

# Chemical Equilibrium on Low Dimensional Supports: Connecting the Microscopic Mechanism to the Macroscopic Observations

D. Bullara<sup>1</sup> · Y. De Decker<sup>1</sup>

Received: 13 March 2015 / Accepted: 30 June 2015 / Published online: 10 July 2015  
© Springer Science+Business Media New York 2015

**Abstract** Classical chemical thermodynamics predicts that the equilibrium composition of a reactive system is entirely defined by the equilibrium constants of the different reactions involved. In this paper we show that for nonlinear reactions taking place on a low-dimensional support this is not true anymore: the equilibrium state depends on the mechanistic details of the chemical processes, so that even two reactions having the same mean field kinetics and equilibrium constants can reach a different equilibrium composition, depending on the microscopic mechanism. We illustrate this point by simulations and mathematical analyses of a simple autocatalytic scheme, and we propose a theoretical route to discriminate between the different cases.

**Keywords** Low-dimensional systems · Stochastic dynamics · Nonlinear phenomena · Reactive systems

## 1 Introduction

Equilibrium thermodynamics is an extremely powerful and useful theoretical framework, with applications ranging from chemistry [1,2], to the management of energy resources [3,4], to astrophysics [5,6], to cite but a few examples. Much of the well-deserved success of this theory can be attributed to its intrinsically universal character. Macroscopic systems subject to a given class of time-independent boundary conditions are supposed to reach a state that coincides with the extremum of the same state function. A major strength of equilibrium thermodynamics is that this property is supposed to hold, irrespective of the microscopic details of the components of the system or of the processes taking place.

However, this fundamental feature can sometimes be put into danger. The kinetics and the equilibrium state of chemical reactions occurring in low-dimensional systems are for example

---

✉ Y. De Decker  
ydedecke@ulb.ac.be

<sup>1</sup> Center for Nonlinear Phenomena and Complex Systems (CENOLI), Université libre de Bruxelles, CP 231, 1050 Brussels, Belgium

known to deviate substantially from the above precepts [7]. Previous studies revealed that such deviations can be related to the nonlinearity of the reaction and to the dimensionality of the support [8]. However, despite all the knowledge gathered on this class of systems over the last decades, there is as of today no simple general criterion upon which one could predict whether equilibrium thermodynamics will hold or not. In this work, we show that such a rule could find its roots in the properties of the space of spatial configurations of the system at hand. Since these critical properties are related directly to the mechanism of the elementary steps, one can assess *a priori* the validity of the traditional macroscopic approach.

We will consider as an illustration the case of reactions taking place on one-dimensional supports. Such situations are encountered for example in the channels of a zeolite [9,10], or in carbon nanotubes [11,12]. More particularly, our focus will be on a typical example of nonlinear reaction, the trimolecular autocatalytic scheme



which is common to many chemical schemes giving rise to complex behavior such as temporal oscillations or pattern formation [13]. Classical thermodynamics tells us that the ratio of the molar fractions of X and Y is given, at equilibrium, by

$$\left(\frac{x_X}{x_Y}\right)_{eq} = K_{eq}, \quad (2)$$

independently of the details of the kinetics behind reaction (1). From a mechanistic point of view, however, such effectively trimolecular reactions are usually understood as being the result of a succession of two or more monomolecular or bimolecular steps. In Sect. 2, we present different reaction mechanisms of this type that can lead to a globally trimolecular process. We first discuss them in the framework of a mean-field description, and then focus on their microscopic implementation on a one-dimensional support. Section 3 is devoted to a presentation of the results of stochastic (Kinetic Monte Carlo) simulations of these different realizations. We show that some of the mechanisms can lead the system to its equilibrium composition (2), while others cannot. We explain and rationalize these results in Sect. 4 and 5, with an analytical study based on the blending of combinatorial arguments and properties of the space of microscopic configurations. We finish by summarizing our main results and by pointing towards future developments (Sect. 6 and 7).

## 2 Mechanisms for the Trimolecular Reaction

We start by discussing how different mechanisms can lead to an effectively trimolecular reaction kinetics within the mean-field hypothesis. Under ideal conditions, the dynamical evolution of the molar fractions  $x_X$  and  $x_Y$  undergoing reaction (1) is typically given by

$$\frac{d}{dt}x_Y = -\frac{d}{dt}x_X = J, \quad (3)$$

where  $J$  is the rate of reaction (1) given by the *mass action law*

$$J = k_f x_X^2 x_Y - k_b x_X^3, \quad (4)$$

with  $k_f$  and  $k_b$  being the kinetic constants of the forward and backward processes. Equation (4) strictly implies that reaction (1) is an elementary process, with a molecularity of 3.

It would be more realistic from a chemical point of view to think of reaction (1) as a combination of, say, the following three elementary steps:



involving the intermediate species E and F. If reactions (5) and (7) evolve in both ways on a faster time scale than reaction (6), the molar fractions  $x_E$  and  $x_F$  of the intermediate species can be respectively replaced by the equilibrium conditions on (5) and (7). If these values are small, so that the changes over time of  $x_E$  and  $x_F$  are negligible, the species E and F can be considered in a stationary state. Under such *pre-equilibrium steady state approximation*, the rate of the global process reads:

$$J_I = k_I K_E x_X^2 x_Y - k_{-I} K_F^{-1} x_X^3, \quad (8)$$

with  $k_I$  and  $k_{-I}$  being the forward and backward kinetic constants of (6), and  $K_E$  and  $K_F$  the equilibrium constants of (5) and (7).

Schemes (5)–(7) is clearly not the only one compatible with (1). For example, another sequence of acceptable monomolecular and bimolecular steps is:



where the intermediate species D and F appear. Assuming the pre-equilibrium steady state approximation for both D and F gives the global rate

$$J_{II} = k_{II} K_D x_X^2 x_Y - k_{-II} K_F^{-1} x_X^3, \quad (12)$$

where  $k_{II}$ ,  $k_{-II}$  and  $K_D$  respectively are the forward and backward kinetic constants of (10) and the equilibrium constant of (9).

Both  $J_I$  and  $J_{II}$  have the same form as  $J$ , which makes both schemes (5)–(7) and (9)–(11) chemically acceptable schemes for the trimolecular reaction (1). However, the reactive events involved in these schemes are different, and this fact needs to be taken into account when the two above mechanisms are translated into a set of microscopic local rules in low-dimensional systems.

We model such spatially restricted environments as linear sequences of boxes with no-flux boundary conditions (i.e., as lines). Each box is occupied either by a particle X or by a particle Y. We also assume that the volume of these species is conserved when they are chemically bound into the intermediate species D, E and F, so that these latter three particles each occupy two boxes. Finally we also consider that all the particles are impenetrable, so that swapping of adjacent species is ruled out.

Since the global mechanism is trimolecular, it is useful to focus on the local behavior of triplets of adjacent particles. For convenience we will discuss the forward sequence of reaction: the backward series of events can be easily understood using similar ideas.

Because chemical reactions are short-ranged events, the different elementary steps involve particles that occupy adjacent positions on the support. Schemes (5)–(7) requires one particle of X and one of Y to be close to each other in order to form E. Therefore, any reactive triplet must have at least one XY or, assuming the reaction to be symmetric in space, one YX sequence. Now, the intermediate species E needs to react with a nearby X. As an example,



### 3 Stochastic Simulations

To investigate the properties of Case I and Case II, we performed Kinetic Monte Carlo simulations of the above two schemes on one-dimensional lattices with no-flux boundary conditions. For the sake of simplicity we set  $K_{eq} = 1$ , so that the equilibrium values for  $x_X$  and  $x_Y$  predicted by the mass action law both are 0.5. The initial condition of each simulation was constructed by independently assigning to each box of the lattice a Y or an X with probability  $p$  and  $1 - p$ , respectively. We also investigated the role played by the linear size  $L$  of the system.

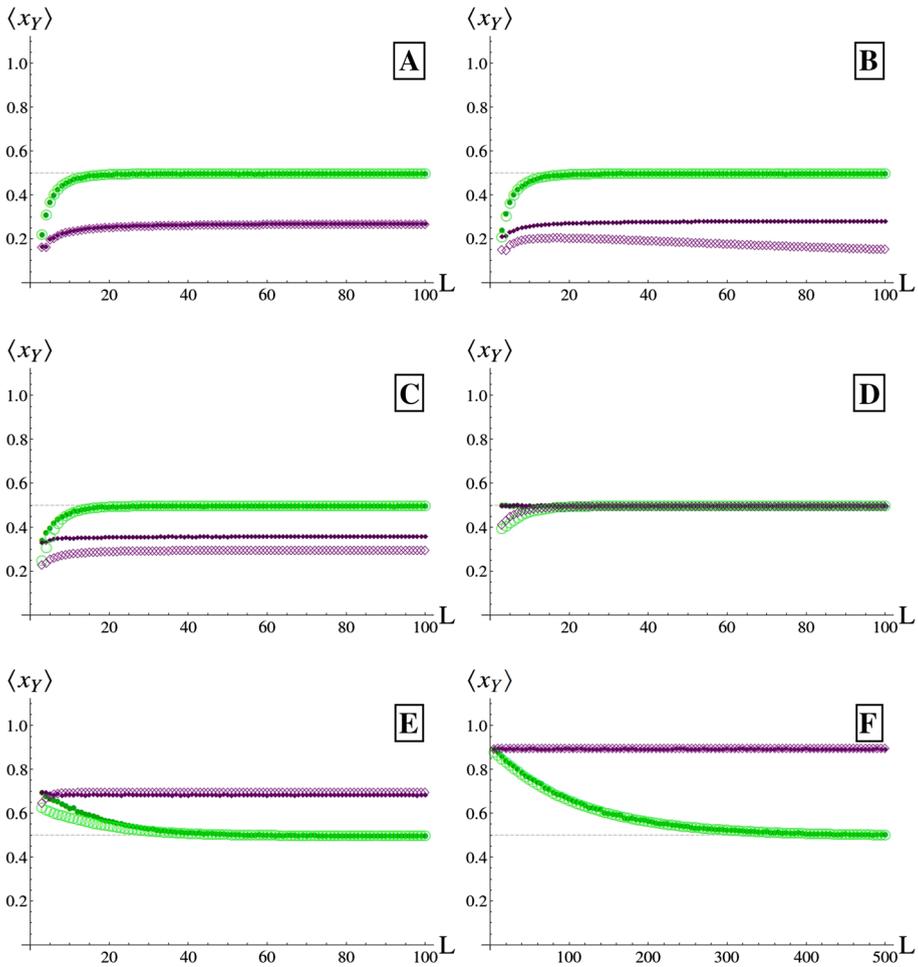
Figure 1 compares the ensemble-averaged equilibrium values  $\langle x_Y \rangle$  of  $x_Y$  for the two mechanisms. The plots show the dependence of  $\langle x_Y \rangle$  on the lattice size, for different choices of the initial condition parameter  $p$ . We observe that despite the fact that Case I and Case II are very similar and derive from the same global reaction, they present quite different behaviors when simulated on a one-dimensional lattice. On one hand, the thermodynamic prediction is reached by Case II for a large range of initial conditions in the macroscopic limit. Substantial deviations are observed only when the initial condition is rich in Y. On the other hand, the equilibrium reached in Case I never follows the mass action law. The final state displays a strong, non-trivial dependence on the initial condition. A closer look reveals that the equilibrium composition tends to remain stuck at its initial value, as the latter becomes more concentrated in Y. It is worth noticing that in both cases, the final value also depends on the system size (see Fig. 1). As the linear size  $L$  of the lattice increases the final state asymptotically approaches a value, which coincides with the macroscopic limit discussed above. The rate of convergence to this limiting value depends on the initial condition: as the initial concentration in Y increases, so does the value of  $L$  above which the macroscopic limit can be obtained. Figure 2 summarizes the differences between the two schemes in such a macroscopic limit.

The very fact that such systems sometimes do not follow the mass action law should not come as a surprise. It is indeed well known that when nonlinear processes take place on low-dimensional supports, the mean field hypothesis is put in danger and so are the predictions of traditional (non)equilibrium thermodynamics [8]. Systematic analyses of different simple schemes revealed that the extent of the deviations depends on the dimensionality of the underlying network and on the molecularity of the processes involved [14]. Such explanations are here inadequate, since both mechanisms are trimolecular and are simulated on the same support. If it is not the dimensionality or the molecularity, then what makes Case I and Case II behave so differently?

### 4 Combinatorial Approach

To unveil the reasons behind the observed differences, we will use a combined statistical and combinatorial approach inspired by a previous study devoted to the strong dependence on initial conditions of Case I [15]. In this work, the authors used the properties of the transition matrix appearing in the master equation, which rules the stochastic evolution of the simulations. Let us first illustrate these ideas qualitatively on a simple case.

The central point is to analyze the space of all the possible spatial configurations of particles. The total number of configurations for the problem hereby studied in a system of size  $L$  is given by the combinatorial factor:

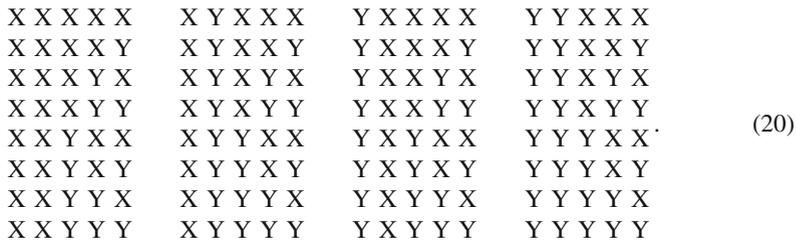
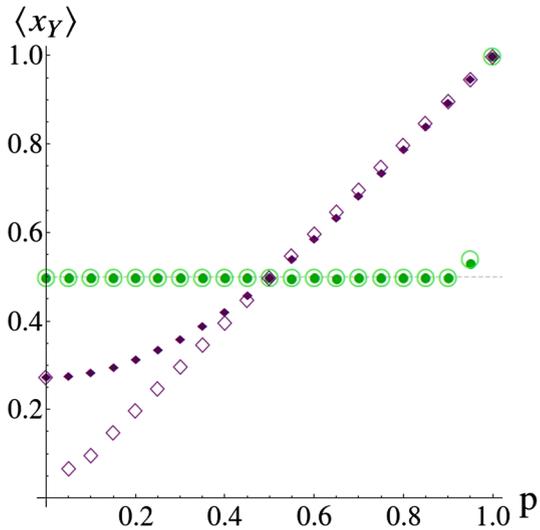


**Fig. 1** Average simulated and theoretical values of  $\langle x_Y \rangle$  at equilibrium as a function of different lattice sizes  $L$ , for Case I (purple diamonds) and Case II (green circles). Full-small marks refer to computational values, calculated by averaging the outcome of 5000 independent Kinetic Monte Carlo simulations, with equal probabilities for the forward and backward reactions and with no flux boundary conditions. Each simulation runs for 5000 time units and the equilibrium value is averaged over the last 50 points of each simulation. Each of the six plots correspond to a different value of  $p$ . **a**  $p = 0$ ; **b**  $p = 0.1$ ; **c**  $p = 0.3$ ; **d**  $p = 0.5$ ; **e**  $p = 0.7$ ; **f**  $p = 0.9$ . Empty-big marks refer to the theoretical estimations calculated with Eq. (39). The gray-dashed line marks the ideal value predicted by the mass action law (2), with a kinetic ratio  $K_{eq} = 1$  (see also Eqs. (17)–(18)) (Color figure online)

$$C_L = \sum_{n=0}^L \binom{L}{n}. \tag{19}$$

We will also refer to  $C_L$  as the *size of the configurational space*. As an example, for  $L = 5$  there are 32 configurations that read:

**Fig. 2** Steady state molar fraction of Y as a function of the initial condition, as obtained from 5000 realizations of the kinetic Monte Carlo simulations. The initial condition corresponds to randomly placed particles of Y and X with probabilities  $p$  and  $(1 - p)$ , respectively. The size of the system is in each case  $L = 1000$ . Colors and symbol legends are the same as in Fig. 1 (Color figure online)



The next step is the analysis of the transitions between these states. Although the possible configurations (20) are the same for both Case I and Case II, the transitions between them are very different because the microscopic mechanism is not the same. To show this, one can construct a graph involving the configurations in the following way. We consider each configuration to be the node of a network and link two nodes if it is possible to transform one configuration into the other by a single reactive event. In such a graph, a given node will have  $L$  different links to other nodes if any of the  $L$  lattice sites can undergo a reaction. Otherwise it will have a *self loop* and a number of links equal to the number of reactive particles. Given the specificities of the reaction schemes and the no-flux boundary conditions, It should be clear at this point that:

1. Any two linked configurations only differ by one single particle;
2. If a link exists from the configuration  $i$  to the configuration  $j$ , an equivalent link also exists from  $j$  to  $i$  (we will denote this situation by drawing a single non-directed link between  $i$  and  $j$ );
3. Every configuration will have a self loop;
4. Frozen (time-invariant) configurations will appear as isolated nodes linked only to themselves.

Figure 3 shows how such a graph looks like when  $L = 5$  for Case I and Case II. It appears that the two cases lead to substantially different pictures. Case I segregates the configurational space into several intraconnected networks of different length on one hand, and isolated nodes



cases, the final molar fraction is equal to the initial one, since the system remains *frozen* in its initial state.

With these properties in mind, one can reproduce the behavior of the equilibrium state for the two cases. Indeed, for many different realizations of the simulation algorithm, the ensemble average of  $x_Y$  will be given by

$$\langle x_Y \rangle = \sum_{i=1}^N P_i x_Y^{(i)} + \left( 1 - \sum_{i=1}^N P_i \right) x_Y^{(0)}, \quad (21)$$

where  $P_i$  is the probability to start in an intraconnected network  $i$ , for which the mean molar fraction of Y is  $x_Y^{(i)}$ . Note that  $\left( 1 - \sum_{i=1}^N P_i \right)$  is the probability to start in an isolated configuration, for which  $x_Y$  is thus equal to the initial condition:  $x_Y^{(0)}$  is therefore the average molar fraction calculated among all the isolated states.

All these quantities could in principle be accessed analytically or numerically, although their calculation is not always simple or straightforward. In the next section we will analytically derive some of them, which we will then use to find approximate laws for the behavior observed in the simulations.

## 5 Analytical Results

Our goal is to obtain analytical explicit expressions of (21) that will help us to understand the differences between the two reactive schemes. To this end, we will focus on the behavior of the systems starting in a configuration belonging to the largest intraconnected network of the graph representation introduced previously.

### 5.1 Mean Molar Fraction of the Largest Network

In this subsection we will explicitly calculate the equilibrium molar fraction reached by the largest graph in both Case I and II. This corresponds to starting the system in a globally reactive configuration, i.e., a configuration in which there is no frozen “islands” of particles. We start by calculating the transition matrix  $T_L$  within the subregion of the configurational space corresponding to the largest connected network. In this case  $T_L$  is non-negative and irreducible by definition. Since it will prove to be also stochastic, we can use the ergodicity ensured by the Frobenius theorem to calculate the first moment of any equilibrium quantity by averaging over the total number of globally active configurations.

The original derivation for Case I has been performed by Provata et al. [15]. We will shortly repeat it hereafter for the sake of completeness and for a better comparison with the rest of our presentation.

#### 5.1.1 Case I

The microscopic rules of Case I forbid many of the configurations, because they require two Xs to surround a third particle for it to be reactive. This excludes the possibility for any configuration with one or more blocks of adjacent Ys to be either obtained or removed by the chemical reactions.

Because of this limitation and the no-flux boundary conditions, the maximum number of Ys allowed in a globally reactive configuration of size  $L$  is  $(L - 1)/2$ , corresponding

to an alternating sequence of X and Y. Note that the largest graphs always correspond to configurations where the two boundaries are X species. If the number  $n$  of Ys is less or equal to this upper limit, it will be possible to create reactive configurations by placing the  $n$  Ys in  $L - n - 1$  different positions (see [15]), so that the above constraints are satisfied. The number of reactive configurations for each value of  $n$  is therefore given by

$$C_{L,n}^I = \binom{L - n - 1}{n}, \tag{22}$$

with the total number of reactive configurations being

$$C_L^I = \sum_{n=0}^{\frac{L-1}{2}} C_{L,n}^I. \tag{23}$$

For example, the only 5 allowed configurations for  $L = 5$  are:

$$\begin{array}{ccccc} X & X & X & X & X \\ X & X & X & Y & X \\ X & X & Y & X & X \\ X & Y & X & X & X \\ X & Y & X & Y & X \end{array} \tag{24}$$

The transition matrix for this 5-boxes example reads:

$$T_5^I = \begin{pmatrix} \frac{2}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & 0 \\ \frac{1}{5} & \frac{3}{5} & 0 & 0 & \frac{1}{5} \\ \frac{1}{5} & 0 & \frac{4}{5} & 0 & 0 \\ \frac{1}{5} & 0 & 0 & \frac{3}{5} & \frac{1}{5} \\ 0 & \frac{1}{5} & 0 & \frac{1}{5} & \frac{3}{5} \end{pmatrix}. \tag{25}$$

Because of the above mentioned restrictions, the mean molar fraction of the largest network in Case I is given by

$$x_Y^{(1),I} = \frac{1}{L} \frac{\sum_{n=0}^{\frac{L-1}{2}} C_{L,n}^I n}{C_L^I}. \tag{26}$$

### 5.1.2 Case II

The microscopic rules of Case II allow a particle to react only if it is adjacent to a block of two Xs. Since the reaction is reversible and can proceed either to the left or to the right, any configuration containing at least one XX block can eventually be converted into any other configuration, with the only limitation that at least one XX block must remain. Hence, the only frozen configurations are those where all the Xs are separated by one or more Y particles.

Consequently, the maximum number of Ys allowed in a globally reactive configuration of size  $L$  is  $L - 2$ , which is far greater than for Case I. Again, reactive configurations are possible if the number  $n$  of Y particles is less or equal to the upper limit. To find the total number of such configurations, it is easier to calculate the number of frozen configurations for each  $n$  and then subtract it from the total number of configurations. According to the

*pigeonhole principle* [16], if the number of Xs is greater than  $n + 1$  there must be at least two X particles close to each other, resulting in a globally reactive configuration. If the number of Xs is less or equal to  $n + 1$ , then it is possible to create frozen configurations where each X is placed between two Y particles. For each given  $n$  and with no-flux boundary conditions, there are  $n + 1$  such positions to place each X (before the first Y, and after each Y present in the system) so that the number of reactive configurations for each value of  $n$  is

$$C_{L,n}^{II} = \left[ \binom{L}{n} - \binom{n+1}{L-n} \right]. \tag{27}$$

The total number of reactive configurations in a lattice of size  $L$  is therefore given by:

$$C_L^{II} = \sum_{n=0}^{L-2} C_{L,n}^{II}. \tag{28}$$

Consider again the simple case where  $L = 5$ . The above equation predicts the presence of 19 configurations, which read explicitly:

$$\begin{array}{cccccc}
 \text{X X X X X} & \text{X X Y X Y} & \text{X Y Y X X} & \text{Y X Y X X} \\
 \text{X X X X Y} & \text{X X Y Y X} & \text{Y X X X X} & \text{Y Y X X X} \\
 \text{X X X Y X} & \text{X X Y Y Y} & \text{Y X X X Y} & \text{Y Y X X Y} \\
 \text{X X X Y Y} & \text{X Y X X X} & \text{Y X X Y X} & \text{Y Y Y X X} \\
 \text{X X Y X X} & \text{X Y X X Y} & \text{Y X X Y Y} & 
 \end{array} \tag{29}$$

The transition matrix among these configurations is:

$$T_5^{II} = \begin{pmatrix}
 0 & \frac{1}{5} & \frac{1}{5} & 0 & \frac{1}{5} & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \frac{1}{5} & 0 & 0 & \frac{1}{5} & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & \frac{2}{5} & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 \\
 0 & \frac{1}{5} & 0 & \frac{2}{5} & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 \\
 \frac{1}{5} & 0 & 0 & 0 & \frac{4}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & \frac{1}{5} & 0 & 0 & 0 & \frac{4}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & \frac{4}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & \frac{4}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{3}{5} & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{4}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & \frac{1}{5} & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{5} & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{5} & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & \frac{4}{5}
 \end{pmatrix} \tag{30}$$

As a consequence of the above, the mean molar fraction of the largest connected network in Case II is given by

$$x_Y^{(1),II} = \frac{1}{L} \frac{\sum_{n=0}^{L-2} C_{L,n}^{II} n}{C_L^{II}}. \tag{31}$$

### 5.2 Probability to Select the Largest Network

We will here derive the probabilities  $P_1$  that the system starts in the largest network, for both Case I and Case II.

The initial condition in the simulations is built by assigning to each box in the lattice a Y or an X, with probability  $p$  and  $(1 - p)$ , respectively. Given that every box is filled independently from the others, we can estimate the probability to choose any initial configuration containing  $n$  particles Y as  $p^n (1 - p)^{L-n}$ . The probability to select the largest subgraph with such an initial condition is calculated by multiplying the number of configurations with  $n$  Ys in the largest network by the probability to obtain one such configuration, and then summing over all the possible values of  $n$  allowed in the network. For Case I and Case II we obtain

$$P_1^I = \sum_{n=0}^{\lfloor \frac{L-1}{2} \rfloor} C_{L,n}^I p^n (1 - p)^{L-n}, \tag{32}$$

$$P_1^{II} = \sum_{n=0}^{L-2} C_{L,n}^{II} p^n (1 - p)^{L-n}. \tag{33}$$

Figure 4 shows the trend of  $P_1^I$  and  $P_1^{II}$  for the six  $p$  values of Fig. 1. It can be seen that while the probability to initiate the system in the largest network approaches 1 in Case II as  $L$  increases, it vanishes as  $L$  increases for Case I.

### 5.3 Average Equilibrium Composition

The equilibrium molar fraction is given by Eq. (21), which for Case I and Case II reads:

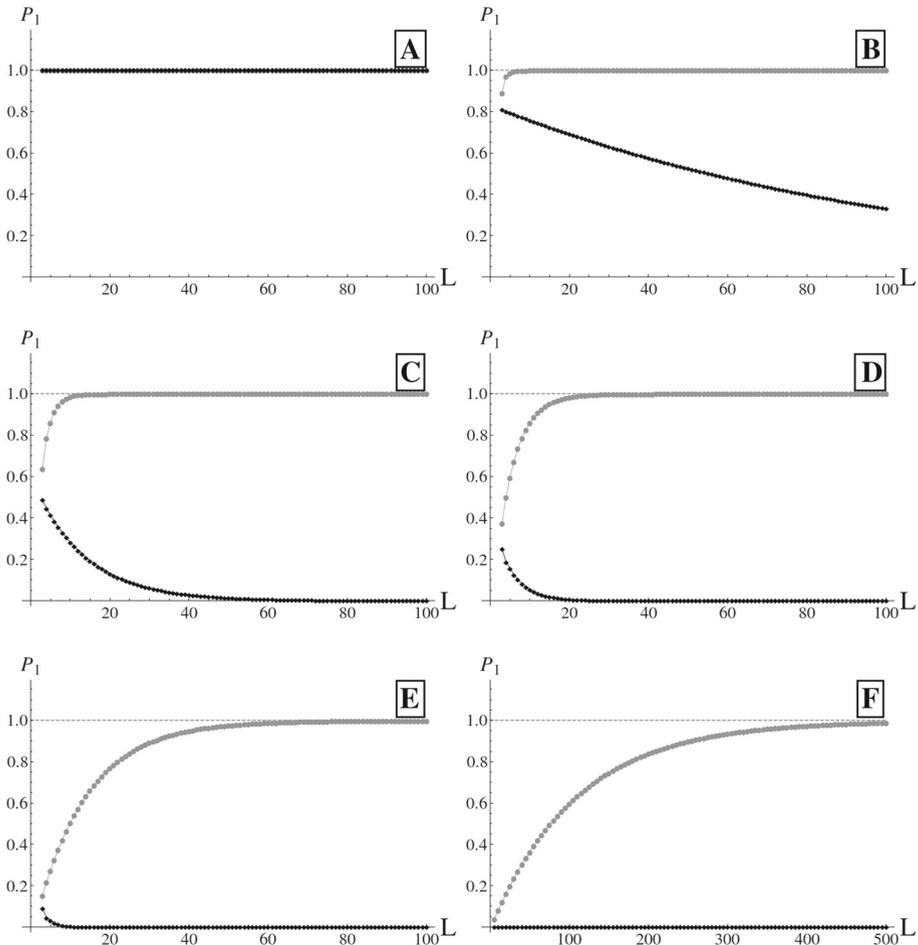
$$\langle x_Y \rangle^I = P_1^I x_Y^{(1),I} + (1 - P_1^I) x_Y^{(0),I} + \sum_{i=2}^N P_i^I (x_Y^{(i),I} - x_Y^{(0),I}), \tag{34}$$

$$\langle x_Y \rangle^{II} = P_1^{II} x_Y^{(1),II} + (1 - P_1^{II}) x_Y^{(0),II}. \tag{35}$$

Let us start with Case II for simplicity. In view of the results of the previous subsections, the only term left to be known in Eq. (35) is the average composition of all the isolated configurations  $x_Y^{(0),II}$ . One could be tempted to calculate this average simply as the mean value among the isolated states, but this would give a bad estimate of the simulated value. The problem is that not all the isolated configurations are selected with the same probability: for example if  $p$  is large, the isolated configurations which are more populated in  $p$  will be selected more often. A good approximation of  $x_Y^{(0),II}$  should take into account this probability, and would therefore read:

$$x_Y^{(0),II} = \frac{\sum_{n=0}^{L-2} \left[ \binom{L}{n} - C_{L,n}^{II} \right] n p^n (1 - p)^{L-n} + \sum_{n=L-2}^L \binom{L}{n} n p^n (1 - p)^{L-n}}{L (C_L - C_L^{II})}. \tag{36}$$

where  $\sum_{n=0}^{L-2} \left[ \binom{L}{n} - C_{L,n}^{II} \right] n p^n (1 - p)^{L-n}$  is the probability to select a frozen configuration for initial compositions for which both reactive and frozen configurations are possible (i.e., for  $0 \leq n \leq L-2$ ), and  $(C_L - C_L^{II})$  is the number of frozen configurations. The rationale behind this formula is the following. The first term in the numerator of Eq. (36) represents the probability to select a frozen configuration with a number  $n$  of Y particles, for which both



**Fig. 4**  $P_1^I$  (black diamonds) and  $P_1^{II}$  (gray circles) as a function of  $L$ . The light gray dashed line marks the unitary probability. Each of the six plots correspond to a different value of  $p$ . **a**  $p = 0$ ; **b**  $p = 0.1$ ; **c**  $p = 0.3$ ; **d**  $p = 0.5$ ; **e**  $p = 0.7$ ; **f**  $p = 0.9$ . The other parameters are the same as in Fig. 1

active and frozen configuration are possible. The second term accounts for values of  $n$  for which only frozen configurations are possible. The sum of these configurational estimates is then simply divided by the total number of frozen configurations,  $(C_L - C_L^{II})$ .

For Case I the situation is more complicated, because in principle we also need to calculate the probabilities  $P_i^I$  and the mean compositions  $x_Y^{(i),I}$  for every network in the configurational space. This calculation is rather complicated and we did not find a simple, yet analytical and general way to perform it for every lattice size  $L$ . However our scope here is to use the qualitative structure of the configurational space to keep our analysis simple by finding suitable approximations. Since Case I produces a very fragmented graph with many disconnected small networks, we can push this concept to its limit and consider that any network except the largest one is a disconnected point. We expect this approximation to hold if the term  $\sum_{i=2}^N P_i^I (x_Y^{(i),I} - x_Y^{(0),I})$  in Eq. (34) vanishes, which is true if all the probabilities  $P_i^I$  tend to

zero, or if the mean molar fraction within each smaller network is close to the average molar fraction of the ensemble of isolated nodes. Under this approximation we calculate  $x_Y^{(0),I}$  as

$$x_Y^{(0),I} \simeq \frac{\sum_{n=0}^{\frac{L-1}{2}} \left[ \binom{L}{n} - C_{L,n}^I \right] n p^n (1-p)^{L-n} + \sum_{n=\frac{L-1}{2}}^L \binom{L}{n} n p^n (1-p)^{L-n}}{L(C_L - C_L^I)}, \quad (37)$$

using the same reasoning as for Case II.

Notice that if the configurational space consisted solely of isolated nodes (i.e., if it were a fully disconnected graph), the estimated value of  $\langle x_Y \rangle$  would take the remarkably simple form

$$\langle x_Y \rangle^{\text{FD}} = \frac{1}{L} \frac{\sum_{n=0}^L \binom{L}{n} p^n (1-p)^{L-n} n}{\sum_{n=0}^L \binom{L}{n} p^n (1-p)^{L-n}} = p. \quad (38)$$

Figure 1 clearly shows that Case I fulfills this prediction for large values of  $p$ , which indicates that even the brutal approximation of a fully disconnected configurational space is not a bad one in such cases. In fact we can speculate more about this result and propose a further simplification of Eqs. (34) and (35). From a careful analysis of Case I and Case II, we can individuate two limiting behaviors

- All the networks in the configurational space have vanishing probability to be selected as an initial condition (with the sum of these probabilities nevertheless being equal to 1), which makes the system behave as a collection of isolated nodes (see Case I). In this case  $\langle x_Y \rangle \approx \langle x_Y \rangle^{\text{FD}} = p$ ;
- The biggest network has  $P_1 \approx 1$ , meaning  $\langle x_Y \rangle \simeq x_Y^{(1)}$ .

We can therefore try to approximate the behavior of the system in the intermediate cases as a linear function of  $P_1$ , which amounts to substituting the terms  $x_Y^{(0),I}$  and  $x_Y^{(0),II}$  in Eqs. (34) and (35) with  $\langle x_Y \rangle^{\text{FD}}$  and neglecting the contributions of the smaller networks in Eq. (34):

$$\langle x_Y \rangle \approx P_1 x_Y^{(1)} + (1 - P_1) \langle x_Y \rangle^{\text{FD}}. \quad (39)$$

Equation (39) has been used to derive the theoretical curves in Figs. 1 and 2 and is observed to give reasonable estimates for sufficiently large values of  $p$ , and in the macroscopic limit.

## 6 Discussion

We now want to discuss further the meaning of the results we obtained in the previous sections. For the sake of illustration, we will here focus on the macroscopic limit  $L \rightarrow \infty$ .

Remember that, for Case II, the space of configurations consists in one large intraconnected network (so that  $N = 1$ ) and several isolated nodes. More precisely, it has been shown (see Fig. 4) that in the macroscopic limit,  $P_1$  converges to 1, with a convergence rate inversely proportional to the initial molar fraction of Y. Moreover, it also appears that  $x_Y^{(1)}$  tends to 0.5 in the same limit. Since the probability to start in a configuration out of the biggest network vanishes, one finds that  $\langle x_Y \rangle \approx 0.5$  independently of the initial condition for Case II (See, again, Fig. 2). The fact that we re-obtain the mass action prediction in such limit is thus the consequence of having an almost perfectly connected configurational space.

For Case I, there can be many intraconnected networks of non-negligible size. It is not always easy to estimate  $P_i$  for each connected network. In principle, should one of these probabilities converge to 1, we would have again a situation similar to Case II, with the whole configurational space being spanned by this network. However this is not the case, and it is fairly easy to show that even the probability to start in the largest network rapidly vanishes as the lattice size increases, especially for those initial conditions which are more concentrated in Y. if  $P_i$  tends to zero for all  $i$ , the configurational space can be seen as a collection of isolated points. One then finds that  $\langle x_Y \rangle \approx \langle x_Y \rangle^{FD}$  so that the equilibrium composition is the same as the initial one, as observed in the simulations (Fig. 2). In other words, Case I deviates from the mass action law because the rules are such that most of the configurations of the system are disconnected from each other.

In the end, the different equilibrium properties of Case I and Case II can all be traced back to differences in the structure of the configurational space. It is now interesting to try and shed some light on the physicochemical reasons behind these differences and why some schemes may produce a very fragmented structure like in Case I, while others give rise to a much more connected network as in Case II.

A good starting point is to have a look again at the configurational space of Case I for  $L = 5$  in Fig. 3. The five configurations in the largest network all have two particles in common: the first and the last ones are an X in each case. This reflects the fact that these particles cannot react according to scheme (15) because of the no-flux boundary conditions. The links among those configurations are therefore due to the remaining particles, which are in a reactive configuration and allow for transitions from one node to another. Similarly, all the configurations in each network with three nodes share three inactive particles, and the links are due to the two remaining (active) ones.

The isolated nodes represent from this point of view a limiting case, where none of the lattice sites can undergo a reaction. The possibility to create such a globally frozen configuration does not however automatically imply a strong divergence from the mean field prediction: this can be a rare occurrence and the total number of frozen states can be vanishingly small as the configurational space increases (see Case II). The very reason for the fragmentation of the configurational space is rather the presence of *local clusters of non-reactive of particles*, which prevent any network to grow large enough so as to invade the majority of the configurational space.

The above conclusion means that, generally speaking, one might expect that the reactions allowing for frozen local structures will not follow the mass action law. Consequently, the validity of the traditional thermodynamic approach can often be guessed *a priori* from the microscopic mechanism, without entering into a detailed analysis such as the one we presented here. For example, consider the following cluster:



submitted to the dynamics of either Case I or Case II. The central box (the YYY cluster) cannot produce any reaction in any case. However, in order for this cluster to qualify as a frozen island the Ys at the borders must also be inactive. This is true for Case I, for which an XYX configuration is needed, but not for Case II, since the XXY configurations are reactive. Playing with the different local configurations, one arrives to the conclusion that Case I allows for the existence of inactive clusters, so that its configurational space is clustered and is hence expected to deviate from the mass action law. Case II on the contrary is able to produce a cascading process that “attacks” the initially frozen blocks from the borders and activate them. Local frozen islands are thus not permitted and the only inactive configurations are the globally frozen ones, which cannot be obtained by reaction. In consequence, Case II is

expected to follow the traditional law, as long as the initial condition is itself not a globally frozen state.

## 7 Conclusions

To conclude, we show here that for low-dimensional systems, the deviation from classical equilibrium thermodynamics can be traced back to the existence of local non-reactive islands in a system. A connection between the spatial features of such problems and their chemical composition can be obtained by analyzing the properties of the space of configurations. Simple but effective corrections to the mass action law can be derived just by looking at the structure of that space.

Although the present results are still far from forming a general theory, they suggest that the network properties of the configurational space play a pivotal role in quantifying the distance from ideal mean field predictions, more than dimensionality or coordinancy. We already highlighted how the relative size of the largest network can provide useful information to obtain accurate predictions of the non mean field behavior, even in extreme situations such as the one-dimensional lattices considered here. It would be herefore interesting to investigate the possibility of using more advanced tools from graph and network theory to derive thermodynamically relevant quantities directly from the reaction mechanism itself. This knowledge could be used not only to predict the equilibrium state reached by a system, but also to predict the non-equilibrium behavior of more complex schemes where multimolecular steps appear. Additionally, it would also be interesting to further explore the connections between the ideas we put forward here and the nonextensive statistical thermodynamics theory proposed by C. Tsallis [17], which naturally takes into account the non-ergodicity of reactive systems.

**Acknowledgments** We thank Astero Provata for stimulating discussions. D. B. is particularly thankful to Samuele Sommariva for suggesting the combinatorial importance of the pigeonhole principle on part of the present work.

## References

1. Soldatov, V.S.: Application of basic concepts of chemical thermodynamics to ion exchange equilibria. *React. Funct. Polym.* **27**, 95 (1995)
2. Leitner, J., Voňka, P., Mikulec, J.: Application of chemical thermodynamics to the description of processes of special inorganic materials preparation. *J. Mater. Sci.* **24**, 1521 (1989)
3. Züttel, A.: Hydrogen storage methods. *Naturwissenschaften* **91**, 157 (2004)
4. Hansen, A.C., Zhang, Q., Lyne, P.W.L.: Ethanol-diesel fuel blends: a review. *Bioresour. Technol.* **96**, 277 (2005)
5. Potekhin, A.Y., Chabrier, G.: Thermodynamic functions of dense plasmas: analytic approximations for astrophysical applications. *Contrib. Plasma Phys.* **50**, 82 (2010)
6. Sagi, E., Bekenstein, J.D.: Black holes in the tensor-vector-scalar theory of gravity and their thermodynamics. *Phys. Rev. D* **77**, 024010 (2008)
7. Trzaskowski, B., Adamowicz, L.: Chloromethane and dichloromethane decompositions inside nanotubes as models of reactions in confined space. *Theor. Chem. Acc.* **124**, 95 (2009)
8. Tretyakov, A., Provata, A., Nicolis, G.: Nonlinear chemical dynamics in low-dimensional lattices and fractal sets. *J. Phys. Chem.* **99**, 2770 (1995)
9. da Machado, M.S., Pérez-Pariente, J., Sastre, E., Cardoso, D., de Guereu, A.M.: Selective synthesis of glycerol monolaurate with zeolitic molecular sieves. *Appl. Catal. A Gen.* **203**, 321 (2000)
10. Lu, J., Aydin, C., Liang, A.J., Chen, C., Browning, N.D., Gates, B.C.: Site-isolated molecular iridium complex catalyst supported in the 1-dimensional channels of zeolite HSSZ-53: characterization by spectroscopy and aberration-corrected scanning transmission electron microscopy. *ACS Catal.* **2**, 1002 (2012)

11. Wang, N., Guan, L.: A chemical combination reaction within single-walled carbon nanotubes. *Nanoscale* **2**, 893 (2010)
12. Nair, N., Strano, M.S.: One-dimensional nanostructure-guided chain reactions: harmonic and anharmonic interactions. *Phys. Rev. B* **80**, 174301 (2009)
13. Prigogine, I., Lefever, R.: Symmetry breaking instabilities in dissipative systems. II. *J. Chem. Phys.* **48**, 1695 (1968)
14. Bentz, J.L., Kozak, J.J., Abad, E., Nicolis, G.: Efficiency of encounter-controlled reaction between diffusing reactants in a finite lattice: topology and boundary effects. *Phys. A* **326**, 55 (2003)
15. Provata, A., Turner, J.W., Nicolis, G.: Nonlinear chemical dynamics in low dimensions: an exactly soluble model. *J. Stat. Phys.* **70**, 1195 (1992)
16. Herstein, I.N.: *Topics in Algebra*. Blaisdell Publishing Company, New York (1964)
17. Tsallis, C.: Possible generalization of Boltzmann-Gibbs entropy. *J. Stat. Phys.* **52**, 479 (1988)