Markov state modeling of sliding friction

F. Pellegrini,[1,2] François P. Landes,[3] A. Laio,[1] S. Prestipino,[4,5] and E. Tosatti[1,2,3,*]

[1]*SISSA, Via Bonomea 265, I-34136 Trieste, Italy*

[2]*CNR-IOM Democritos National Simulation Center, Via Bonomea 265, I-34136 Trieste, Italy*

[3]*International Centre for Theoretical Physics (ICTP), Strada Costiera 11, I-34151 Trieste, Italy*

[4]*Università degli Studi di Messina, Dipartimento di Scienze Matematiche ed Informatiche, Scienze Fisiche e Scienze della Terra, Contrada Papardo, I-98166 Messina, Italy*

[5]*CNR-IPCF, Viale F. Stagno d'Alcontres 37, I-98158 Messina, Italy*

(Received 11 May 2016; published 2 November 2016)

Markov state modeling (MSM) has recently emerged as one of the key techniques for the discovery of collective variables and the analysis of rare events in molecular simulations. In particular in biochemistry this approach is successfully exploited to find the metastable states of complex systems and their evolution in thermal equilibrium, including rare events, such as a protein undergoing folding. The physics of sliding friction and its atomistic simulations under external forces constitute a nonequilibrium field where relevant variables are in principle unknown and where a proper theory describing violent and rare events such as stick slip is still lacking. Here we show that MSM can be extended to the study of nonequilibrium phenomena and in particular friction. The approach is benchmarked on the Frenkel-Kontorova model, used here as a test system whose properties are well established. We demonstrate that the method allows the least prejudiced identification of a minimal basis of natural microscopic variables necessary for the description of the forced dynamics of sliding, through their probabilistic evolution. The steps necessary for the application to realistic frictional systems are highlighted.

DOI: 10.1103/PhysRevE.94.053001

I. INTRODUCTION

Despite the relevance of friction between solids from the macroscale to the nanoscale, its physical description still needs theoretical basis and systematic understanding. There is at present no theory of friction, one that should be based on a small number of relevant degrees of freedom and on some governing equation of motion. Even the simplest, classical atomistic nanoscale sliding simulation has too many degrees of freedom, with no available method for the unprejudiced identification of a handful of collective dynamical variables, suitable for a mesoscopic description of the sliding evolution and its relevant events, such as stick slip [1].

In the totally different field of equilibrium biomolecular simulations, powerful tools have been developed, aimed at identifying the metastable conformations, reactions paths, and rates associated to transition events between them. In particular, Markov state models [2–6] (MSMs) have emerged as a key technique, with clear theoretical foundations and great flexibility. In that approach, the phase space trajectory of a large collection of molecular entities is projected onto a much smaller discrete set of states that are deemed typical, and the dynamics is reduced to Markovian jumps between these states. So far MSMs were mostly applied to systems at equilibrium, where a stationary measure is defined, detailed balance holds, and the Markov description is natural. In friction on the other hand one deals with a strongly nonequilibrium dynamics, characterized by violent events, and where even in steady-state sliding, detailed balance is broken. Application of MSMs to nonequilibrium systems is still in its infancy, with just a few instances related to periodic driving [7] or cycle detection [8].

Here we show that the MSM framework can be extended, despite the difficulty represented by a time-growing phase space, to describe the forced dynamics of steady-state sliding friction. The resulting procedure, including the identification of the few slow collective variables ("excitations") relevant to steady-state sliding, the recognition of Markovian behavior, and the transfer matrix time evolution of probabilities, represents in our view a first step towards formulating a theory of friction, a methodological advance which we deem quite important.

To illustrate concretely this idea while keeping complications down to a minimum, we choose the simplest atomistic tribological model, the one-dimensional Frenkel-Kontorova (FK) model [9] in its atomic stick-slip regime [10,11]. For this model we explicitly show how the MSM construction leads to the identification of a handful of macrostates within which the basic frictional dynamics can be described.

II. MARKOV STATE MODELING

The normal procedure to build a MSM starts from running a long, "ergodic" molecular dynamics, exploring all the relevant configurations in phase space a sufficiently large number of times. By coarse-graining phase space, one then classifies all the configurations explored in this dynamics in a finite number of *microstates*. Each microstate is assumed to contain only structurally similar configurations. The classification in microstates is normally achieved by a clustering technique, such as k-means [12], or by performing a geometric partition of the configuration space, for example, by Voronoi tessellation.

These partitioning techniques require a *metric*, that is, a quantitative measure of the similarity and distance between configurations; the quality of the partition strongly depends on this choice. Only in very simple model systems can one use a metric taking into account all the coordinates of the system; in real world applications one must define the metric in a subset

*Corresponding author: tosatti@sissa.it

of the coordinates (for example, the coordinates of the solute). The choice of the metric is the single really arbitrary step in the procedure, to be done with utmost physical care.

Once the microstates are built, one estimates from the molecular dynamics trajectory the *transfer operator* Π^τ, namely, the transition probability matrix between pairs of microstates in a time τ. If the dynamics is ergodic and an equilibrium distribution is sampled, the matrix Π^τ has a single eigenvalue equal to one, with all its other eigenvalues being smaller than one. The components of the eigenvector associated with the eigenvalue one are proportional to the long-time probability of observing each microstate. The eigenvectors of Π^τ associated with eigenvalues close to one are associated with the slow modes of the system, while the smaller eigenvalues correspond to fast, irrelevant modes of motion. If there is a clear gap in between high and low eigenvalues in the spectrum of Π^τ, one can represent the dynamics in the reduced basis of the eigenvectors associated with the eigenvalues before the gap. This dimensional reduction can be carried out keeping under control the error it introduces by using standard techniques [13,14]

In order to apply this conceptual framework to the study of nonequilibrium problems, particularly of friction, the normal procedure must be modified and extended in several points. In short, the frictional dynamics does not reach an equilibrium, but only a steady state. The configuration space that is explored grows (approximately) linearly with the simulation time, which makes its sampling and subsequent clustering problematic. The solution is dividing the long evolution into intervals, deemed equivalent between them, so that results from each interval can be added on top of each other. The final test of whether the procedure is meaningful is the stability of results against extension of the interval size. This manner of building the microstates is possibly the crucial conceptual novelty which allowed us to deal with nonequilibrium steady states.

Since the metric is defined by a large number of microscopic variables, we did not build the microstates by the tessellation techniques used in standard MSM [15]. We built them instead by means of a recently proposed clustering algorithm [16], which associates each microstate to each meaningful peak of a (possibly multidimensional) probability distribution. Further coarse graining into an even smaller, minimal number of *macrostates* can then be implemented using standard methods [13,14], finally yielding the most compact MSM description of friction. In the space of these few macrostates, the time evolution of observables such as frictional work and displacement still reproduced the main features of the original frictional dynamics, represented in the FK model by the kink-antikink populations. This procedure and these steps are applicable to sliding systems of higher and generic complexity, which we expect to be described by similar, yet to be built, MSM approaches.

III. THE FRENKEL-KONTOROVA MODEL

The one-dimensional FK model, Fig. 1(a), our test case, consists of a chain of particles dragged over a sinusoidal potential $V(x) = A\cos(2\pi x/a)$. Nearest neighbor springs of stiffness k link L classical particles of mass m and positions x_i whose spacing a is commensurate with the periodic potential. Each particle is dragged by a spring of constant κ moving with

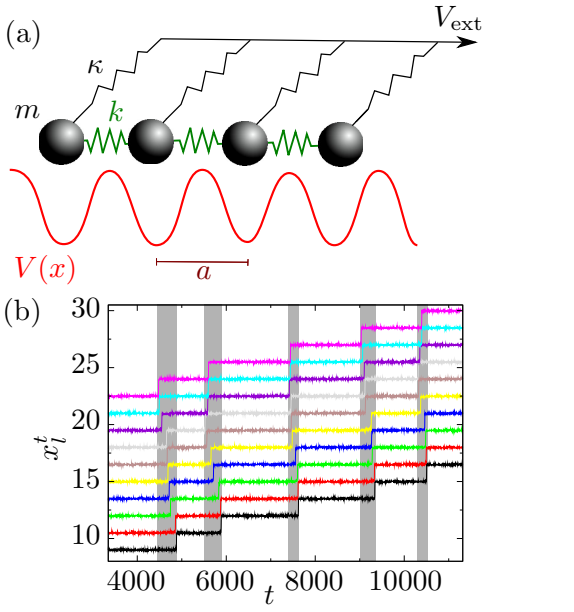

FIG. 1. (a) Schematic of the FK system. (b) Sample of steady-state motion of $L = 10$ particles with parameters $k = 0.04$, $A = 0.1$, $a = 1.5$, $m = 1$, $\gamma = 1$, $\beta = 500$, $\kappa = 0.01$, and $v_{\rm ext} = 0.001$. The white and gray backgrounds represent stick and slip time domains, respectively.

constant velocity $v_{\rm ext}$. Particle motion obeys an overdamped Langevin dynamics (large damping γ), in a bath of inverse temperature $\beta = 1/k_B T$:

$$x_l^{t+dt} = x_l^t + \left[\frac{2\pi A}{\gamma m a}\sin\left(\frac{2\pi x_l^t}{a}\right) + \frac{\kappa}{\gamma m}\left(v_{\rm ext}t - x_l^t\right) - \frac{k}{\gamma m}\left(2x_l^t - x_{l-1}^t - x_{l+1}^t\right)\right]dt + \sqrt{\frac{2dt}{\gamma m\beta}}\,f^t, \quad (1)$$

where f^t is an uncorrelated Gaussian distribution and dt is the elementary time step (here $dt = 10^{-2}$). Our input for the MSM procedure is a long steady-state trajectory of the chain, obtained by integrating these equations, mostly for the simplest case of $L = 10$ (but also $L = 15,20$) and a sufficient duration of 10^6 time units.

As is well known [10], in a wide range of parameters the chain sliding alternates long sticking periods during which particles are close to their respective potential minima with fast slips during which one or more lattice spacings are gained. This kind of atomic stick-slip motion is well established for, e.g., the sliding of an atomic force microscope tip on a crystal surface [17] – of course, involving in that case three-dimensional displacements of larger complexity. A slip event involves the formation of kink-antikink defects (large deviations of the interparticle distance from the equilibrium value) that propagate along the chain and enable the global movement. A sample of steady-state sliding evolution can be seen in Fig. 1(b), showing the finer details of each particle's motion for a few slip events.

Based on the long steady-state trajectory in the original positional phase space, the building of the MSM involves the following steps: definition of the transfer operator and transfer

matrix (TM), choice of phase space metric, clustering into microstates, and construction of macrostates and their reduced TM dynamics.

IV. THE TRANSFER OPERATOR

Our MSM analysis is based on the transfer operator (TO) formalism [5]. Denote by $\Pi^\tau(X \to X')$ the probability to go from a configuration $X^t = X$ at time t to $X^{t+\tau} = X'$ at time $t+\tau$. While Π^τ is a continuous operator which takes infinite time to sample, we built a coarse-grained TO by partitioning the configuration space into microstates (ensembles of similar configurations) $\{c_\alpha, \alpha = 1, \ldots, n_c\}$. Between these microstates the restricted TO is a finite $n_c \times n_c$ transfer matrix with the generic element $\Pi^\tau_{\alpha\beta} = \int_{X\in c_\alpha}\int_{X'\in c_\beta} dX\, dX' P(X)\Pi^\tau(X \to X')$, the probability to go from c_α to c_β in time τ. This TM contains less detail than Π^τ. Being simpler, it is more informative and can be sampled with satisfactory statistics in finite time. In general $\Pi^\tau_{\alpha\beta}$ depends on the lag time τ, but there are techniques to control these variations. (See Sec. 1.1 of the Supplemental Material [18].) We calculated the eigenvalues $\{\lambda_i\}$ and left eigenvectors $\{\vec{\chi}_i\}$ of the TM. Because we are not in equilibrium, detailed balance does not hold, the TM is not symmetric, and the eigenvalues are not necessarily real. However, they still satisfy $|\lambda_i| \leqslant 1$ by the Perron-Frobenius theorem. The largest (modulus-wise) eigenvalue is exactly 1, and if the evolution is ergodic there is only one such eigenvalue. The eigenvector $\vec{\chi}_1$ represents the invariant, steady-state distribution, endowed with nonzero sliding current. The eigenvectors $\vec{\chi}_i$ with $|\lambda_i| \simeq 1$ form the so-called Perron cluster [13]. They characterize the long-lived excitations of the steady state, which decay with long characteristic times $\tau_i = -\tau/\ln|\lambda_i| \gg \tau$, while oscillating with period $\tau/\arctan(\mathrm{Im}\lambda_i/\mathrm{Re}\lambda_i)$.

V. CHOICE OF THE METRIC

The phase space explored under steady sliding does not possess a preexisting metric; it also grows linearly with time and is thus very poorly sampled (except for variables internal to the slider). To overcome the sampling problem, the natural prescription is to break up the dynamics into many long intervals between which memory is lost and Markovian behavior can be verified, while at the same time good statistics can be accumulated within each interval. For the building of a viable metric there is no systematic approach, and the choice must be guided by physical considerations; even if arbitrary, one should then try to make that choice as judicious as possible. In the FK model, we took our metric to include internal variables, the bond lengths $b_l^t = (x_{l+1}^t - x_l^t - a)/a$, and an external variable, the center of mass coordinate (CM): $x_{\mathrm{CM}}^t = \frac{1}{Na}\sum_{l=1}^L x_l^t$. The sampling of x_{CM} can be improved by considering multiple independent slices (intervals) of the steady-state evolution, each long enough that all relevant events have occurred (here, frictional slips), setting then "absorbing" boundary conditions for any transition from and to the outside of each range. While this route could be followed without problems, we preferred here to set artificial periodic boundary conditions for this slice of phase space, a choice which provides a more compelling picture of steady-state sliding, and where the error involved in the transition rates can anyway be reduced at will by extending the slice size. In the FK model one can exploit in addition the substrate periodicity and differences in x_{CM} are taken modulo na for a chosen integer $n > 1$. Under slow driving, we found that $n = 2$ is sufficient for a correct description of slips by a (atomic slip), and states divide into even and odd x_{CM}. If slips of $2a$, $3a$ or more became frequent, we would simply choose a larger n. The full set of steady-state sliding data was used to generate many independent configurations, all treated in the same manner. Summing up, the metric we adopt defines the distance between configurations at times s and t as

$$d_{st} = \left[\left(x_{\mathrm{CM}}^s - x_{\mathrm{CM}}^t\right)_{\mathrm{mod}\,2}\right]^2 + \sum_{l=1}^{L-1}\left(b_l^s - b_l^t\right)^2. \qquad (2)$$

We stress again that the choice of metric and the handling of phase space growth are by no means unique and must be decided case by case. Yet we believe that a very similar physical description would emerge from different choices, so long as they are made on sensible physical grounds. In that broad sense, our description is minimally prejudiced.

VI. MICROSTATES

In the next step, configurations whose relative distance is small were collected together, in n_c microstates. Microstates were built by the Density Peak algorithm [16], which efficiently traces them as maxima of the probability density in phase space. Given a distance d_{st} between two configurations X^s and X^t we estimate the local density ρ^s in X^s by counting the number of configurations within a cutoff d_c, $\rho^s = \sum_t \theta(d_c - d_{st})$, where θ is the step function. One then computes the distance δ^s between X^s and the closest configuration of higher density, $\delta^s = \min_{\rho^t > \rho^s} d_{st}$ and identifies the microstate centers as the n_c points with the highest product $\delta^s\rho^s$. All remaining points are assigned to the microstate of highest local density. This clustering technique allows finding microstates of variable volume in phase space and well-defined cluster centers (configurations often visited), both desirable features in building a MSM.

We used samples of $N = 10^5$ configurations (separated by the lag time τ) and clustered them using the metric (2). The optimal lag time τ was determined by studying the evolution of the spectrum of the clustered TM with τ (see Sec. 1 of Ref. [18]). We found a plateau around the value $\tau = 10 = 1000dt$ (dt being our time unit) showing that the dynamics is Markovian in this range. With $\tau = 10$, the algorithm detected about $n_c \simeq 100$ microstates (clusters). Below the largest eigenvalue $\lambda_1 = 1$, the spectrum of the $n_c \times n_c$ TM is characterized by a second eigenvalue λ_2 (see Fig. 2), corresponding to a relaxation time of $\simeq 600$, separated by a gap from other eigenvalues with shorter relaxation times. The significance of the eigenmodes χ_i is clarified by considering the probability distribution $P(O,t)$ of an observable O at time t, starting from a system prepared in the mixed state P_α^0 (probability vector to be in c_α at $t = 0$). We have

$$P(O,t) = P^{\mathrm{ss}}(O) + \sum_{i>1} f_i g_i(O) e^{-t/\tau_i}, \qquad (3)$$

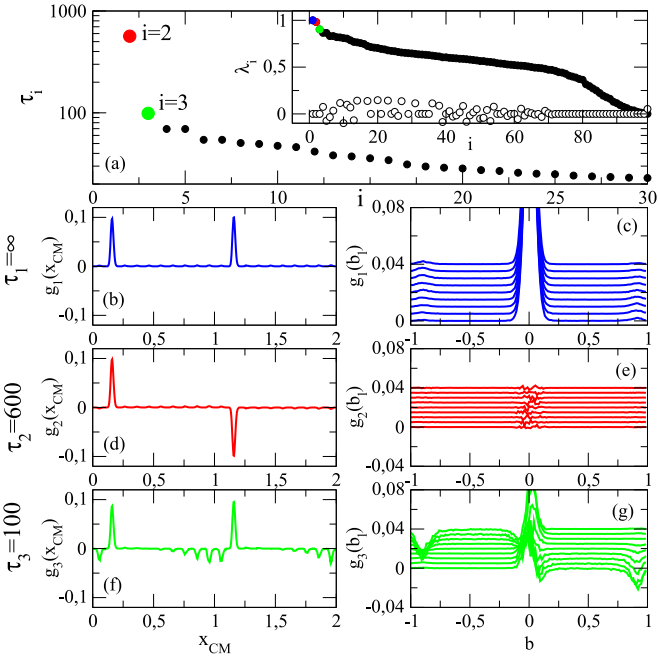

FIG. 2. (a) Characteristic time scales and eigenvalues (in the inset, imaginary part in white) of the TM (averaged over 10 realizations with $N = 10^5$ each). (b), (d), and (f) Probability distribution $g_1(x_{CM})$ and perturbations $g_i(x_{CM})$ for the first three eigenvectors of the TM. (c), (e), and (g) These same functions for the bonds b_l, $g_i(b_l)$ (spaced vertically for clarity).

where $f_i = \sum_\alpha \chi_i^\alpha P_\alpha^0 / P_\alpha^{ss}$ accounts for the initial condition, and

$$g_i(O) = \sum_\alpha \chi_i^\alpha P(O|\alpha), \tag{4}$$

where $P(O|\alpha)$ is the probability distribution of O in microstate α, $P^{ss}(O) = g_1(O)$ the steady-state distribution of O, and P_α^{ss} the steady-state probability to visit microstate α. The $g_i(O)$ for $i > 1$ represent "perturbations" of $P^{ss}(O)$, each decaying within the lifetime τ_i.

These functions provide a direct insight into the slow eigenmodes: we now consider the first three eigenstates in detail and the observables contained in our metric; any other observable could be similarly analyzed to inspect its relevance to the slow eigenstates. In Figs. 2(b), 2(d), and 2(f) we plot $g_i(x_{CM})$: the steady state χ_1 consists of one large peak per period plus nine smaller peaks, corresponding to the relaxed chain state and defect combinations, respectively. The second eigenvector χ_2 presents exactly the same features, except for a factor -1 in the second period: combinations $\chi_1 \pm \chi_2$ represent the chain CM sticking either in an odd or even position (the global sign of an eigenvector is arbitrary). The second eigenvector is thus representative of the main advancing motion of the chain, namely, the slip. Indeed, $t_2 = -\tau/\log(\lambda_2) \simeq 600$ is about half the sticking time, which is the average jump time (see Fig. 1). In Figs. 2(c), 2(e), and 2(g) we plot g_i $(i = 1,2,3)$ for the bonds lengths b_l. In the steady state χ_1 each bond length has high peaks around its value at rest (0) and smaller ones around $b \simeq \pm 0.9$, reflecting the infrequent appearance of excitations, which are kinks or antikinks. The second mode χ_2 shows a flat distribution, since all the difference with χ_1 lies in the CM degree of freedom, x_{CM}. In fact, χ_3 displays small central peaks and more pronounced lateral ones, corresponding to the creation (and destruction for negative contributions) of a kink or antikink (depending on the sign of b_l) [10], excitations with shorter lifetimes. Indeed, $t_3 \simeq 100$ is comparable to the half-lifetime of kinks and antikinks (the next eigenstates exhibit similar time scales and features, not shown). Furthermore, we can see how the peaks are around $b \simeq 0.9$ for the first b_l and around $b \simeq -0.9$ for the last ones, implying that the chain tends to be elongated in its head and compressed in its tail. This shows that kink-antikink pairs are more likely to be formed in the center of the chain, intrinsic to the slip mechanism for this system. At this stage one can already identify the kink and antikink populations as the relevant collective variables of sliding, together with x_{CM}. While the basic features of commensurate FK stick slip are already contained in the first few long-lived eigenmodes, there are still too many large eigenvalues for an extensive analysis to be possible (no gap is apparent). A more accurate description must involve a sharper definition and quantitative analysis of the whole Perron cluster.

VII. MACROSTATES

The number of microstates must be reduced for a reasonable coarse-grained theory. In the next and final step, the n_c microstates were coarse-grained and grouped into macrostates, according to the kinetics of intermicrostate jumps. The idea is that microstates that have different structures but are very well connected kinetically should belong to the same macrostate. A well-established approach for that is the (Robust) Perron Cluster Cluster Analysis (PCCA+) [13,14] (see also Sec. 2.1 of Ref. [18] and Ref. [19]). For technical reasons, and strictly for this purpose, we symmetrized the TM, thus temporarily neglecting the nonequilibrium breaking of detailed balance (the exact asymmetrical TM could also be used if necessary; see more recent developments [20,21] and Sec. 2.1 of Ref. [18]). We found that relevant macrostates could be further reduced from $n_c \sim 100$ down to as little as $\tilde{n}_c = 6$. Moreover, whereas n_c grows with system size L, $\tilde{n}_c = 6$ is much more stable against L: we found a consistent description of the system with $\tilde{n}_c = 6$ also for $L = 15$ and $L = 20$, and the detection of the optimal $\tilde{n}_c$ consistently yielded $\tilde{n}_c \in [5,9]$ (see Ref. [18] and Figs. 3 and 4). In Fig. 3 we present the six macrostates $\{\tilde{c}_\alpha\}$, displaying some of the configurations (microstate cluster centers) which they contain.

Macrostates $\tilde{c}_1$ and $\tilde{c}_4$ include the relaxed chain microstates, along with some single excitations at the tips; $\tilde{c}_2$ and $\tilde{c}_5$ contain mostly single kinks, while $\tilde{c}_3$ and $\tilde{c}_6$ contain mostly antikinks. The microstates with (kink, antikink) pairs are spread between groups, with neighboring pairs belonging to $\tilde{c}_{1,4}$ and extended pairs to others. The only difference between the triplets of $\tilde{c}_{1,2,3}$ and $\tilde{c}_{4,5,6}$ is in the value of x_{CM}, respectively, $x_{CM} \approx 0.15$ and 1.15. Overall, this description provides a qualitative understanding of the basic mechanisms of slips complementary to that of our kinetic analysis and allows us to directly read the kink and antikink populations as the collective variables describing sliding (not relying on any choice of distance). In Fig. 3(b) observation of the 6×6 TM $\tilde{\Pi}_{\alpha\beta}$

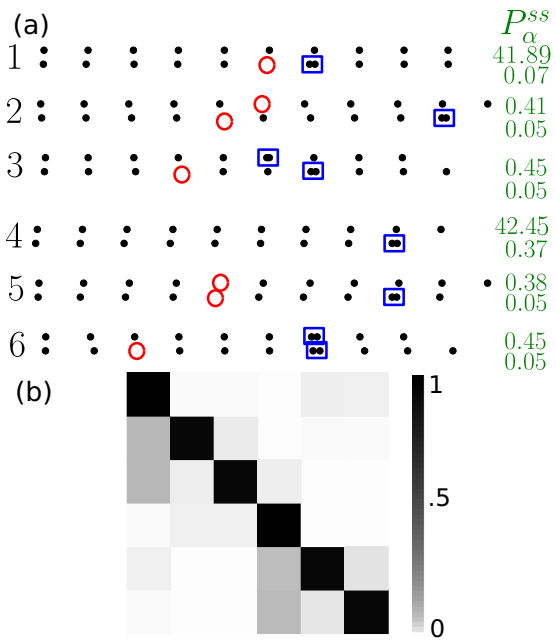

FIG. 3. (a) Selection of microstates inside the six macrostates identified with PCCA+: the atoms positions relative to potential minima (black dots) display kinks (red circles) and antikinks (blue squares). For clarity, the probability P^{ss}_α is multiplied by 100, and we show only 12 of the $\sim$100 microstates. (b) Representation of the reduced transition matrix $\widetilde{\Pi}_{\alpha\beta}$ (gray scale proportional to magnitude).

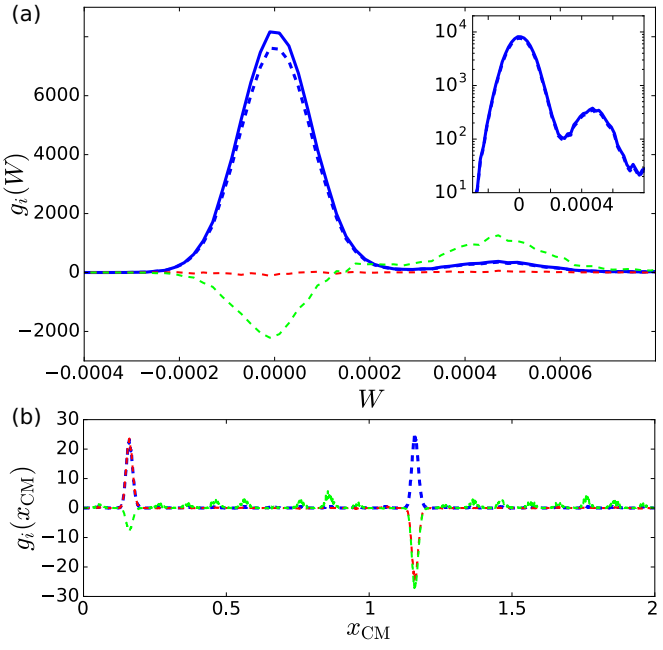

FIG. 4. Comparison of (a) work $g_i(W)$ and (b) center of mass $g_i(x_{CM})$ distributions for $i = 1,2,3$ (blue, red, green). Solid lines: raw data; dashed lines: $\tilde{n}_c = 6$ macrostate results. Inset: Blowup of $g_i(W)$, highlighting the excess probability for $W > 0$, signaling the positive frictional work. Note how $i = 3$ (green) is the excitation of the steady state (blue) that populates the $W > 0$ tail. Note how (b) is similar to Figs. 2(b), 2(d), and 2(f).

reveals, e.g., that motion (looping through states $\tilde{c}_{1;4}$) occurs only through excited states $\tilde{c}_{2,3;5,6}$. Additional details about the role of each macrostate is given in Sec. 2.3 of Ref. [18].

VIII. SLIDING FRICTION DESCRIBED BY MACROSTATE EVOLUTION

The coarse-grained model thus found with PCCA+ provides a probabilistic (Markovian) description of friction as a MSM. In the macrostate representation, the probabilities $P^t_\alpha = P(X^t \in \tilde{c}_\alpha)$ evolve in time as $P^{t+\tau}_\alpha = \widetilde{\Pi}_{\alpha\beta} P^t_\beta$. For the whole construction to be satisfactory, this coarse-grained frictional evolution should reproduce the quantitative aspects of raw data.

We compare in Fig. 4(a) the distribution of frictional work $P^{ss}(W)$ in the raw data (using all individual configurations) with that computed in our MSM, i.e., the steady state $g_1(W)$ relative to $\widetilde{\Pi}_{\alpha\beta}$. For that, to each macrostate $\tilde{c}_\alpha$ we associated its distribution $P(O|\alpha)$, i.e., the distribution of O restricted to the individual configurations of macrostate $\tilde{c}_\alpha$. Using Eq. (4), we could obtain any $g_i(O)$ of our MSM $\widetilde{\Pi}_{\alpha\beta}$. The average particle current in the reduced basis is $\langle J \rangle = 6.82 \times 10^{-4}$, close to the exact value $\langle J \rangle = 6.66 \times 10^{-4} = v_{ext}/a$. Similar agreement was found for the steady state of x_{CM}. Strikingly, not only the steady state but also the excitations $g_i(x_{CM})$ within our MSM reproduced well those of the n_c states description; see Fig. 4(b) and Fig. 4 of Ref. [18]. The lifetimes corresponding to these modes, and more precisely the decay of the correlation functions of various observables also matched well the respective correlation functions evaluated on the raw data (not shown).

We conclude that MSM provides a viable description of sliding friction in the FK model, including all the important excitations and slow events that govern the phenomenon. The remaining fast variables and more ephemeral excitations are in this manner systematically integrated away to the best possible level.

IX. CONCLUSIONS

The Markov State Model method, so far developed for the equilibrium evolution of large-scale molecular systems, can be naturally extended to nonequilibrium dynamics under the action of external forces. Among nonequilibrium phenomena, the physics of sliding friction is in desperate need of a description, with coarse-grained variables and their time evolution constructed in the least prejudiced manner. We formulated and carried out the first analysis of frictional sliding through the MSM method extended to nonequilibrium, whose result is just that.

Starting with a steady-state dynamical simulation, the main initial step is, as in all MSM applications, the choice of a metric in phase space, after which the approach built without further bias or approximations the microstates that track the effective dynamical variables of frictional sliding. As needed, the number of microstates could be reduced down to a handful of macrostates by successive application of standard methods such as PCCA+.

In the FK model, chosen for this first demonstration, the coarse graining is sharp enough to capture not only overall

steady-state observables such as average dissipated power or average current, but also their modes of excitations and their correlations, as shown by the excitations $g_i(O)$. This brings all the important, slow dynamical features under control in a manner which, as far as we know, is unprecedented for violent and nonlinear frictional motion. All these steps are generically applicable to more complex and realistic sliding problems. Further developments could introduce biased sampling favoring exploration of rare transitions, to improve the statistics more efficiently [22]. Summing up, this work introduces a theoretical approach to friction not limited to nanoscale sliding but potentially extendable to other driven systems.

ACKNOWLEDGMENTS

Work was carried out under ERC Advanced Research Grant No. 320796 MODPHYSFRICT. COST Action MP1303 is also acknowledged. F.L. thanks R. Dandekar for useful discussions.

[1] E. Gnecco and E. Meyer, *Fundamentals of Friction and Wear on the Nanoscale* (Springer, New York, 2015).
[2] F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weikl, Proc. Nat. Acad. Sci. USA **106**, 19011 (2009).
[3] C. R. Schwantes, R. T. McGibbon, and V. S. Pande, J. Chem. Phys. **141**, 090901 (2014).
[4] F. Noé and F. Nüske, Multiscale Model. Simul. **11**, 635 (2013).
[5] G. R. Bowman, V. S. Pande, and F. Noé, in *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation* (Springer, New York, 2014).
[6] C. Schütte and M. Sarich, Eur. Phys. J. Spec. Top. **224**, 2445 (2015).
[7] H. Wang and C. Schütte, J. Chem. Theory Comput. **11**, 1819 (2015).
[8] F. Knoch and T. Speck, New J. Phys. **17**, 115004 (2015).
[9] Y. I. Frenkel and T. A. Kontorova, Phys. Z. Sowietunion **13**, 1 (1938).
[10] O. M. Braun and Y. S. Kivshar, Phys. Rep. **306**, 1 (1998).
[11] M. Paliy, O. Braun, T. Dauxois, and B. Hu, Phys. Rev. E **56**, 4025 (1997).
[12] G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé, J. Chem. Phys. **139**, 015102 (2013).
[13] P. Deuflhard and M. Weber, Linear Algebra Appl. **398**, 161 (2005).
[14] M. Weber and S. Kube, Lect. Notes Comput. Sci. **3695**, 57 (2005).
[15] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and F. Noé, J. Chem. Phys. **134**, 174105 (2011).
[16] A. Rodriguez and A. Laio, Science **344**, 1492 (2014).
[17] A. Vanossi, N. Manini, M. Urbakh, S. Zapperi, and E. Tosatti, Rev. Mod. Phys. **85**, 529 (2013).
[18] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.94.053001 for technical details on the calculation.
[19] https://bitbucket.org/flandes/msm_fk_densitypeakalgo_rpcca/ (2016).
[20] S. Roeblitz, Ph.D. thesis, *Statistical Error Estimation and Grid-free Hierarchical Refinement in Conformation Dynamics*, Freie Universität Berlin (2009).
[21] N. D. Conrad, M. Weber, and C. Schütte, ZIB Report 15-40 (2015).
[22] J. D. Chodera, N. Singhal, V. S. Pande, K. A. Dill, and W. C. Swope, J. Chem. Phys. **126**, 1 (2007).