A graph-theoretical kinetic Monte Carlo framework for on-lattice chemical kinetics

Michail Stamatakis and Dionisios G. Vlachos[a)]
Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716, USA

(Received 3 March 2011; accepted 12 May 2011; published online 7 June 2011)

Existing kinetic Monte Carlo (KMC) frameworks for the simulation of adsorption, desorption, diffusion, and reaction on a lattice often assume that each participating species occupies a single site and represent elementary events involving a maximum of two sites. However, these assumptions may be inadequate, especially in the case of complex chemistries, involving multidentate species or complex coverage and neighboring patterns between several lattice sites. We have developed a novel approach that employs graph-theoretical ideas to overcome these challenges and treat easily complex chemistries. As a benchmark, the Ziff-Gulari-Barshad system is simulated and comparisons of the computational times of the graph-theoretical KMC and a simpler KMC approach are made. Further, to demonstrate the capabilities of our framework, the water-gas shift chemistry on Pt(111) is simulated. © *2011 American Institute of Physics*. [doi:10.1063/1.3596751]

I. INTRODUCTION

Since the advent of the lattice kinetic Monte Carlo (KMC) approach, attributed to Bortz *et al.*,[1] numerous studies have employed KMC simulation for catalytic systems. KMC studies have focused on characterizing non-equilibrium (kinetic) phase transitions in model as well as realistic systems, understanding the effect of adsorbate-adsorbate interactions on surface coverage and reaction rates,[2–6] or more recently investigating diffusion in zeolites[7,8] and detailed chemistries using kinetic rates obtained from *ab initio* or first-principles calculations.[9–18] The transition probabilities for various elementary reaction steps were generalized in Reese *et al.*[19] Reviews of KMC simulation methodologies appear in Refs. 20 and 21, whereas accelerated and coarse-grained methods have recently been reviewed in Ref. 22.

Simulation of more complex reaction mechanisms using existing KMC frameworks, can produce inaccurate results, due to the simplistic representation of adsorbate binding and elementary reactions. It is often assumed that an adsorbate occupies a single site; this approach cannot account for the multidentate nature of several adsorbates. Further, previous studies are frequently limited to reactions that involve 1 or 2 sites.[3,4,14,15,23,24] On the other hand, the multisite approach of Hansen and Neurock[9,11] entails reaction detection on the basis of at most two sites, with the possibility of reactants appearing in reflected or intermediate sites. Yet, this approach is inadequate in cases where products appear in sites other than those just mentioned; in such cases, specific spatial arrangements of multidentate reactant species need to be explicitly considered. Such challenges call for a KMC framework, capable of efficiently accounting for these complexities.

In this paper, we present a graph-theoretical KMC approach that provides freedom and specificity in defining elementary events, thereby being capable of capturing more complex and realistic processes. The generality of the formalism makes this approach applicable to a vast array of lattice chemical kinetics, and the use of efficient data structures and algorithms makes the computational overhead comparable to that of simpler KMC approaches. In the following, we discuss this framework in detail using simple examples to illustrate the key ideas and procedures. We further simulate a prototype system devised by Ziff *et al.*[25] in order to compare the results and the performance of our approach with a simpler KMC capable of simulating 1- and 2-site processes only. Finally, we demonstrate the capabilities of the graph-theoretical KMC by simulating the water-gas shift chemistry on the Pt(111) surface using mainly density functional theory (DFT) input.

II. METHODS

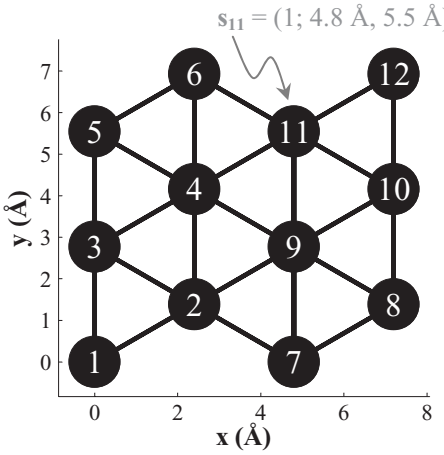
In our discussion of the graph-theoretical KMC approach, we will use a simple prototype model to illustrate the underlying ideas and procedures. Thus, consider a catalytic surface that is initially partially covered by species A* and B*. These species can diffuse and react with each other, thereby forming a bidentate species AB**. The latter can leave the surface by desorption.

A. Lattice representation

In order to represent the catalytic surface we assume that there exist well-defined sites on which the adsorbates A*, B*, and AB** are bound. Each one of these sites is defined on the basis of its type and position. Thus, each site s_i is represented by a three-element vector: the first element denotes the site type, an integer ranging from 1 to S_T, and the subsequent elements the site's x- and y-coordinates,

$$\mathbf{s_i} \in \{1, 2, \ldots, S_T\} \times \mathbb{R}^2, \qquad i \in \mathcal{S} = \{1, 2, \ldots, S_L\},$$
$$\mathbf{s} = \left\{(s_{i,j})_{j=1}^{3}\right\}_{i=1}^{S_L}, \tag{1}$$

[a)]Author to whom correspondence should be addressed. Electronic mail: vlachos@udel.edu. Tel.: 302-831-2830.



Site Number	Number of Neighbors	List of Neighbors					
1	6	2	3	5	6	12	8
2	6	1	4	3	9	7	6
3	6	4	5	1	2	8	10
4	6	3	6	5	11	9	2
5	6	6	1	3	4	10	12
6	6	5	2	1	7	11	4
7	6	8	9	11	12	6	2
8	6	7	10	9	3	1	12
9	6	10	11	7	8	2	4
10	6	9	12	11	5	3	8
11	6	12	7	9	10	4	6
12	6	11	8	7	1	5	10

FIG. 1. Illustration of a small periodic lattice, with a single site type and 6-fold coordination. Site 11 is represented by the vector $\mathbf{s_{11}}$ = (site type; x-coordinate, y-coordinate) = (1; 4.8 Å, 5.5 Å). The table on the right shows the adjacency list of the lattice graph. Thus, the neighbors of site 3 are sites 1, 2, 4, 5, as well as 8 and 10 (due to the periodicity).

where $\mathcal{S}$ is the index set of all sites on the lattice and S_L denotes the number of these sites. The lattice is then represented as a graph $\mathcal{L}$ in which each vertex is a site index and each edge introduces a neighboring relation between two sites,

$$\mathcal{L} = (\mathcal{S}, \mathcal{E}), \tag{2}$$

where $\mathcal{E}$ contains two element subsets of $\mathcal{S}$. Given the graph $\mathcal{L}$, one can form the adjacency list of the graph, which encodes the lattice structure as a series of lists of all the neighbors of each and every site. Note that even though all lattice sites are on a plane, the graph is not necessarily planar, since the edges may intersect at points other than the lattice sites.

Figure 1 presents a small periodic lattice with a single site type and 6-fold coordination. In line with Eq. (1), site 11 is represented as (1; 4.8 Å, 5.5 Å). The table on the right of this figure shows the adjacency list for this lattice graph. For instance, site 3 has a total of six neighbors, namely, sites 1, 2, 4, 5, as well as 8 and 10 (due to the periodicity).

B. State of the system

Having specified a lattice, we now need to define a state variable, which contains all information about where each adsorbate is located. Thus, we need to know whether an adsorbate binds to a single site (monodentate species) or more (multidentate species). In the latter case, we will need to specify the sites occupied by the adsorbate as well as the orientation of the molecule.

Thus, suppose that there exist N_S different surface species. The number of sites that species k occupies is denoted by d_k and can take non-negative integer values,

$$\mathbf{d} \in \{1, 2, \ldots\}^{N_S}. \tag{3}$$

If k is a monodentate species, then $d_k = 1$; for multidentate species $d_k > 1$. For instance, in our prototype example, there are three surface species (A*, B*, and AB**), the first two monodentate and the third bidentate, so that $\mathbf{d} = [1\ 1\ 2]$.

At any particular time instance, a lattice site can be either unoccupied or occupied. In the former case, the site is considered to be "occupied" by a "free site" species, also referred to as a vacancy. In the latter case, one needs to know which adsorbate is bound to which site and with what configuration. Consequently, at any time the surface sites can be occupied by S_L entities at maximum, an entity being a distinct free site or adsorbate. We thus label each entity that exists on the lattice by an integer and represent the state of each site with a three element vector $\boldsymbol{\sigma_i}$: the first element provides the entity label, the second gives the species number, and the third gives the subunit number that occupies the site in consideration. The state of the system is then given by an $S_L \times 3$ array $\boldsymbol{\sigma}$,

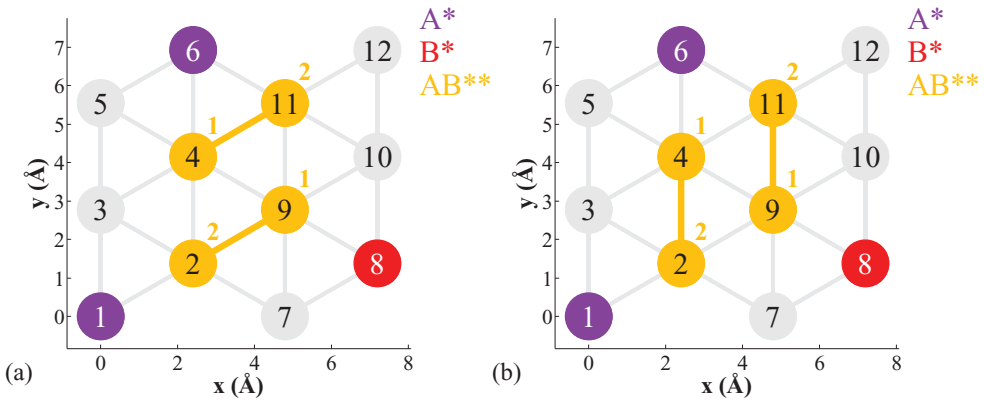
$$\begin{aligned} \boldsymbol{\sigma_i} &\in \{1, 2, \ldots, S_L\} \times \{0, 1, \ldots, N_S\} \\ &\quad \times \{1, 2, \ldots, \max(\mathbf{d})\} \quad i \in \mathcal{S} \\ \boldsymbol{\sigma} &= \left\{(\sigma_{i,j})_{j=1}^{3}\right\}_{i=1}^{S_L}. \end{aligned} \tag{4}$$

We further introduce an inverse mapping which, for every entity η, gives the species and the sites occupied by this entity,

$$\begin{aligned} \omega_\eta &= [\theta; \mathbf{v}] \text{ s.t. } \sigma_{v_j,1} = \eta,\ \sigma_{v_j,2} = \theta, \\ \sigma_{v_j,3} &= j,\ \forall\, j = 1, \ldots, d_\theta, \end{aligned} \tag{5}$$

where η is the entity number, θ the corresponding species, and $\mathbf{v}$ the vector with the sites occupied.

In order to clarify the purpose of labeling the entities that exist on the lattice, recall our prototype model, two possible configurations of which are portrayed in Fig. 2. In the two lattice plots, sites are either empty (grey) or occupied by the species A*, B*, and AB** (purple, red, and orange). The difference between the two lattice configurations is only in the orientation of the two adsorbates; yet, species and dentate information is identical in both cases. In order to distinguish the two configurations, all entities are labeled so that A* occupying site 1 is entity 1 (see Table c in Fig. 2), the adsorbate AB** occupying sites {9, 2} is entity 2, empty site 3 is entity 3, AB** occupying sites {4, 11} is entity 4, etc. Thus, the pairs of sites {9, 2} and {4, 11} are occupied by the same entity in Fig. 2(a), whereas in Fig. 2(b), it is the pairs {4, 2} and {9, 11} that are occupied by the same entity. Hence, the state of the system in each case has been unambiguously specified.



(c)

Site Number	Entity Number	Species	Dentate
1	1	1	1
2	*2*	3	1
3	3	0	1
4	*4*	3	2
5	5	0	1
6	6	1	1
7	7	0	1
8	8	2	1
9	*2*	3	2
10	9	0	1
11	*4*	3	1
12	10	0	1

(d)

Site Number	Entity Number	Species	Dentate
1	1	1	1
2	*2*	3	1
3	3	0	1
4	*2*	3	2
5	5	0	1
6	6	1	1
7	7	0	1
8	8	2	1
9	*4*	3	2
10	9	0	1
11	*4*	3	1
12	10	0	1

FIG. 2. The state of each site is completely defined by three variables: species number and dentate bound to that site as well as entity number. One can see why an explicit numbering of the entities is required by considering the examples of panels (a) and (b). Each circle represents a single site; grey circles are vacant sites; colored circles represent sites occupied by the species A*, B*, and AB** (numbered 1 to 3, respectively). Site numbers are shown in the center of the circles, and dentate numbers on a top right position outside the circles. In both cases, the species and dentate information is exactly the same; what changes is the orientation of the two AB** adsorbates. Thus, in the lattice of panel (a), sites 2 and 9 are occupied by entity 4 and sites 4 and 11 by entity 10. In panel (b), however, sites 2 and 4 are occupied by entity 4 and sites 9 and 11 by entity 10. (c, d) Array $\boldsymbol{\sigma}$ for the two configurations of panels (a, b), respectively.

Note that the state specification is unique up to a permutation of the labels of the entities; thus, the adsorbate AB** occupying sites {9, 2} in Fig. 2(a) could have been relabeled as the 10th entity and the empty site 12 as the 2nd entity without changing the observed lattice configuration.

Note that the two tables in Fig. 2 show array $\boldsymbol{\sigma}$ for each configuration (Eq. (4)). The inverse mapping (Eq. (5)) would consist of the species and site numbers pertinent to each entity. Thus, for the configuration of Fig. 2(a), $\omega_1 = [1;\ 1]$, $\omega_2 = [3;\ 2,\ 9]$, $\omega_3 = [0;\ 3], \ldots, \omega_{10} = [0;\ 12]$.

C. Elementary step representation

The state of the lattice can change through the occurrence of adsorption, desorption, reaction, or diffusion events. Thus, we need a way to represent such events in a general way.

Each event, referred to as an elementary step, is thus represented by a connected graph with specified site types as well as initial and final coverage patterns. For our purposes, no position data are required for the sites of the elementary reaction graphs. Thus, one does not need to define multiple orientations for a reaction pattern as done in previous KMC frameworks (for instance in the CARLOS code, Ref. 23); as long as the neighboring of the entities on the lattice is similar to that of the elementary reaction pattern, the latter will be detected as a legitimate lattice process. It is still straightforward to add spatial information in the framework if required for a particular application. Thus, if elementary step k involves $S_{R,k}$ sites, each site will be assigned a site type (similarly to Eq. (1)),

$$\xi_{k,i} \subseteq \{0, 1, 2, \ldots, S_T\}, \qquad i \in \Xi_k = \{1, 2, \ldots, S_{R,k}\}, \tag{6}$$

where site type 0 means that a site of any type can be involved in the elementary step k. Then, one defines the graph of this elementary step as follows:

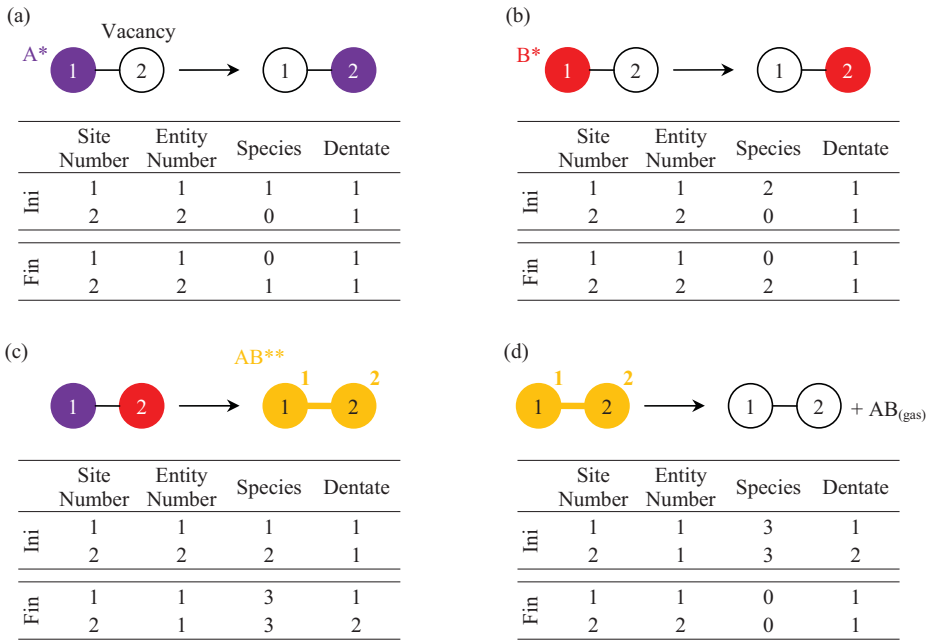
$$\mathcal{R}_k = (\Xi_k, \mathcal{E}_k). \tag{7}$$

Furthermore, the initial and final coverage patterns of elementary step k can be defined as the states of the elementary step graph in accordance with Eq. (4),

$$\sigma_{k,i}^{ini}, \sigma_{k,i}^{fin} \in Z \subseteq \{1, 2, \ldots, S_{R,k}\} \times \{0, 1, \ldots, N_S\} \times \{1, 2, \ldots, \max(\mathbf{d})\}, \quad i \in \Xi_k. \tag{8}$$

Note that an elementary reaction affects the coverage of the corresponding sites but not their types or neighboring structure; in other words, this formalism does not account for catalyst reconstruction.

Going back to our prototype model, the four elementary events it entails are shown in Fig. 3: diffusion of species A*,



(a)	Site Number	Entity Number	Species	Dentate
Ini	1	1	1	1
	2	2	0	1
Fin	1	1	0	1
	2	2	1	1

(b)	Site Number	Entity Number	Species	Dentate
Ini	1	1	2	1
	2	2	0	1
Fin	1	1	0	1
	2	2	2	1

(c)	Site Number	Entity Number	Species	Dentate
Ini	1	1	1	1
	2	2	2	1
Fin	1	1	3	1
	2	1	3	2

(d)	Site Number	Entity Number	Species	Dentate
Ini	1	1	3	1
	2	1	3	2
Fin	1	1	0	1
	2	2	0	1

FIG. 3. The graph representation of each of the four elementary steps (panels a–d) for the prototype model. The tables below each of the elementary step-schematics show the initial and final coverage patterns (Eq. (8)) marked as "Ini" and "Fin."

diffusion of B*, reaction between A* and B* to form the bidentate species AB**, and desorption of the latter leaving two empty sites. For each of these elementary steps, the initial and final coverage patterns $\sigma_{k,i}^{ini}$, $\sigma_{k,i}^{fin}$ (see Eq. (8)) are shown in the tables below the schematics.

In order to demonstrate the extensive capabilities of the graph-theoretical KMC approach in defining elementary events, we also consider an example of an elementary reaction involving two bidentate species (Fig. 4(a)). For simplicity, in this example a single site type is considered. This reaction would be conventionally represented as A** + B** → C** + D**; however, this representation does not convey information about the specific configuration that would result in the realization of this reaction. On the other hand, the graph representation shown in Fig. 4 explicitly requires that dentate 2 of A** neighbors with dentate 1 of B** for the reaction to proceed. Thus, only two out of the three neighboring A** and B** pairs appearing in Fig. 4(b) can react. The sequence of elementary reaction events appears in Fig. 4(b)–4(d). The upper left pair reacts first (c) followed by the pair on the right (d). After these two events, no more reactions can occur. Finding which lattice adsorbates can react according to a given elementary reaction patterns is a core procedure of the graph theoretical lattice KMC. This procedure is described in detail in Sec. II D, where we also discuss the sequence of elementary events shown in Figs. 4(e) and 4(f).

Thus, the initial and final coverage information (Eq. (8)) along with the neighboring structure of the sites involved (Eq. (7)), and the kinetic constant, completely specify the elementary step. The kinetic rate constant just mentioned can be calculated according to (classical) transition state theory as[14,15,26,27]

$$k_{TST} = \frac{Q^{\neq}}{Q_R}\frac{k_B T}{h}\exp\left(-\frac{\Delta E^{\neq}}{k_B T}\right), \tag{9}$$

where h is Planck's constant; k_B is Boltzmann's constant; T the temperature; $Q^{\neq}$ and Q_R are the partition functions of the activated complex and reactants, respectively; and $\Delta E^{\neq}$ is the energy barrier (total energy of the transition state minus that of the initial state). For immobile adsorbed species, the partition function has only vibrational components, whereas for gas-phase species, translational and rotational degrees of freedom have to be included as well,

$$\begin{aligned} Q_{surf} &= q_{vib}, \\ Q_{gas} &= q_{trans,3D} q_{rot} q_{vib}. \end{aligned} \tag{10}$$

Expressions for these components are provided from statistical mechanics textbooks[28] and the information needed for computing them can be obtained from quantum chemistry calculations. Finally, the partition function of the reactants Q_R is by assumption equal to the product of the partition functions of each reactant species. Thus, one can calculate the rate of any elementary step that involves gas and surface species. For instance, for an Eley Riedel reaction, $X_{(gas)} + Y^* \rightarrow Z^*$, the rate would be (see Sec. 1 of Ref. 29),

$$\begin{aligned} k_{Eley-Riedel} = &\frac{q_{vib}^{\neq}}{q_{vib,X_{(gas)}} q_{rot,X_{(gas)}} q_{trans2D,X_{(gas)}} q_{Y,vib}} \\ &\times \frac{p_X A_{st}}{\sqrt{2\pi m_X k_B T}} \exp\left(-\frac{\Delta E^{\neq}}{k_B T}\right). \end{aligned} \tag{11}$$

The partition functions can be calculated from standard statistical mechanical expressions.[28] A more detailed discussion of the rates for several different types of elementary steps is given in Sec. 1 of Ref. 29.

To account for lateral interactions, the forward and backward activation energies can be parameterized in terms of the coverage in the neighborhood of the reaction pattern. This can

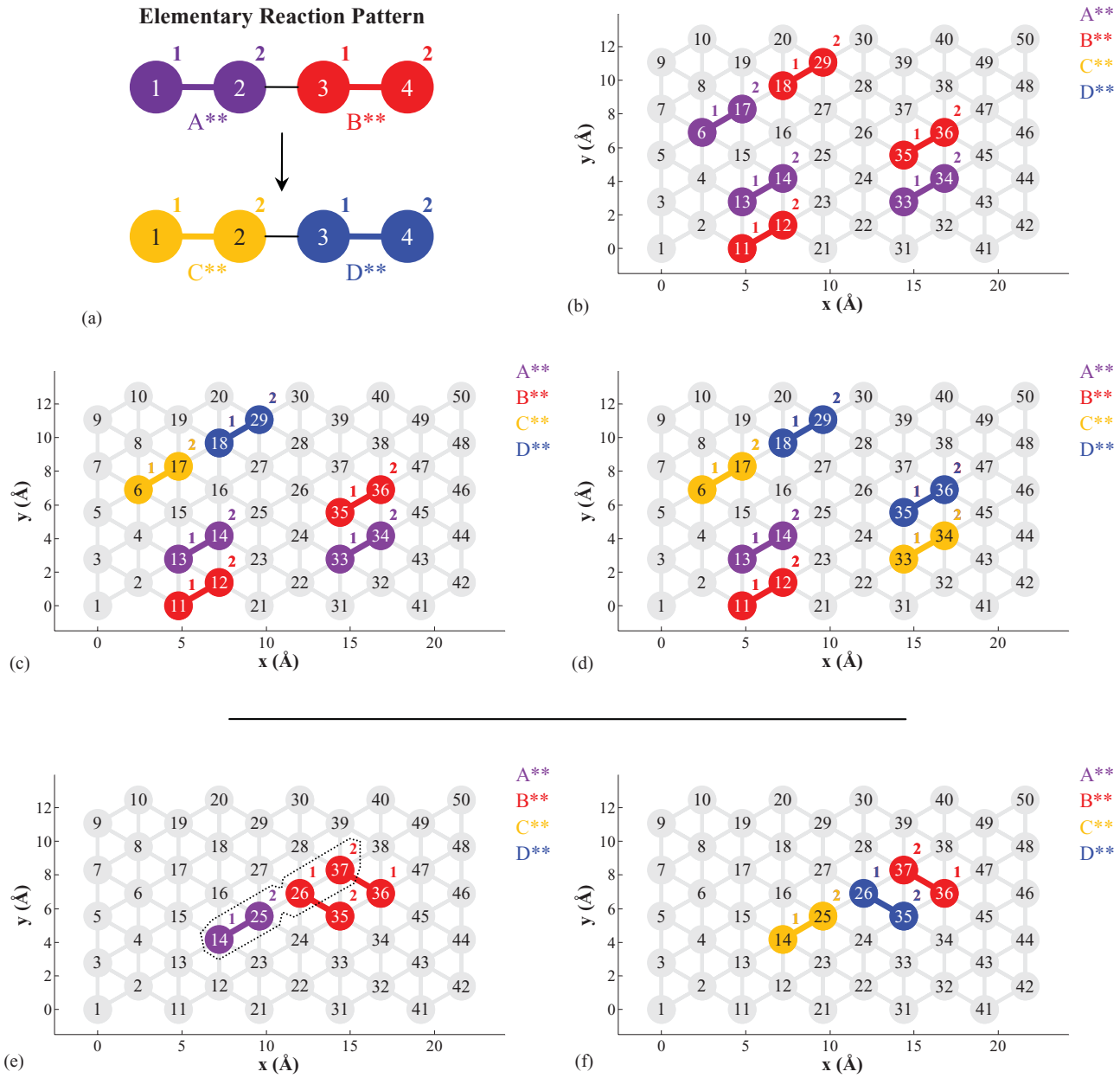


FIG. 4. Example demonstrating the representation of elementary steps as graphs and their mapping to lattice processes. (a) The elementary step involves two bidentate adsorbates, species A** and B**, that get converted to species C** and D**. This is a four-site process (site labeling appears in the top-right corner outside the circles representing the sites). Dentate numbering appears in the center of the circles. (b), (c), (d) Two out of the three neighboring pairs of A** and B** comply with the pattern of panel (a). The top left pair reacts first (c) followed by the pair on the right (d). (e, f) For the starting configuration shown in panel (e), a single reaction event can occur (panel f).

be done using cluster expansion techniques.[30–32] For simplicity, in practice one avoids expanding the activation energies; rather, a cluster expansion of the energy of the system is introduced from which the difference in the initial and final state energies can be deduced. The latter difference is the reaction energy,

$$\Delta E_{rxn}(\boldsymbol{\sigma}) = \Delta E_{rxn,0}(\boldsymbol{\sigma}) + F(\boldsymbol{\sigma}), \tag{12}$$

where the $\Delta E_{rxn,0}$ term includes no lateral interactions (as if the entities in the lattice were infinitely separated), and $F(\boldsymbol{\sigma})$ is the lateral interaction term computed through the cluster expansion. The simplest cluster expansion is the Ising model with nearest-neighbor pairwise additive iteractions,

$$\Delta E_{rxn}(\boldsymbol{\sigma}) = \sum_{i=1}^{S_L} h_i \sigma_i + \underbrace{\frac{1}{2}\sum_{i=1}^{S_L}\sum_{\substack{j=1\\ j\neq i}}^{S_L} J_{i,j}\sigma_i\sigma_j}_{F(\boldsymbol{\sigma})}. \tag{13}$$

Then, the forward and backward activation energies can be calculated as follows:

$$\begin{aligned} \Delta E^{\neq}_{fwd}(\boldsymbol{\sigma}) &= \Delta E^{\neq}_{fwd,0} + (1-\omega)\cdot F(\boldsymbol{\sigma}), \\ \Delta E^{\neq}_{bwd}(\boldsymbol{\sigma}) &= \Delta E^{\neq}_{bwd,0} - \omega \cdot F(\boldsymbol{\sigma}), \end{aligned} \tag{14}$$

where ω is a proximity factor,[33] expressing how reactant- or product-like the transition state is. For the extreme values $\omega = 0$ or $\omega = 1$, the transition state is reactant-like or product-like, respectively.

D. Mapping to lattice processes

Having defined the elementary steps, the basic idea is then to solve a subgraph isomorphism problem for each elementary step defined, in order to identify the lattice processes that can take place. Thus, in order to find feasible lattice processes corresponding to elementary step k, one needs to find mappings of the elementary step sites to the lattice sites, such that the site types and coverages of elementary step k match those of the lattice. More specifically, a lattice process is a mapping between the vertex sets of $\mathcal{R}_k$ (the subgraph) and $\mathcal{L}$ (the "large" graph),

$$\mathcal{M} : \Xi_k \rightarrow \mathcal{S}, \tag{15}$$

such that,

1. $\mathcal{M}$ is a subgraph isomorphism, namely for every pair p, q of neighboring sites in the edge set of the pattern $\mathcal{E}_k$ (see Eq. (7)), there exists a pair $\mathcal{M}(p)$, $\mathcal{M}(q)$ of neighboring sites in $\mathcal{E}$ of $\mathcal{L}$, namely in the lattice,

$$\underset{\substack{1\le p\le S_{R,k}\\ 1\le q\le S_{R,k}}}{\forall} \{p, q\} \in \mathcal{E}_k \Rightarrow \{\mathcal{M}(p), \mathcal{M}(q)\} \in \mathcal{E}, \tag{16}$$

2. the site types of p on the pattern and $\mathcal{M}(p)$ on the lattice are the same,

$$s_{\mathcal{M}(p),3} = \xi_{k,p} \qquad 1 \le p \le S_{R,k}, \tag{17}$$

3. there is a mapping between the elementary step entities and the lattice entities,

$$\mathcal{F} : \{1, 2, \ldots, N_{RE}\} \rightarrow \{1, 2, \ldots, N_E\}, \tag{18}$$

such that,

$$\sigma_{\mathcal{M}(p),1} = \mathcal{F}\left(\sigma^{ini}_{k,p,1}\right) \qquad 1 \le p \le S_{R,k}, \tag{19}$$

namely the coverage patterns of the elementary step and the lattice match.

Note that there are as many such mappings $\mathcal{M}$ (Eq. (15)) as lattice processes.

To exemplify the mapping of elementary reaction patterns to lattice processes consider the configuration shown in Fig. 2(a). Table I lists the twelve possible lattice processes according to the elementary step patterns of Fig. 3, showing the mappings between sites and entities. For instance, process number 5 (A + B reaction) involves sites 1 and 8; thus, sites 1 and 2 of pattern (c) in Fig. 3 are mapped to lattice sites 8 and 10. This process requires the participation of two lattice entities belonging to species A* and B*, respectively. Thus, pattern entities 1 and 2 are mapped to entities 8 and 10. As another example, consider process 12 (AB desorption), which involves lattice sites 4 and 11, occupied by lattice entity 4 (species AB**). Thus, the mappings in this case are: sites mapping $\{1, 2\} \rightarrow \{4, 11\}$ and entity mapping $\{1\} \rightarrow \{4\}$.

TABLE I. List of processes for the prototype model in the configuration of Fig. 2(b).

Process number	Elementary step	Sites mapping $\mathcal{M}$	Entities mapping $\mathcal{F}$
1	AB desorption	$\{1, 2\} \rightarrow \{9, 2\}$	$\{1\} \rightarrow \{2\}$
2	A diffusion	$\{1, 2\} \rightarrow \{1, 3\}$	$\{1, 2\} \rightarrow \{1, 3\}$
3	A diffusion	$\{1, 2\} \rightarrow \{1, 5\}$	$\{1, 2\} \rightarrow \{1, 5\}$
4	A diffusion	$\{1, 2\} \rightarrow \{1, 12\}$	$\{1, 2\} \rightarrow \{1, 10\}$
5	A + B reaction	$\{1, 2\} \rightarrow \{1, 8\}$	$\{1, 2\} \rightarrow \{1, 8\}$
6	A diffusion	$\{1, 2\} \rightarrow \{6, 5\}$	$\{1, 2\} \rightarrow \{6, 5\}$
7	A diffusion	$\{1, 2\} \rightarrow \{6, 7\}$	$\{1, 2\} \rightarrow \{6, 7\}$
8	B diffusion	$\{1, 2\} \rightarrow \{8, 7\}$	$\{1, 2\} \rightarrow \{8, 7\}$
9	B diffusion	$\{1, 2\} \rightarrow \{8, 10\}$	$\{1, 2\} \rightarrow \{8, 9\}$
10	B diffusion	$\{1, 2\} \rightarrow \{8, 3\}$	$\{1, 2\} \rightarrow \{8, 3\}$
11	B diffusion	$\{1, 2\} \rightarrow \{8, 12\}$	$\{1, 2\} \rightarrow \{8, 10\}$
12	AB desorption	$\{1, 2\} \rightarrow \{4, 11\}$	$\{1\} \rightarrow \{4\}$

To further demonstrate the specificity in the detection of patterns let us consider the two cases shown in Fig. 4. For the configuration shown in Fig. 4(b), only two out of the three A** and B** pairs can react, as already discussed in Sec. II C. For these two, the mappings of elementary reaction sites to lattice sites are $\{1, 2, 3, 4\} \rightarrow \{6, 17, 18, 29\}$ and $\{1, 2, 3, 4\} \rightarrow \{33, 34, 35, 36\}$. Each of these mappings is a lattice process and has a probability of occurrence in the next dt time interval. Eventually both processes occur (Figs. 4(c) and 4(d)) and the reactant species A** and B** are substituted by the products C** and D** on the lattice. Note that the adsorbates on sites $\{13, 14\}$ and $\{11, 12\}$ cannot react since their orientation is not compatible with the elementary reaction pattern; in other words, there does not exist a mapping $\mathcal{M}$ that satisfies Eq. (16).

Further, consider another configuration shown in Fig. 4(e), that demonstrates the importance of the third condition for the mapping to exist (Eqs. (18) and (19)). Suppose that species A** is entity 1 and B** is entity 2 in the elementary step. Also, suppose that the adsorbate occupying the pair of lattice sites $\{14, 25\}$ is labeled as entity 26, the one on sites $\{26, 35\}$ is entity 36 and the one on $\{36, 37\}$ is entity 47. Obviously, a lattice process involving sites $\{14, 25, 26, 37\}$ cannot exist since sites 26 and 37 are occupied by different adsorbates. However, if one was to check only the site type, species, and dentates occupying each of these sites, then one would have concluded incorrectly that the elementary reaction could take place. The requirement that a mapping between elementary step entities and lattice entities exist prevents such problems. In particular, for sites $\{14, 25, 26, 37\}$ such a mapping does not exist, since there is an ambiguity as to where to map entity 2 (B**) of the elementary step. On the other hand, such a mapping between entities exists for the lattice process defined as $\mathcal{M}([1\ 2\ 3\ 4]) = [14\ 25\ 26\ 35]$. The mapping between entities in this case is: $\mathcal{F}([1\ 2]) = [26\ 36]$, and the lattice process takes place in the next Monte Carlo step as shown in Fig. 4(f).

E. Event statistics

Whenever a mapping (15) is found, there is the chance that elementary event i will happen sometime in the future. In the general case, in which the rates are time-dependent, the inter-arrival time for event i is calculated as[34]

$$p(\tau_i) = k_i^{TST}(t+\tau_i)\exp\left[-\int_0^{\tau_i} k_i^{TST}(t+\tau_i')d\tau'\right]. \quad (20)$$

Thus, in order to calculate a time for which the next event will happen, one has to solve the equation,

$$\int_0^{\tau_i} k_i^{TST}(t+\tau')d\tau' + \ln(1-u) = 0, \quad (21)$$

where u is a uniformly distributed random number. Note that if the rate is not time dependent, the inter-arrival time follows the exponential distribution,

$$\tau_i \sim \mathcal{E}xp\left(k_i^{TST}\right). \quad (22)$$

Thus, in order to generate a random inter-arrival time, one uses the mapping,

$$\tau_i = -\frac{1}{k_i^{TST}}\ln(1-u), \quad (23)$$

where u is a random number sampled from the continuous uniform distribution: $u \sim \mathcal{U}(0,1)$.

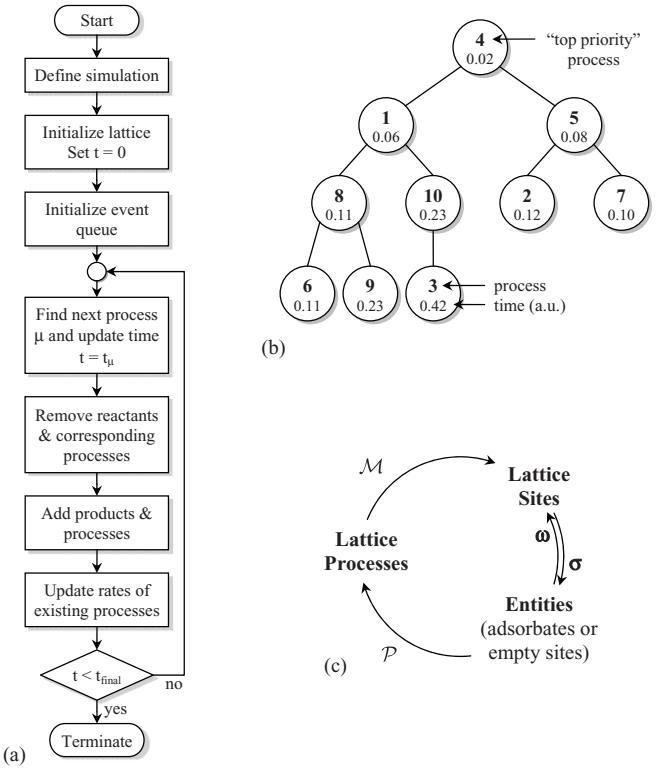


FIG. 5. (a) Flowchart of the KMC algorithm. (b) A binary heap structure in which the occurrence times for all processes are stored. (c) The flow of information in the KMC algorithm. Structures σ, ω, $\mathcal{M}$, and $\mathcal{P}$ are explained in the text (Eqs. (4), (5), (15), and (24) respectively).

F. Pseudocode

The graph-theoretical KMC method is outlined in the following pseudocode and schematically shown in Fig. 5(a). An open source freely distributed FORTRAN implementation can be found at our webpage.[35]

0. Start
1. Define simulation lattice (Eqs. (1) and (2)), conditions, participating species and elementary steps (Eqs. (6)–(9))
2. Initialize the lattice state (Eqs. (4) and (5)) and set the time clock to $t = 0$
3. Find all elementary events that can happen (Eq. (15)) and generate a random time at which they will take place (Eq. (21) or, for constant rates, Eq. (23)); put these times in an event-queue
4. While $t < t_{final}$
   - 4a. Find the process μ that will occur next and update the time
   - 4b. Remove the reactants from the lattice as well as all the processes which they participate in from the event-queue
   - 4c. Add the products of process μ in the lattice and find elementary events in which products participate; for each of these events calculate a random inter-arrival time and include it in the event-queue
   - 4d. Update the rates of existing processes in the case where energetic interactions exist
5. Repeat
6. **Terminate**

In order to minimize the computational overhead, several optimization strategies were used. Absolute reaction times were used (as opposed to time increments with respect to the current t), and these times were stored in an index priority queue, an optimization scheme that has also been used in well-mixed KMC algorithms.[36] This form of queue was implemented as a binary heap structure (chapter II.6 in Ref. 37), an example of which appears in Fig. 5(b). The latter is a special type of a complete binary tree with the property that a node has "priority" over all its children. In our case, each node contains information about the time at which each lattice process occurs, and thus, the top node always returns the time the process that will be simulated next (in step 4.a). In other words, step 4.a requires constant (computational) time whereas removing (adding) an element from (into) the heap during steps 4.b and 4.c has complexity $\mathcal{O}(\log(n))$, with n being the number of processes stored in the queue at a given instance. Thus, in the example of Fig. 5(b), there are 10 processes stored in the queue, and the 4th one is about to occur in 0.02 time units.

Moreover, the algorithm uses a 2-dimensional array $\mathcal{P}$ that stores the processes in which each entity participates. Thus, for entity η, which may be an adsorbate or an empty site, $\mathcal{P}_{\eta,0}$ gives the number of processes in which this entity is involved and $\mathcal{P}_{\eta,m}$, $m = 1,\ldots,\mathcal{P}_{\eta,0}$ gives these processes. Note that $\mathcal{P}$ satisfies,

$$\sigma_{s,1} = \eta \quad \forall \text{ site } s \in \mathcal{M}(\Xi_k)_{\mathcal{P}_{\eta,m}} \quad \forall\, m = 1,\ldots,\mathcal{P}_{\eta,0}, \quad (24)$$

TABLE II. The $\mathcal{P}$ structure for the processes of Table I.

Entity	Number of processes	List of processes
1	4	2, 3, 4, 5
2	1	1
3	2	2, 10
4	1	12
5	2	3, 6
6	2	6, 7
7	2	7, 8
8	5	5, 8, 9, 10, 11
9	1	9
10	2	4, 11

where k is the elementary step type for process $\mathcal{P}_{\eta,\mathrm{m}}$ and consequently s must be a lattice site occupied by the entity η. Therefore, using $\mathcal{P}$ with the lattice state array and the inverse mapping (Eqs. (4) and (5)), the algorithm can identify the processes in which each entity participates, as well as the lattice sites involved in each process (Fig. 5(c)). This scheme allows for the efficient retrieval of information pertaining to which processes each entity or site is involved in, thereby accelerating the procedures of steps 3, 4.b, 4.d, and 4.c. To exemplify the use of this data structure, Table II shows the information contained in $\mathcal{P}$ for the processes of Table I, pertaining to the configuration of Fig. 2(a).

Finally, for the elementary reaction pattern search of steps 3 and 4.c, Ullmann's algorithm for subgraph isomorphism was used.[38] The algorithm is applicable to non-planar graphs as required in the general case. In specific cases where the lattice can be represented by a planar graph, one can use even more efficient algorithms.[39] Furthermore, since in step 4.c the only change in the lattice state pertains to the addition of the products of process μ, the elementary reaction pattern search is further facilitated by searching only for the elementary events in which each product participates locally in the neighborhood of the newly added entity. This is achieved by building two data structures right after step 1: (i) a dependency array that contains the elementary events in which each species participates; this way only the relevant elementary events are considered in step 4.c. (ii) An array that gives the "pattern level" for each elementary event and for each entity thereof. The "pattern level" $\ell_{\mathrm{k,j}}$ for entity j participating in elementary event k is expressed as follows:

$$\ell_{\mathrm{k,j}} = \max_{\mathrm{p}\in\Xi_{\mathrm{k}}:\ \sigma^{\mathrm{ini}}_{\mathrm{k,p,1}}\neq \mathrm{j}} \ \min_{\mathrm{q}\in\Xi_{\mathrm{k}}:\ \sigma^{\mathrm{ini}}_{\mathrm{k,q,1}}=\mathrm{j}} d_{\mathcal{R}_{\mathrm{k}}}(\mathrm{p},\mathrm{q}), \tag{25}$$

where $d_{\mathcal{R}_k}(\mathrm{p},\mathrm{q})$ denotes the graph distance between vertexes p and q of graph Ξ_{k}. Thus, the pattern level has the following meaning: starting from the sites that entity k occupies in pattern k, it gives the neighboring level of the farthest neighbor, where the level of the nearest neighbors is 1, that of the next nearest neighbors is 2 and so on. For instance, in the elementary step pattern shown in Fig. 6(a), if species A** and B* are entities 1 and 2, respectively, the pattern level with respect to entity 1 is $\ell_{\mathrm{k,1}} = 1$ since the farthest neighbor is site 3, whose minimum distance from A** is 1, the distance between the second dentate (site 2) and that site. On the other hand, for entity 2, $\ell_{\mathrm{k,1}} = 2$, since the farthest neighbor from site 3 (occupied by B*) is site 1. This way, the pattern search is localized, since in the implementation discussed only the relevant neighboring sites are considered in solving the subgraph isomorphism problem of step 4.c. As shown in Fig. 6(b), if adsorbate B* was a newly added entity, one has to consider candidate sites up to the next nearest neighbors, while solving the graph isomorphism problem to find whether the elementary reaction can be mapped to a lattice process. If, however, adsorbate A** was the newly added entity, one would have to search the nearest neighbors of sites 20 and 31 occupied by A** (Fig. 6(c)). In this case both searches will detect a lattice process that complies with the elementary reaction pattern of Fig. 6(a), which is executed at the next KMC step (Fig. 6(d)).

III. COMPUTATIONAL RESULTS AND DISCUSSION

A. Ziff-Gulari-Barshad model system

The Ziff-Gulari-Barshad (ZGB) system is a prototype surface reaction model, whose development was motivated by the investigation of kinetic phase transitions in the CO oxidation reaction.[25] Here we will apply the graph-theoretical KMC method to reproduce the results by Ziff *et al.*[25], thereby showing that for chemistries that contain only 1-site and 2-site elementary events, our method is equivalent to the simpler "traditional" KMC. Furthermore, we will investigate the performance by comparing the computational times of the two KMC implementations.

The reactions considered in the ZGB model are shown in Table III, where A, B, and C denote the CO, O, and CO_2 species. Note that the algorithm used by Ziff *et al.*[25] operates in discrete time; thus, in order to ensure the equivalence of our simulation setup with that of the cited paper, we need to cautiously assign the values of the kinetic rates.

Ziff *et al.*[25] assume that the gas-phase consist of a mixture of A and B_2 and no inert exists, so that the species partial pressures are related as $\mathrm{P_{B2}} = \mathrm{P_{Total}} - \mathrm{P_A}$; further $\mathrm{P_{B2}} = \mathrm{y_{B2}P_{Total}}$ and $\mathrm{P_A} = \mathrm{y_A P_{Total}}$, with y denoting molar fractions. Since the algorithm of the cited paper operates in discrete time, at every trial one selects randomly the type of molecule that hits the surface. That molecule is B_2 with probability proportional to $\mathrm{y_{B2}}$, and A with probability proportional to $\mathrm{y_A}$. Thus, in continuous time (our implementation), the propensities are taken to be proportional to $\mathrm{P_{B2}}$ and $\mathrm{P_A}$. The factor of 10 appearing in the first propensity is an arbitrary constant that scales time. Finally, the factor 1/4 that appears in the propensity for the two site events (B_2 adsorption and A* + B* reaction) is due to the 4-fold coordination of each site on the rectangular lattice. The kinetic constant of the A* + B* reaction is taken to be large compared to the other constants, in order to reproduce the instantaneous oxidation assumed in Ziff *et al.*[25]

The result of a KMC simulation for a particular value of $\mathrm{P_A}$ appears in Fig. 7(a), where the lattice state is plotted. Grey and black points denote molecules of A and atoms of B, respectively. For this parameter set, about half of the surface area is covered with B atoms and there exist some clusters of A. For higher values of $\mathrm{P_A}$, these clusters cover the entire surface area, in a transition that occurs discontinuously

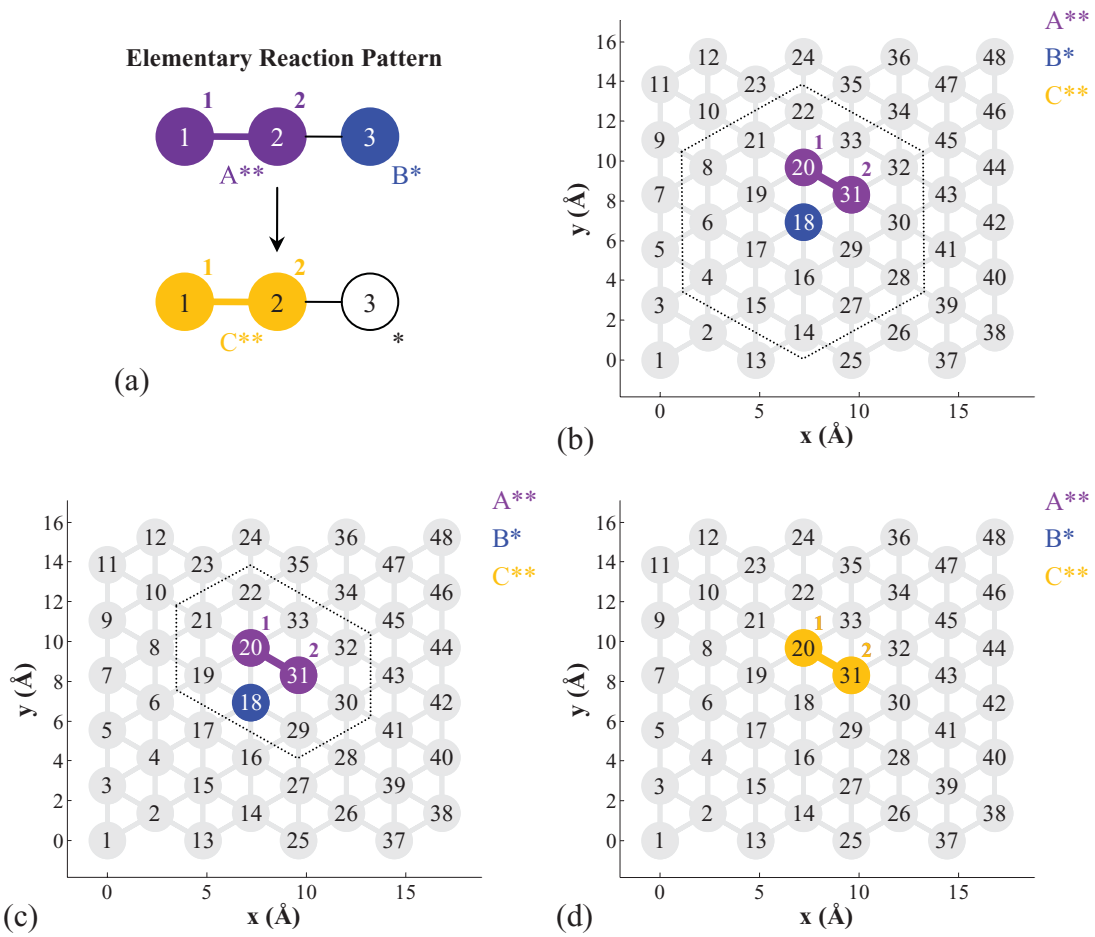


FIG. 6. Example demonstrating the pattern level concept and the localized pattern search. The elementary step shown in (a) involves a bidentate species A** that reacts with a monodentate species B* to produce species C** and an empty site. The pattern levels with respect to A** and B* are 1 and 2, respectively. Thus, if one searches for the pattern based on knowledge that B* occupies site 18, then one will have to consider only the sites enclosed in the hexagon of panel (b) for solving the graph isomorphism problem. On the other hand, if one knows that species B** occupies sites 20 and 31, one has to consider the sites marker in panel (c). In both cases a lattice process will be detected and executed at the next KMC step (d).

(Fig. 7(b)). The results of these simulations (Fig. 7(b)) are in perfect agreement with the results of the paper of Ziff *et al.*[25]

Finally, Fig. 7(d) shows a comparison of the computational times for the graph-theoretical KMC and a KMC that can only simulate 1- and 2-site processes. The latter implementation goes through each site and neighbors thereof upon initialization, and calculates the propensities of the 1- and 2-site events. The simulation proceeds similarly to Gillespie's direct method:[40] to determine when the next process will occur, an exponentially distributed random number r_1 is generated with rate parameter equal to the total sum of the propensities; to determine which process will occur, a uniform random number r_2 is generated and the partial sums of the propensities of processes 1 to k are evaluated. The first k for which the ratio of the partial sum over the total sum is greater than r_2 gives the process about to occur. For efficiency the propensities and their partial sums are stored in a binary tree. Thus, the searching for the process that will occur next is done in a divide-and-conquer approach that takes $\mathcal{O}(\log_2(N_{proc}))$ time (where N_{proc} is the number of individual lattice processes). More information about this scheme can be found in Appendix 7.2 by Gibson and Bruck[36] and Sec. 6.2 in Chatterjee and Vlachos.[22] Furthermore, once a lattice process has occurred, only the propensities of the sites involved and their neighbors are updated, thereby performing the minimum number of update operations.

TABLE III. Elementary events in the ZGB model.

	Elementary event	$k_{rxn}(s^{-1})$
1	$A_{(g)} + * \rightarrow A*$	$10P_A$
2	$B_{2(g)} + 2* \rightarrow 2B*$	$(1/4)10P_{B2}$
3	$A* + B* \rightarrow C_{(g)} + 2*$	$(1/4)10^5$

The simulations of Fig. 7(d) were performed in an Intel Core 2 Duo E8300 processor running at 2.83 GHz. The code was compiled using the Intel® Visual Fortran Compiler Professional Edition 11.1 using the "Maximize Speed" optimization option. Each run utilizes a single core (no parallelization) and the times reported are averages from triplicate runs. It is observed that the computational times of both KMC approaches scale approximately linearly with respect to the number of lattice sites for the same simulated time interval. In practice there is a small overhead associated with the search and update operations of larger binary trees, as evidenced by the exponents of the exponential fits, which are greater than

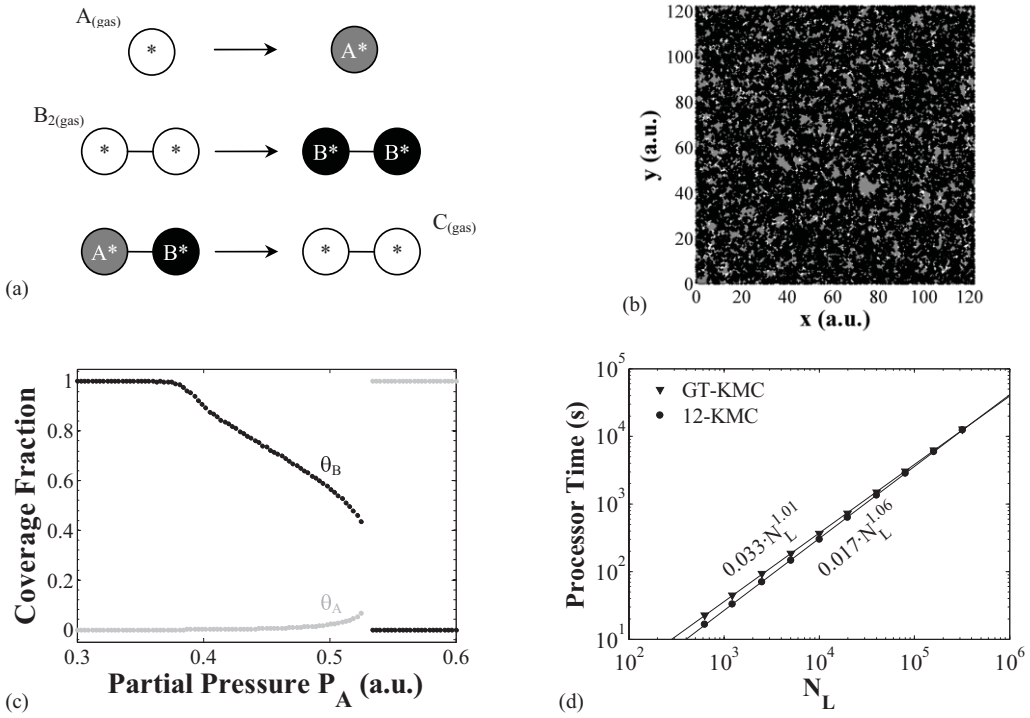


FIG. 7. (a) Elementary reaction patterns for the ZGB system. (b) A snapshot of 300 × 300 lattice at t = 100, for $P_A = 0.525$. (c) The phase diagram of coverage versus partial pressure of A, showing the kinetic phase transition at around $P_A = 0.53$. (d) Comparison of the computational times of the graph-theoretical KMC (GT-KMC) and a KMC that can simulate 1- and 2-site processes only (see text for more details), for the ZGB system with $P_A = 0.5$. The system was simulated for a total of 100 time units. The times are averages over three runs.

unity. This overhead is smaller in the graph-theoretical KMC where the search is performed in constant time, $\mathcal{O}(1)$, and the update takes at most $\mathcal{O}(\log_2(N_{proc}))$. Thus, as the lattice size increases, the graph-theoretical KMC progressively becomes as efficient as the simpler KMC implementation. The computational times for each run are also reported in Sec. 2 of Ref. 29.

B. Water-gas shift on Pt(111)

The water-gas shift (WGS) reaction has recently attracted attention as a means for producing H_2 from reforming of fossil fuels and biomass. Here we demonstrate how one can use our graph-theoretical KMC to study this reaction mechanism.

There are five gas species and eight surface species that participate in this chemistry, as shown in Table IV. Note that two surface species are multidentate: HCO binds to two sites (top-bridge) and HCOO occupies three sites (top-bridge-top). We assume that each of these species can be found in the preferred binding site of the Pt(111) surface, the lattice representation of which appears in Fig. 8(a). It is worth noting that DFT calculations predict that CO preferentially binds to the fcc hollow sites, even though experimental evidence suggests otherwise.[41] For our purposes, we assume that CO binds to the fcc sites.

Seventeen elementary steps are considered for this chemistry, outlined in Table V along with the corresponding preexponentials, activation, and reaction energies. In the calculation of the latter energies, CO-CO pairwise additive repulsive lateral interactions were taken into account between nearest neighbor fcc sites (see Eq. (13)), since this was observed to be the dominant species on the surface. Note that for some events, the reactants and products may appear in different site types and extra empty sites may also be involved, for example in the reaction between CO* and H* the product HCO** binds to top-bridge configuration on the surface, whereas CO* and H* bind to fcc sites (see Fig. 8(b)); the graph representation of each one of these steps appears in Sec. 3 of Ref. 29. Thus, OH decomposition and disproportionation reactions between COOH and O or OH (reactions 7, 11, and 12 in Table V) involve more than two sites in a linear

TABLE IV. Gas (a) and surface (b) species participating in the water-gas shift chemistry.

Gas species	
CO	
CO_2	
O_2	
H_2O	
H_2	
Surface species	**Binding site(s)**
CO*	fcc
O*	fcc
H_2O*	top
OH*	top
H*	fcc
HCO**	top-bridge
HCOO***	top-bridge-top
COOH*	top

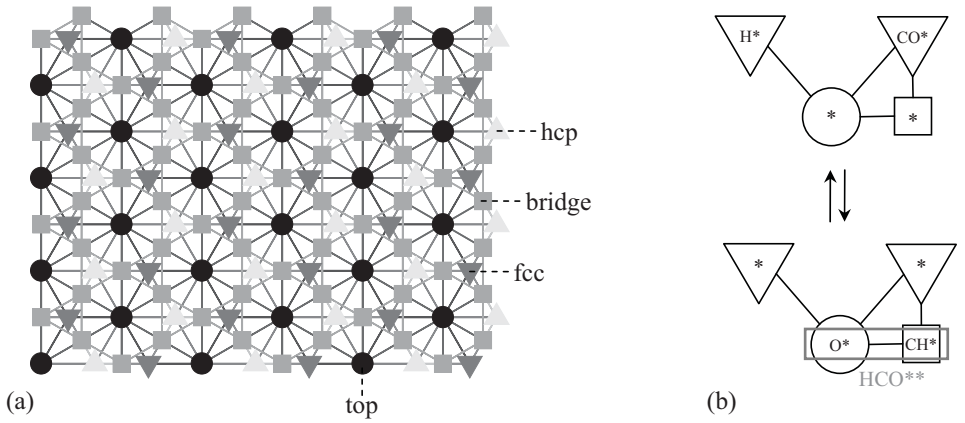


FIG. 8. (a) The Pt(111) lattice for the water-gas shift simulations. (b) An example of an elementary reaction pattern: formyl formation.

connectivity pattern; further, OH-OH disproportionation, HCO formation, and all reactions involving HCOO involve more than 3 sites in complex patterns (reactions 8, 13–17 in Table V).

The rate constants for each step are computed from TST expressions (see Sec. 1 of Ref. 29), for which the input is found in the literature or calculated through quantum chemistry software. Specifically, we use DFT as implemented in the SIESTA code,[42] with which we model bound configurations on four layers of metals (16 atoms), with the bottom two layers fixed, and the top two layers and the adsorbates relaxed. The vacuum region between slabs is ∼10 Å. A $p(2 \times 2)$ unit cell and surface Monkhorst Pack meshes of $5 \times 5 \times 1$ k-point sampling in the surface Brillouin zone are used. The DZP basis set and the PBE-GGA functional were chosen for all calculations. Thus, we compute the total energies and vibrational frequencies of surface, transition state, and gas-phase species (summarized in Sec. 4 of Ref. 29). The vibrational frequencies are subsequently used in the calculation of vibrational partition functions and zero-point energy corrections. For the rotational partition functions of gas molecules, the moments of inertia are read from Table A6.2 in Ref. 43. For the translational partition functions, the mass of each gas molecule is required, which is trivial to calculate. From this information, the rate of each elementary event can be calculated (see Eq. (9)).

Simulations for this system are performed for $P_{H_2O} = 0.10$ bar and CO/H_2O gas molar fraction of 1/2. In preliminary simulations the adsorption and desorption processes of CO and H_2O were found to be in partial equilibrium and 6 orders of magnitude faster than all other processes. Thus, to accelerate the simulations both forward (adsorption) and reverse (desorption) rates were divided by 100; all results shown here were obtained with these latter rates. A lattice with surface area around 6120 Å^2 containing 5520 sites (of any site type) is used. The results of these simulations are portrayed in Fig. 9: panel (a) shows the lattice coverages of CO* and H_2O* at 650 K, normalized with respect to the number of fcc and top sites for the two species, respectively. The system is observed to quickly reach a stationary state

TABLE V. Elementary events and rate parameters of the water-gas shift chemistry for T = 650 K.

	Elementary event	A_{fwd} (s^{-1})	A_{fwd}/A_{bwd}	$E_{a,fwd}$ (eV)	ΔE_{rxn} (eV)
1	$CO_{(g)} + * \leftrightarrow CO*$	3.41×10^{5} bar^{-1}	3.43×10^{-9}	0.00	$-1.82 + 0.11N_{CO}$[a]
2	$O_{2(g)} + 2* \leftrightarrow 2\,O*$	1.18×10^{8} bar^{-1}	4.84×10^{-8}	0.00	−2.32
3	$H_{2(g)} + 2* \leftrightarrow 2\,H*$	2.87×10^{7} bar^{-1}	4.65×10^{-6}	0.00	−1.00
4	$CO_{2(g)} + 2* \leftrightarrow CO* + O*$	6.65×10^{4} bar^{-1}	5.66×10^{-8}	1.23	$-0.18 + 0.11N_{CO}$
5	$H_2O_{(g)} + * \leftrightarrow H_2O*$	7.20×10^{5} bar^{-1}	1.69×10^{-7}	0.00	−0.29
6	$H_2O* + * \leftrightarrow OH* + H*$	4.48×10^{12}	6.39	0.65	0.40
7	$OH* + * \leftrightarrow O* + H*$	2.36×10^{13}	19.8	0.79	−0.31
8	$2\,OH* \leftrightarrow O* + H_2O*$	3.09×10^{11}	3.09	0.00	−0.71
9	$CO* + OH* \leftrightarrow COOH* + *$	4.58×10^{11}	3.90×10^{-2}	$0.48 - 0.055N_{CO}$	$-0.17 - 0.11N_{CO}$
10	$COOH* \leftrightarrow H* + CO_{2(g)}$	5.28×10^{14}	8.96×10^{9} bar	0.67	0.04
11	$COOH* + O* \leftrightarrow OH* + CO_{2(g)}$	6.93×10^{11}	4.53×10^{8} bar	0.45	0.35
12	$COOH* + OH* \leftrightarrow H_2O* + CO_{2(g)}$	1.04×10^{13}	1.40×10^{9} bar	0.10	−0.36
13	$H* + CO* \leftrightarrow HCO**$	1.98×10^{11}	2.35×10^{-2}	$1.35 - 0.055 \cdot N_{CO}$	$1.08 - 0.11N_{CO}$
14	$HCO** + O* \leftrightarrow HCOO***$	6.15×10^{11}	4.89×10^{-2}	1.06	−0.54
15	$HCOO*** \leftrightarrow CO_{2(g)} + H* + 2*$	5.97×10^{13}	1.54×10^{10} bar	0.91	−0.36
16	$HCOO*** + O* \leftrightarrow CO_{2(g)} + OH* + 3*$	1.19×10^{12}	7.76×10^{8} bar	1.71	−0.05
17	$HCOO*** + OH* \leftrightarrow CO_{2(g)} + H_2O* + 3*$	6.30×10^{12}	2.40×10^{9} bar	0.92	−0.76

[a]This term denotes CO-CO pairwise repulsive interactions of 0.11 eV per CO molecule adsorbed in a neighboring fcc site. The value was obtained from Ref. 45: 15 kcal/mol per monolayer that corresponds to 0.11 eV per each one of the 6 possible neighboring CO molecules.

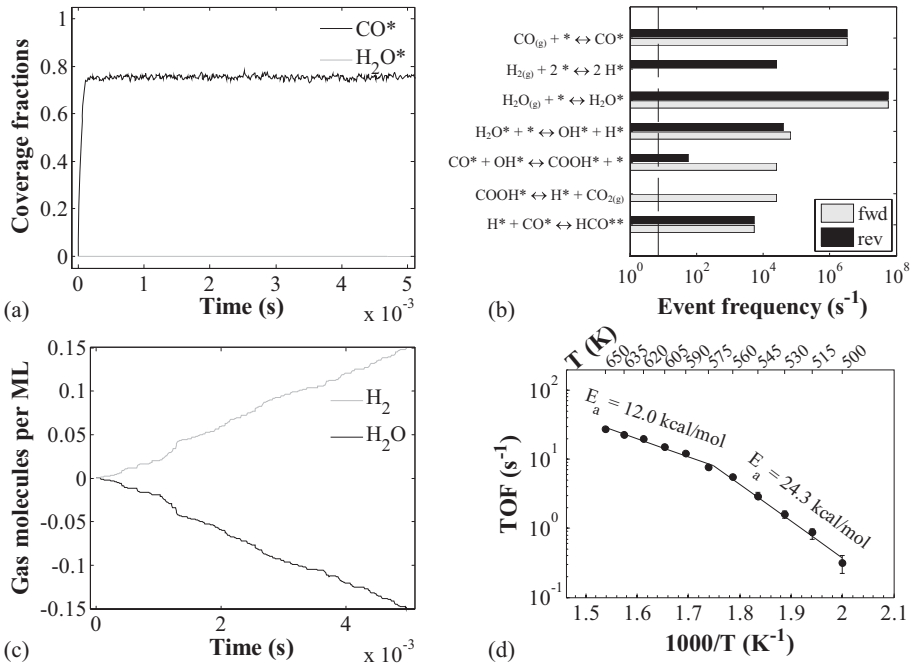


FIG. 9. (a) Coverage fractions for CO* and H_2O* as a function of time. The surface is almost poisoned by CO. (b) Number of gas species molecules produced or consumed per time per monolayer. (c) Arrhenius plot for the Pt (111) water-gas shift model. The apparent activation energy is ~32 kcal/mol. (d) Frequencies for the elementary events observed during the course of the simulation. The vertical line corresponds to the frequency of a single event $= (t_{fin} - t_{ini})^{-1}$. For all simulations T = 650 K (except in panel d), $P_{H_2O} = 0.10$ bar, $P_{CO} = 0.05$ bar.

(equilibration time less than 5 ms for all conditions investigated) in which carbon monoxide almost poisons the surface, whereas water coverages are low due to this species' weak binding on Pt. Note that diffusion was neglected for the simulations shown here. The effect of this process was assessed by allowing CO and H to hop between neighboring sites and the results obtained were practically the same with those obtained in the absence of diffusion (see Sec. 5 of Ref. 29).

Furthermore, Fig. 9(b) shows the frequencies of the elementary events that took place in the course of the simulation (reaction channels not mentioned in the bar graph did not fire within the accessible time scale). It is observed that H_2O and CO adsorption-desorption are partially equilibrated processes, validating our treating them as such to accelerate the simulations. Further, HCO formation-decomposition is in partial equilibrium too, indicating that HCO does not contribute to the overall water-gas shift chemistry. On the other hand, H_2O dissociation and COOH formation and subsequent decomposition towards adsorbed H and CO_2 gas are non-equilibrated processes, indicating that the overall chemistry on Pt(111) proceeds through the carboxyl pathway. H_2O dissociation appears to be the rate-determining step (RDS), having a partial-equilibrium ratio (PE ratio) of 0.79 at 530 K and 0.62 at 650 K (the PE ratio is defined as the ratio between the forward rate over the sum of forward and backward rates). The subsequent step in the chemistry, namely the carboxyl formation, remains irreversible (PE ratio of 1.0 at both 530 and 650 K). These observations are in agreement with the results of Grabow *et al.*[33] where it was reported that the direct COOH decomposition accounts for the majority of CO_2 produced and that H_2O dissociation is the RDS.

Furthermore, Fig. 9(c) shows the number of H_2 and H_2O gas-phase molecules produced and consumed, respectively, as a function of time. From the slope of the number of H_2 molecules and given the number of sites on the surface, one can calculate the overall reaction rate per monolayer, for various temperatures. In our calculations of the rate, we discard the first 0.001 s in order to sample the stationary part of the trajectory. The resulting Arrhenius plot is shown in Fig. 9(d). The error bars in this plot quantify the uncertainty in the calculated rate due to the finite number of H_2 molecules produced over the course of a simulation. The smaller the total number of H_2 molecules, the larger the uncertainty. For a detailed description of the error estimation procedure please refer to Sec. 6 of Ref. 29. The apparent activation energy varies from 24.3 kcal/mol for the low temperature range, to 12.0 kcal/mol for higher temperatures. These values compare well with the experimentally observed barriers of 13.9–27 kcal/mol (Ref. 44) and the value of 16.2 kcal/mol (67.8 kJ/mol) calculated by Grabow *et al.*[33] through mikrokinetic mean-field modeling. The observed drop in the activation energy can be attributed to the CO poisoning effects. Thus, for high temperatures the apparent activation energy is close to the barrier of the RDS, namely 15 kcal/mol for the H_2O dissociation, whereas for low temperatures it appears that CO poisoning is prominent; consequently, the rate is determined by the equilibrium between gas and surface CO.

Finally, it is interesting to note that throughout the course of the simulation, no HCOO molecules were observed. This can be attributed to the lack of adsorbed O on the surface, which eliminates the possibility of HCOO being produced through the association of HCO and O (reaction 14 in

Table V), and the high barriers exhibited by the Eley Riedel reactions (15)–(17) in Table V.

IV. CONCLUSIONS

We have presented a novel KMC framework that employs graph theoretical ideas to explicitly account for the neighboring patterns in complex surface kinetics. In this approach, lattice structure and elementary events are represented as graphs, and the identification of lattice processes is done by solving subgraph isomorphism problems during the course of a simulation. Optimized algorithms and data structures are used in order to minimize the computational overhead resulting in computational times comparable to simpler KMC techniques.

We further demonstrated the capabilities of this framework by simulating the water-gas shift reaction on Pt (111). This chemistry proceeds in several different site types, involves species that bind to more than one site and contains elementary reactions with complex geometrical arrangements of reactants and products. KMC simulations using our framework predicted surface coverages and activities as a function of temperature, thereby providing an estimate for the apparent activation energy for the overall chemistry. Moreover using statistical analysis, we identified partially equilibrated reversible events and elucidated the main pathway through which the chemistry proceeds.

The flexibility and specificity in the definition of elementary events within the presented framework makes the latter applicable to a vast array of problems involving complex chemistries and geometries, such as those encountered in vicinal surfaces or the facets of nanoparticles.

ACKNOWLEDGMENTS

The research was partially supported by Grant No. DE-FG02–05ER25702 from the (U.S.) Department of Energy (DOE). M.S. and D.G.V. would like to thank Dr. Ying Chen for performing DFT calculations on the WGS chemistry.

[1] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**(1), 10 (1975).
[2] L. V. Lutsevich, O. A. Tkachenko, and V. P. Zhdanov, Langmuir **8**(7), 1757 (1992).
[3] M. Silverberg, A. Ben-Shaul, and F. Rebentrost, J. Chem. Phys. **83**(12), 6501 (1985).
[4] M. Silverberg and A. Ben-Shaul, J. Chem. Phys. **87**(5), 3178 (1987).
[5] P. Araya, W. Porod, R. Sant, and E. E. Wolf, Surf. Sci. **208**(1–2), L80 (1989).
[6] J. Cortes, E. Valencia, and H. Puschmann, Phys. Chem. Chem. Phys. **1**(7), 1577 (1999).
[7] M. O. Coppens, A. T. Bell, and A. K. Chakraborty, Chem. Eng. Sci. **53**(11), 2053 (1998).
[8] F. J. Keil, R. Krishna, and M. O. Coppens, Rev. Chem. Eng. **16**(2), 71 (2000).
[9] E. W. Hansen and M. Neurock, Chem. Eng. Sci. **54**(15–16), 3411 (1999).
[10] E. W. Hansen and M. Neurock, Surf. Sci. **464**(2–3), 91 (2000).
[11] E. W. Hansen and M. Neurock, J. Catal. **196**(2), 241 (2000).
[12] K. Reuter, D. Frenkel, and M. Scheffler, Phys. Rev. Lett. **93**(11), 116105 (2004).
[13] D. Mei, P. A. Sheth, M. Neurock, and C. M. Smith, J. Catal. **242**(1), 1 (2006).
[14] K. Reuter and M. Scheffler, Phys. Rev. B **73**(4), 045433 (2006).
[15] K. Reuter, "First-principles kinetic Monte Carlo simulations for heterogeneous catalysis: concepts, status and frontiers," in *Modeling Heterogeneous Catalytic Reactions: From the Molecular Process to the Technical System*, edited by O. Deutschmann (Wiley-VCH, Weinberg, 2010).
[16] D. H. Mei, J. C. Du, and M. Neurock, Ind. Eng. Chem. Res. **49**(21), 10364 (2010).
[17] D. H. Mei, M. Neurock, and C. M. Smith, J. Catal. **268**(2), 181 (2009).
[18] H. Meskine, S. Matera, M. Scheffler, K. Reuter, and H. Metiu, Surf. Sci. **603**(10–12), 1724 (2009).
[19] J. S. Reese, S. Raimondeau, and D. G. Vlachos, J. Comput. Phys. **173**(1), 302 (2001).
[20] L. J. Broadbelt and R. Q. Snurr, Appl. Catal., A **200**(1–2), 23 (2000).
[21] J. J. Lukkien, J. P. L. Segers, P. A. J. Hilbers, R. J. Gelten, and A. P. J. Jansen, Phys. Rev. E **58**(2), 2598 (1998).
[22] A. Chatterjee and D. G. Vlachos, J. Comput.-Aided Mater. Des. **14**(2), 253 (2007).
[23] CARLOS Project: a general purpose program for the simulation of chemical reactions taking place at crystal surfaces, available from: http://carlos.win.tue.nl/.
[24] V. P. Zhdanov and B. Kasemo, Surf. Sci. **405**(1), 27 (1998).
[25] R. M. Ziff, E. Gulari, and Y. Barshad, Phys. Rev. Lett. **56**(24), 2553 (1986).
[26] P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).
[27] P. Pechukas, "Statistical approximations in collision theory" in *Dynamics of Molecular Collisions. Part B*, edited by W. H. Miller (Plenum, New York, 1976).
[28] D. A. McQuarrie, *Statistical Mechanics*, 2nd ed. (Harper & Row, New York, 2000).
[29] See supplementary material athttp://dx.doi.org/10.1063/1.3596751 for: (1) Rate expressions for elementary steps involving gas species; (2) Comparison of the computational times of graph-theoretical KMC and a simpler KMC approach; (3) Water-gas shift elementary reaction patterns; (4) Data used in the rate constant calculations; (5) Effect of Diffusion; (6) Error estimates for the Arrhenius plot.
[30] S. D. Miller and J. R. Kitchin, Mol. Simul. **35**(10–11), 920 (2009).
[31] W. Chen, D. Schmidt, W. F. Schneider, and C. Wolverton, Phys. Rev. B **83**(7), 075415 (2011).
[32] J. M. Sanchez, F. Ducastelle, and D. Gratias, Physica A **128**(1–2), 334 (1984).
[33] L. C. Grabow, A. A. Gokhale, S. T. Evans, J. A. Dumesic, and M. Mavrikakis, J. Phys. Chem. C **112**(12), 4608 (2008).
[34] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 2004).
[35] Graph Theoretical KMC Code, available from: http://www.dion.che.udel.edu.
[36] M. A. Gibson and J. Bruck, J. Phys. Chem. A **104**(9), 1876 (2000).
[37] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (MIT, Cambridge, MA, 2009).
[38] J. R. Ullmann, J. ACM **23**(1), 31 (1976).
[39] D. Eppstein, J. Graph Algorithm Appl. **3**(3), 1 (1999).
[40] D. T. Gillespie, J. Phys. Chem. **81**(25), 2340 (1977).
[41] P. J. Feibelman, B. Hammer, J. K. Norskov, F. Wagner, M. Scheffler, R. Stumpf, R. Watwe, and J. Dumesic, J. Phys. Chem. B **105**(18), 4018 (2001).
[42] J. M. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejon, and D. Sanchez-Portal, J. Phys. Condens. Matter **14**(11), 2745 (2002).
[43] J. B. Ott and J. Boerio-Goates, *Chemical Thermodynamics: Advanced Applications*. (Academic, London, 2000).
[44] P. Panagiotopoulou and D. I. Kondarides, Catal. Today **112**(1–4), 49 (2006).
[45] A. B. Mhadeshwar and D. G. Vlachos, J. Phys. Chem. B **108**(39), 15246 (2004).