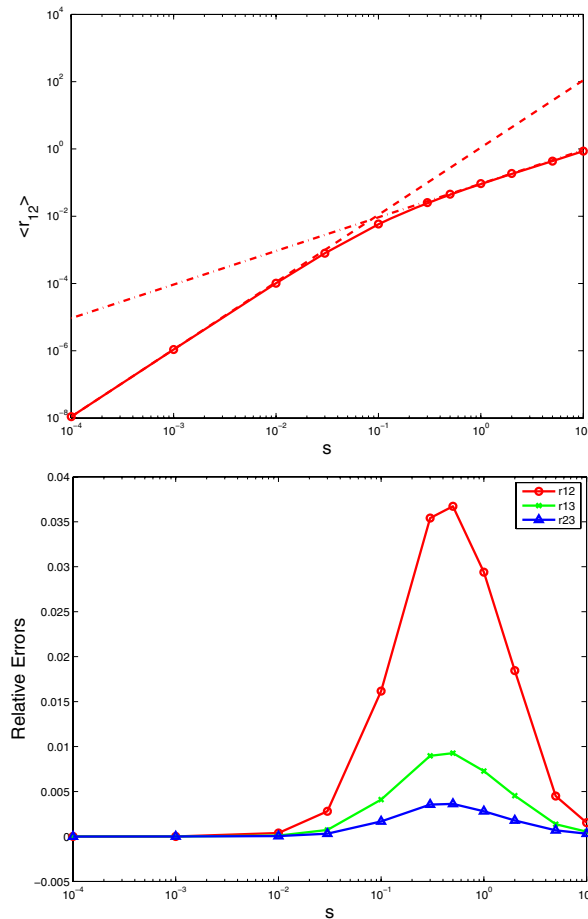


FIG. 7.1. *Top: Log-log plot of the stationary reaction rate  $r_{12}$  versus the grain size  $s$  for the clique of three reactive species. Comparison between the exact solution (circles), the multiplane prediction (solid line), the small- $s$  asymptotic behavior (5.7) (dashed line), and the rate equations' prediction (dash-dotted line). Bottom: Semilog plot of the relative errors of the reaction rates predicted by the multiplane method.*

where  $s$  is the grain size. In Figure 7.1 (top) we plot the stationary reaction rate  $r_{12}$  as a function of the grain size  $s$ . The exact solution is represented by circles, The multiplane prediction is represented by a solid line, the small- $s$  limiting behavior (5.7) is represented by a dashed line, and the rate equations' solution is represented by a dash-dotted line. Note the accurate interpolation of the multiplane method between the two limiting regimes. In Figure 7.1 (bottom) we plot the relative errors of the three stationary reaction rates,  $r_{12}$ ,  $r_{13}$ , and  $r_{23}$ , predicted by the multiplane method. The relative errors reach a maximum in the vicinity of  $s = 1$  and, as predicted, decay in both limits of small and large grains. Note that even at their peak, the relative errors are of the order of a few percents, which is (once again) surprisingly good.

**8. Discussion.** The main goal of this paper was to systematically formulate the multiplane method and shed some light on the reasons for its success, despite its “uncontrolled” nature, as reported in [17]. We considered reaction networks in which every reaction is generated by two reactants. The multiplane method was defined as an

algorithm for producing a closed system of equations for approximating the marginal distributions of all pairs of coreacting species. Thus, the heart of the method is the premise that most of the dynamics are captured by the joint distributions of pairs of coreacting species. Even in the case where a set of reactants constitutes a *clique*, their joint distribution can be approximated by a set of pairwise distributions for the sake of predicting mean reaction rates. Our methodology can easily be adapted to more general situations, such as reactions with more than two reactants. Moreover, while the present analysis treats irreversible reactions, the dissociation of molecules can easily be accounted for by the addition of flux-like terms into the master equation.

We analyzed the method in two extreme situations, weak influxes and fast reactions, on the one hand, and strong influxes and slow reactions, on the other hand. In both limits, the multiplane method was found to yield predictions with relative errors that vanish asymptotically. It should be emphasized that the two limits involve very different analyses, and the asymptotic exactness of the method in both limits has to be attributed to totally different reasons. While each of the limits can be approximated by its own methods (rate equations in one limit and a low-dimensional Markov process in the other limit), the multiplane method has the merit to “interpolate” between those two limits. As such, it makes a good candidate for dealing with intermediate situations in which no simple analysis is available.

The multiplane method can be viewed as part of an abundance of approximation methods for structured probabilistic models [14]. A similar surprising accuracy was observed, for example, in the context of general continuous-time Bayesian networks [19], where no small parameter exists. A natural question is whether such approaches can be justified on more general grounds, for example, by identifying them as optimal within a certain subclass of models.

**Acknowledgment.** We are grateful to Nir Friedman for inspiring discussions.

#### REFERENCES

- [1] B. BARZEL, O. BIHAM, AND R. KUPFERMAN, *Evaluation of the multiplane method for efficient simulations of reaction networks*, Phys. Rev. E (3), 76 (2007), 026703.
- [2] C. BENDER AND S. ORSZAG, *Advanced Mathematical Methods for Scientists and Engineers*, Springer-Verlag, New York, 1999.
- [3] O. BIHAM, I. FURMAN, V. PIRRONELO, AND G. VIDALI, *Master equation for hydrogen recombination on grain surfaces*, Astrophys. J., 553 (2001), pp. 595–603.
- [4] O. BIHAM AND A. LIPSHTAT, *Exact results for hydrogen recombination on dust grain surfaces*, Phys. Rev. E (3), 66 (2002), 056103.
- [5] P. CASELLI, T. HASEGAWA, AND E. HERBST, *A proposed modification of the rate equations for reactions on grain surfaces*, Astrophys. J., 495 (1998), pp. 309–316.
- [6] S. CHARNLEY, *Stochastic theory of molecule formation on dust*, Astrophys. J., 562 (2001), pp. L99–L102.
- [7] S. CHARNLEY, A. TIELENS, AND S. RODGERS, *Deuterated methanol in the Orion compact ridge*, Astrophys. J., 482 (1997), pp. L203–L206.
- [8] C. GARDINER, *Handbook of Stochastic Methods*, 3rd ed., Springer-Verlag, New York, 2004.
- [9] D. GILLESPIE, *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*, J. Computational Phys., 22 (1976), pp. 403–434.
- [10] D. GIVON, R. KUPFERMAN, AND A. STUART, *Extracting macroscopic dynamics: Model problems and algorithms*, Nonlinearity, 17 (2004), pp. R55–R127.
- [11] C. GOMEZ-URIBE AND G. VERGHESE, *Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations*, J. Chem. Phys., 126 (2007), 024109.
- [12] N. GREEN, T. TONIAZZO, M. PILLING, D. RUFFLE, N. BELL, AND T. HARTQUIST, *A stochastic approach to grain surface chemical kinetics*, Astronom. and Astrophys., 375 (2001), pp. 1111–1119.

- [13] T. HASEGAWA, E. HERBST, AND C. LEUNG, *Models of gas-grain chemistry in dense interstellar clouds with complex organic molecules*, *Astrophys. J. Suppl.*, 82 (1992), pp. 167–195.
- [14] D. KOLLER AND N. FRIEDMAN, *Structured Probabilistic Models: Representation, Inference, and Learning*, manuscript.
- [15] T. KURTZ, *Limit theorems and diffusion approximations for density dependent Markov chains*, in *Stochastic Systems: Modeling, Identification and Optimization, I*, North-Holland, Amsterdam, 1976, pp. 67–78.
- [16] T. KURTZ, *Strong approximation theorems for density dependent Markov chains*, *Stochastic Processes Appl.*, 6 (1977/1978), pp. 223–240.
- [17] A. LIPSHTAT AND O. BIHAM, *Efficient simulation of gas-grain chemistry in interstellar clouds*, *Phys. Rev. Lett.*, 93 (2004), 170601.
- [18] D. MCQUARRIE, *Stochastic approach to chemical kinetics*, *J. Appl. Probability*, 4 (1967), pp. 413–478.
- [19] U. NODELMAN, C. SHELTON, AND D. KOLLER, *Continuous time Bayesian networks*, in *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann, San Francisco, 2002, pp. 378–387.
- [20] T. STANTCHEVA AND E. HERBST, *Deuterium fractionation on interstellar grains studied with the direct master equation approach*, *Monthly Notices Roy. Astronom. Soc.*, 340 (2003), pp. 983–988.
- [21] N. VAN KAMPEN, *Elimination of fast variables*, *Phys. Rep.*, 124 (1985), pp. 69–160.