The XPK Package: A Comparison between the Extended Phenomenological Kinetic (XPK) Method and the Conventional Kinetic Monte Carlo (KMC) Method†

Tong-hao Shen, Xin Xu∗

Collaborative Innovation Center of Chemistry for Energy Materials, Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, MOE Laboratory for Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200438, China

(Dated: Received on January 21, 2019; Accepted on January 31, 2019)

Recently, we proposed the extended phenomenological kinetics (XPK) method, which overcomes the notorious timescale separation difficulty between fast diffusion and slow chemical reactions in conventional kinetic Monte Carlo (KMC) simulations. In the present work, we make a comprehensive comparison, based on the newly developed XPK package, between the XPK method and the conventional KMC method using a model hydrogenation reaction system. Two potential energy surfaces with different lateral interactions have been designed to illustrate the advantages of the XPK method in computational costs, parallel efficiency and the convergence behaviors to steady states. The XPK method is shown to be efficient and accurate, holding the great promise for theoretical modelling in heterogeneous catalysis, in particular, when the role of the lateral interactions among adsorbates is crucial.

Key words: Kinetic Monte Carlo, Extended phenomenological kinetics, Surface reaction, Heterogeneous catalysis

I. INTRODUCTION

Modeling the kinetics of surface reactions is essential for understanding the catalytic processes, which, in turn, is critical for the rational design of catalysts. Much effort has been devoted into developing suitable kinetic models, leading to methods of varying accuracy and efficiency such as the Langmuir-Hinshelwood closed form models [1–3], mean-field microkinetic models [4, 5], and kinetic Monte Carlo (KMC) simulation methods [6–10], etc. Among these models the mean-field microkinetic modeling is prevailing due to its simplicity and efficiency. For most heterogeneously catalyzed reactions on the metal surfaces, it usually provides semi-quantitative descriptions of the kinetics at high temperature, and there are a number of cases in which it shows good descriptions of the trends in activity from one catalyst to another [11–16].

Although the mean-field microkinetic method, being a phenomenological kinetic (PK) approach, has won a great popularity in screening some important catalytic processes, it lacks the descriptions of fluctuations, spatial inhomogeneity and spatial correlations of adsorbates. These drawbacks have strongly limited its applications, especially in the operando spatiotemporal understanding at the mesoscopic level. To fill this gap, the KMC model has attracted more and more interest in the recent years due to its accuracy. Beyond the PK model under the mean-field approximation, the KMC model can provide a detailed description of all the elementary processes, including adsorption, desorption, diffusion, and chemical reactions, taking into account the diverse local surface environments in which they occur. In fact, the KMC model has been successfully applied in the investigation of several important systems, such as CO oxidation [17–19], NO oxidation [20], water-gas shift [21, 22], methanol decomposition [23], etc. However, due to the disparity in timescales of various processes, the high computational cost prevents its widespread applications in many other academically interested and industrially important systems. A number of KMC acceleration approaches have been developed, which include tau-leaping [24–26], net-event algorithm [27, 28], and rescaling the reaction rate constants [29–32], etc. Nonetheless, a routine application of the KMC method in realistic catalytic systems to, for example, including the lateral interactions among the adsorbates, remains a great challenge.

Thus, on one hand, PK is efficient but could be physically deficient; while, on the other hand, KMC is physically sound but could be practically infeasible. The extended phenomenological kinetic (XPK) approach [33] developed very recently is one of the choices that try
to close this gap. By assuming a thermal equilibrated surface for a given composition of adsorbed species, the full state Markovian master equations (MME) in the spatial space can be reduced to those in the chemical space [33, 34], i.e., the chemical master equation (CME) [35]. Along this line, the XPK method decomposes the KMC simulations into two steps, the diffusion-only lattice KMC simulations at some given compositions, and later the implicit KMC simulations in chemical space. In the first step, only adsorbate diffusions occur on the lattice, while the rates of chemical reactions are averaged to get a set of apparent rate constants (\(k_{\text{app}}\)). Then these apparent rate constants are used in the simulations in chemical space to evolve the coverages of adsorbed species. As the CME is much simpler than the full state MME, the simulations in chemical space are far more efficient than that in spatial space by definition. However, the efficiency of the XPK method, and its direct comparison with that of conventional KMC simulations, has not been well documented yet.

Recently, we have developed an efficient XPK package. In this work, we make a comprehensive comparison for the performance between the XPK method and the conventional KMC method, in a model system close to the realistic catalytic hydrogenation processes.

II. METHODS AND MODELS

In a heterogeneous catalytic system, where the active sites are mapped to a lattice, the time evolution of the system can be described by the full state MME,

\[
\frac{dP_x(t)}{dt} = -\sum_{y \neq x} K(x \rightarrow y) P_x(t) + \sum_{y \neq x} K(y \rightarrow x) P_y(t)
\]

(1)

Here \(P_x(t)\) is the probability to find the system stay in a surface lattice configuration \(x\) at time \(t\), and \(K(x \rightarrow y)\) is the transition rate from configuration \(x\) to \(y\). As the number of accessible configurations on the lattice is extremely large, Eq.(1) cannot be solved directly in general, and the KMC simulation provides a way to achieve the exact numerical solution of Eq.(1). There are a series of algorithms to carry out a KMC simulation, which can be divided into two classes [36]: the rejection algorithms and the rejection-free algorithms. In many applications, the rejection-free algorithms are computationally more efficient than the others, and thus become more popular nowadays.

There are two stages in each step of a rejection-free KMC simulation:

(i) The transition rates are calculated from current configuration to all accessible configurations, \(\sum_{y \neq x} K(x \rightarrow y)\) gives the total rate of the system to leave configuration \(x\). The time interval that the system stays in the current configuration is then obtained according to

\[
\Delta t = -\frac{\ln(R_N)}{\sum_{y \neq x} K(x \rightarrow y)}
\]

(2)

where \(R_N\) is a random number uniformly distributed in the interval (0,1].

(ii) The system jumps to an accessible configuration randomly, and the transition possibility is proportional to transition rate,

Repeat the steps above, one gets the time evolution of the system. In general, the simulation finishes if the steady state is achieved, which means that the coverages of all the surface species do not change any more with time. The statistic properties can be obtained by analyzing the trajectory, and the probability of each snapshot is proportional to the physical time that the system stays in the respective configuration.

\[
P_x \propto \sum_{y = x} \Delta t (y)
\]

(3)

The summation on the right hand side of Eq.(3) runs over the whole trajectory.

A Fortran pseudo-code is given as follows,

```
Call Initialization()
do i_step = 1, n_step
  ! Calculate total rate
  Call get_total_rate(Rtot)
  ! Calculate and accumulate time interval
  Call random_number(RN)
  dt = -log(RN)/Rtot
  time = time + dt
  ! Pick up a reaction
  Call random_number(Rpick)
  Rpick = Rpick * Rtot
  Call pick_up_reaction(Rpick, ipick)
  ! Update surface
  Call update_surface(ipick)
end do
Call statistics()
```

As the rate of a reaction is only influenced by the lattice environment nearby, its computational cost can be made invariant with respect to different system size by taking into account this nearsightedness. On the other hand, because the number of possible reactions is proportional to the lattice sites, the computational cost in the procedure of picking up a reaction increases with system size. Fortunately, with the help of binary
tree searching method [37], the dependence of computational cost on system size can be reduced to,

\[ \text{FLOP} \propto \log_2 N_{\text{site}} \quad (4) \]

where FLOP denotes the number of floating operations, while \( N_{\text{site}} \) stands for the number of sites on the lattice. For a small or medium sized lattice (e.g. \( N_{\text{site}} \leq 10^6 \)), the FLOPs of searching a binary tree are often smaller than that was needed in updating lattice and calculating reaction rates. Therefore the cost is approximately size independent of each KMC step.

Despite that the FLOP number for each KMC step can be limited by well organizing the code, traditional KMC simulations are still computationally unaffordable for most of the realistic systems. Firstly, the disparity in timescales of diffusion, adsorption, and reactions results in a huge number of steps to achieve the steady state on the lattice, as well as an extremely low efficiency in the time evolution of coverages of surface species. Secondly, according to Eq.(2), the scale of physical time for each snapshot is proportional to the reciprocal of total rates. Therefore, the computational cost of evolving the system for a given time interval increases with system size. Thirdly, as the KMC simulation describes the real time evolution of a Markovian process, it can hardly benefit from the parallel calculation on a modern supercomputer. Although parallel KMC calculations have been demonstrated in several types of applications [38–40], the design of parallel algorithms for general cases is still a non-trivial task.

It has been widely accepted that the very fast adsorbate diffusion is the primary cause of the disparity of timescale in most catalytic systems [34, 41]. On metal surfaces, diffusion rates can be several orders of magnitude higher than chemical reaction rates. Therefore, diffusions dominate a lattice KMC simulation, whereas chemical reactions become ‘rare events’. Notice that coverages of adsorbates change only when a chemical reaction occurs, such that the full state MME can be reduced to the CME rigorously if diffusions are fast enough to reach a quasi-equilibrium state [33, 42]. Based on the CME, the XPK approach further uses a PK-type equation to describe the reaction rates in chemical space.

\[ \frac{\langle R(A) \rangle_{\text{diff}}}{N_{\text{neighbor}} \theta_{r1} \theta_{r2}} \quad (6) \]

in which \( \langle R(A) \rangle_{\text{diff}} \) is the statistical average of the rate for reaction A during the lattice simulations. However, it is too expensive to apply a lattice KMC simulation to each state in chemical space. Therefore, the XPK approach further takes the advantage of the observation that \( k_{\text{app}} \) values for all the reactions are mainly determined by the coverages of a few major adsorbates and the coverage dependences of rate constants can be approximately described by an interpolation method [33]. With these simplifications, lattice KMC simulations are only necessary for a small set of chemical states. Another advantage of the XPK approach is that at the diffusion limit, the surface is considered to be thermal equilibrated before each chemical reaction occurs. Thus the probability distribution of the lattice configurations at a given coverage is time independent. The sampling for these configurations, therefore, can be carried out by averaging on a series of separate trajectories, which facilitates parallel calculations. To illustrate these points, a Fortran pseudo-code for the XPK simulations is given below.

\begin{verbatim}
Call Initialization()
! Diffusion only KMC simulation for Kapp
do i_point = 1, n_interpolation_point
    ! Get Kapp for i-th interpolation point
    ! n_sample samples on n_tasks cores
    do i_sample = 1, n_sample
        if (mod(i_sample, n_tasks) == myid) then
            Call get_Kapp(i_point, Kapp)
        end if
    end do
    Call synchronize(i_point, Kapp)
end do

! Evolution in chemical space
do i_step = 1, n_step
    ! Calculate and accumulate time interval
    Call random_number(RN)
    dt = log(RN)/Rrot
    time = time + dt

    ! Pick up a reaction
    Call random_number(Rpick)
    Rpick = Rpick * Rrot
    Call pick_up_reaction(Rpick, ipick)

    ! Update coverage
    Call update_coverage(ipick)
end do
\end{verbatim}

A model system was designed to test the efficiency of the XPK approach, and to compare XPK with the conventional KMC approach. This model is closely related to the multi-step hydrogenation processes in heterogeneous catalysis. The model hydrogenation reaction is
TABLE I The thermodynamic and kinetic parameters for the designed testing models.

<table>
<thead>
<tr>
<th>Elementary process</th>
<th>PES1</th>
<th>PES2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E_a$/eV</td>
<td>$k$/s$^{-1}$</td>
</tr>
<tr>
<td>$A(g) + ^\ast \rightarrow A^\ast$</td>
<td>-</td>
<td>$1.993 \times 10^5$</td>
</tr>
<tr>
<td>$A^\ast + A(g) + ^\ast \rightarrow A(g)$</td>
<td>1.119</td>
<td>$4.981 \times 10^3$</td>
</tr>
<tr>
<td>$2H^\ast \rightarrow 2H^\ast$</td>
<td>-</td>
<td>$4.501 \times 10^4$</td>
</tr>
<tr>
<td>$2H^\ast \rightarrow 2H^\ast$</td>
<td>0.957</td>
<td>$1.141 \times 10^5$</td>
</tr>
<tr>
<td>$A^\ast + H^\ast \rightarrow AH^\ast + ^\ast$</td>
<td>1.240</td>
<td>$4.802 \times 10^2$</td>
</tr>
<tr>
<td>$AH^\ast + ^\ast \rightarrow A^\ast + H^\ast$</td>
<td>0.940</td>
<td>$1.590 \times 10^5$</td>
</tr>
<tr>
<td>$2H(g) \rightarrow 2H^\ast$</td>
<td>1.180</td>
<td>$1.533 \times 10^3$</td>
</tr>
<tr>
<td>$A^\ast + H^\ast \rightarrow AH^\ast + ^\ast$</td>
<td>0.700</td>
<td>$1.649 \times 10^7$</td>
</tr>
<tr>
<td>$H^\ast + H^\ast \rightarrow H^\ast + H^\ast$</td>
<td>0.700</td>
<td>$1.649 \times 10^7$</td>
</tr>
<tr>
<td>$AH^\ast + ^\ast \rightarrow A^\ast + AH^\ast$</td>
<td>0.700</td>
<td>$1.649 \times 10^7$</td>
</tr>
</tbody>
</table>

where "\*" denotes a surface site, and $\leftrightarrow$ means the reaction is reversible. In addition, the diffusions of all the three adsorbates are taken into account, denoted by ‘PES1’; (ii) simulations with lateral interactions on PES2. The interactions between $A^\ast$ and $A^\ast$, $A^\ast$ and $AH^\ast$, $AH^\ast$ and $AH^\ast$ are 0.06 eV, 0.03 eV and 0.02 eV, respectively. This test is labeled by ‘PES2(0.02)’, corresponding to the interaction between $AH^\ast$ and $AH^\ast$. (c) Similar to case b, but the interaction between $AH^\ast$ and $AH^\ast$ is slightly modified to 0.015 eV, which is labeled by ‘PES2(0.015)’.

The simulations have been performed at 600.0 K, on a periodic hexagonal lattice. A 40$\times$40 grid mesh has been adopted, which is often a good approximation of an infinite slab surface, with an acceptable computational cost for a conventional KMC simulation. All the tests have been carried out on a single computer node, which has two Intel(R) Xeon(R) E5-2676 v3 CPUs with 24 physical cores, and 2 GB memory for each core.
TABLE II Details of KMC and XPK simulations on PES1, PES2(0.02) and PES2(0.015)*.

<table>
<thead>
<tr>
<th></th>
<th>PES1</th>
<th>PES2(0.02)</th>
<th>PES2(0.015)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KMC</td>
<td>XPKb</td>
<td>XPKc</td>
</tr>
<tr>
<td>Total time/s</td>
<td>1.7×10⁵</td>
<td>0.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Total steps</td>
<td>1.1×10⁷</td>
<td>0.0</td>
<td>2.0×10⁵</td>
</tr>
<tr>
<td>Time per 10⁷ steps/s</td>
<td>160.0</td>
<td>0.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

* The KMC simulations ended after 1000 products were yielded, while the XPK simulations ended after the steady state was achieved. No parallel algorithms were applied to XPK.

b For the diffusion-only lattice KMC part.

c For the evolution in chemical space.

III. RESULTS AND DISCUSSION

As discussed in section II, the lattice KMC simulation is a vital part of the XPK method. Therefore, an efficient KMC algorithm is highly important not only for the KMC simulation itself, but also for the XPK simulation. One of the essential requirements is that the computational cost of the algorithm shall not increase for each KMC step as the system size is increased. Thus the first test is the system size dependence of the KMC simulations. The tests were performed on PES1 and PES2(0.02) models, the results are shown in FIG. 2. As expected, with the help of the binary tree searching method, the increase of system size (100 times larger) did not cause significant change of the time cost in both cases. Meanwhile, the time cost of PES2(0.02) model was about 4 times higher than that of PES1, which implies that the KMC module became the bottleneck for the simulations if lateral interactions had to be taken into account.

However, an efficient algorithm for each KMC step cannot cure the suffering of timescale separation, which leads to a dramatic increase of the total number of KMC steps. Although the diffusion rates are usually rescaled to some relatively low values (e.g., 100 times faster than the fastest chemical reaction), the evolution of surface coverage can still be quite slow for a conventional KMC simulation. As the second test, we compared the time costs of KMC and XPK simulations. In this comparison, the system evolved from a clean surface to the steady state. It should be emphasized that, the word ‘steady state’ actually has different meanings for XPK and KMC simulations. In the KMC simulations, the very slow evolution speed in chemical space has prevented setting a critical criterion to check whether the ‘steady state’ is reached or not. In general, it can only be approximated by the desired number of products yielded. Here the time costs of KMC simulations were accumulated until 1000 AH₂ were produced. On the contrast, the XPK method directly solves the CME, which leads to a much faster evolution speed. Therefore, a more precise analysis can be applied to the steady state in XPK simulations. Here we adopted the change of averaged coverage as a criterion. In each simulation, the coverages of the major adsorbates were averaged in every 10⁷ steps, and the steady state was believed to be achieved if the coverage changes were less than 0.0005. To make sure the accuracy of the sampling and interpolation procedure in the XPK method, the diffusion-only KMC simulations were carried out for 1.6×10⁷ steps at each interpolation point of the chosen coverage, and the interval between neighboring points was set to 0.04 for coverage change. After obtaining results for all the interpolation points, a linear interpolation method was applied to approximate the reaction rates in chemical space.

As shown in FIG. 3, in all three cases, the XPK simulations consumed much shorter time as compared to the KMC simulations. For PES1, there is no lateral interaction between adsorbates, such that the rate for each reaction is a constant. Therefore, the $k_{app}$ was obtained immediately by using Eq.(6), as there was no need to do the diffusion-only KMC simulations. In this case, the time cost of the XPK simulations was 5 orders of magnitude smaller than that of the KMC simulations. After including the lateral interactions, the computational cost of the XPK simulations significantly increased on the PES2. It could be found that this increase mainly came from the diffusion-only KMC simulations, for which the time cost was 3 orders of

DOI:10.1063/1674-0068/cjpc1901013 ©2019 Chinese Physical Society
magnitude higher than the evolution time in chemical space. The main reason for the big difference in time costs was that the explicit lattice KMC simulations were much more complicated than implicit KMC simulations. Table II gives details of KMC and XPK simulations for their efficiencies on PES1, PES2(0.02) and PES2(0.015). The KMC simulations ended after 1000 products were yielded, while the XPK simulations ended after the steady state was achieved. As summarized in Table II, the time cost for each step in the diffusion-only KMC simulation was quite similar to that of the conventional KMC simulation, while it was about 300 times higher than that for the evolution in chemical space.

The efficiency of XPK can be further improved by utilizing the parallel computing techniques. As mentioned in section II, the sampling in the diffusion-only KMC simulation, which is the most time consuming part in the XPK method, can be carried out by averaging a number of trajectories. During this procedure, the data communication is only needed after all the trajectories are ended. Therefore, a high parallel efficiency can be explored. The strong scaling of the XPK simulations on PES2(0.02), as well as the speedup is plotted in FIG. 4. The XPK simulations showed a very good strong scaling up to 16 cores for parallel calculations. Meanwhile, the elapsed time reduced from $1.7 \times 10^4$ s of a serial run shown in FIG. 3 and Table II to $1.2 \times 10^3$ s of a 16-core parallel computation. The parallel efficiency remained about 90% if 16 cores were employed. It is worth to note that, a massively parallel calculation with the XPK approach is available by definition, though it is not necessary for this simple test case.

Besides the high computational efficiency, the fast evolution in chemical space makes the XPK method more valuable than the conventional KMC in many aspects. With the help of the XPK simulations, the analysis of many properties becomes much simpler and thus applicable. As an example, the time evolution of the coverages of the major adsorbates and the number of products yielded are summarized in FIG. 5. As a comparison, the results obtained in conventional KMC simulations are also shown. The solid lines denote the coverages of the major adsorbates, i.e., $A^*$ for PES1, $AH^*$ for PES2(0.02) or PES2(0.015), while the dash lines refer to the number of $AH_2$ produced. In addition, the horizontal line of red dash dots labeled the coverages of the major adsorbates at the steady state obtained in XPK simulations. For the time evolution shown in FIG. 5(a), both the XPK and KMC simulations achieved the steady state quickly, but the products yielded very slowly. Therefore, producing $1000 \, AH_2$ was far enough to reach the steady state in PES1. The cases in PES2 were quite different. For both PES2(0.02) and PES2(0.015) models, the coverage convergences were very slowly, as shown in FIG. 5 (b) and (c), respectively. These different behaviors mainly resulted from the different shapes of PES1 and PES2, yet also were strongly influenced by the lateral interactions for the latter. As compared to PES2(0.02), the interactions between two neighboring $AH^*$ adsorbates in PES2(0.015) were shifted by only 0.005 eV. However, this small modification led to a much slower evolution, and the KMC simulation failed to achieve the steady state even when $1000 \, AH_2$ were produced (see FIG. 5(c)). To ensure that the XPK and KMC simulations evolved to the same steady state on PES2(0.015), we performed an additional KMC simulation, which ended after 2000 $AH_2$ were produced (see FIG. 5(d)). In this long time KMC simulation, the coverage of $AH^*$ achieved the value in steady state after about 1150 $AH_2$ yielded, and then fluctuated around the horizontal line of red dash dots.

Here we emphasize that, although the coverage arrived at the right value finally, the increase of the num-

DOI:10.1063/1674-0068/cjcp1901013
FIG. 5 Steady state tests for KMC and XPK simulations on (a) PES1, (b) PES2(0.02), and (c, d) PES2(0.015) where KMC were carried out to yield 1000/2000 AH$_2$ products, respectively. The solid lines stand for coverages of the major species on the surface. The dashed lines denote the number of AH$_2$ produced. The horizontal lines with red dash dots show the coverages of the major species at the steady states, and the vertical line with black dots in (d) gives the time that 1000 AH$_2$ are obtained.

ber of simulation steps might not be the solution to the convergence problem in the KMC simulations. First, it is because that the computational cost increases linearly with the number of the simulation steps, and thus it becomes unaffordable if a huge number of simulation steps is needed. Secondly, it is difficult to determine whether the steady state is achieved or not. As shown in FIG. 5 (b) and (c), the behaviors of two systems are quite similar, thus without XPK and longtime KMC simulations as the references, it is difficult to realize that PES2(0.02) has achieved the steady state while PES2(0.015) has not when 1000 AH$_2$ have been produced. Thirdly, in a realistic system, the timescale separation is much more serious, and thus the KMC simulation is even not applicable. On the contrary, the XPK method evolves the system in chemical space directly, which ensures a high efficiency in describing surface reactive systems.

IV. CONCLUSION

In the present work, we have demonstrated the high efficiency of the XPK approach based on the newly developed XPK package using a model hydrogenation reaction system. Two potential energy surfaces with different lateral interactions have been designed to illustrate the differences in computational costs, parallel efficiency and the convergence behaviors to steady states, between XPK and the conventional KMC approaches. The results showed that, for a KMC simulation, the efficiency of each step can be made quite satisfactory in our current implementation, but the large number of steps needed for evolving system can lead to an expensive time cost. Even worse, the slow evolution in chemical space can result in troubles in determining the steady state. On the contrary, the XPK method provides a more efficient way to directly evolve the system in chemical space, by separating the surface diffusion and chemical reactions. By applying a linear interpolation method, accurate results can be obtained with a limited number of diffusion-only lattice KMC simulations. The implicit KMC simulations are much cheaper than the diffusion-only KMC simulations, therefore, a long time evolution in chemical space can be obtained, and more detailed analyses can be made available. Furthermore, the time cost of the diffusion-only KMC simulations, which is the most expensive part in an XPK
simulation, can be significantly reduced by using the parallel computing techniques.

Based on these observations, we conclude that, in the fast diffusion limit, the XPK approach offers a highly efficient alternative to the conventional KMC method, without losing accuracy. By overcoming the notorious timescale separation difficulty between diffusion and chemical reactions, XPK drastically reduces the time cost of a simulation. Having these advantages, the XPK method significantly extends the scope of applications of detailed surface simulations, and is expected to be a powerful tool to investigate the surface catalytic reaction systems, in particular, when lateral interactions among adsorbates play a key role.

V. ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (No.21688102) and the National Key Research and Development Program of China (No.2018YFA0208600).