scientific reports

OPEN

Sampling rare conformational transitions with a quantum computer

Danial Ghamari[1,2], Philipp Hauke[3], Roberto Covino[4✉] & Pietro Faccioli[1,2✉]

Structural rearrangements play a central role in the organization and function of complex biomolecular systems. In principle, Molecular Dynamics (MD) simulations enable us to investigate these thermally activated processes with an atomic level of resolution. In practice, an exponentially large fraction of computational resources must be invested to simulate thermal fluctuations in metastable states. Path sampling methods focus the computational power on sampling the rare transitions between states. One of their outstanding limitations is to efficiently generate paths that visit significantly different regions of the conformational space. To overcome this issue, we introduce a new algorithm for MD simulations that integrates machine learning and quantum computing. First, using functional integral methods, we derive a rigorous low-resolution spatially coarse-grained representation of the system's dynamics, based on a small set of molecular configurations explored with machine learning. Then, we use a quantum annealer to sample the transition paths of this low-resolution theory. We provide a proof-of-concept application by simulating a benchmark conformational transition with all-atom resolution on the D-Wave quantum computer. By exploiting the unique features of quantum annealing, we generate uncorrelated trajectories at every iteration, thus addressing one of the challenges of path sampling. Once larger quantum machines will be available, the interplay between quantum and classical resources may emerge as a new paradigm of high-performance scientific computing. In this work, we provide a platform to implement this integrated scheme in the field of molecular simulations.

Molecular dynamics (MD) simulations enable us to investigate the structure and dynamics of molecular systems at high spatial and temporal resolution[1]. Despite their large success, MD simulations face the challenge of sampling rare thermally activated re-organizations of complex systems, e.g., conformational changes, folding, and phase transitions[2]. Indeed, in a typical simulation, an exponentially large fraction of the computational time is employed to simulate thermal fluctuations in meta-stable states, rather than sampling the fast transition paths—the stochastic jumps between states—which are rare events[2].

A wide spectrum of clever enhanced sampling methods have been developed over the last two decades to overcome the rare events sampling problem[3]. Some of these techniques reach a high computational efficiency by introducing history-dependent biasing forces that drive the system out of its thermal equilibrium, thus promoting the escape from meta-stable states (see, e.g.,[4–7]). The biasing forces depend on collective variables (CVs), which should encode the essential low-dimensional features of a molecular rare event[8]. In practice, identifying optimal CVs is a very hard problem, and in realistic conditions sub-optimal CVs will affect the quality of the sampling and the accuracy of the mechanistic understanding emerging from the simulations[8].

As an alternative approach, Transition Path Sampling (TPS)[9] is a Markov Chain Monte Carlo scheme that in principle samples the transition path ensemble without involving any biasing force, nor a choice of CVs. In TPS, plain MD simulations generate a trial move, i.e., the attempt to generate a new transition path. For instance, in the so-called shooting move, a new trajectory is initialized from a configuration randomly selected from the last stored transition path[10]. Yet, when applied to complex transitions occurring in large configuration spaces with rugged energy landscapes, TPS faces two challenges: efficiently generating viable trial trajectories at an acceptable computational cost and reducing the correlation of generated paths[11].

[1]Department of Physics, University of Trento, Via Sommarive 14, Trento 38123, Italy. [2]INFN-TIFPA, Via Sommarive 14, Trento 38123, Italy. [3]INO-CNR BEC Center & Department of Physics, University of Trento, Via Sommarive 14, Trento 38123, Italy. [4]Frankfurt Institute for Advanced Studies, Ruth-Moufang-Straße 1, Frankfurt am Main 60438, Germany. ✉email: covino@fias.uni-frankfurt.de; pietro.faccioli@unitn.it
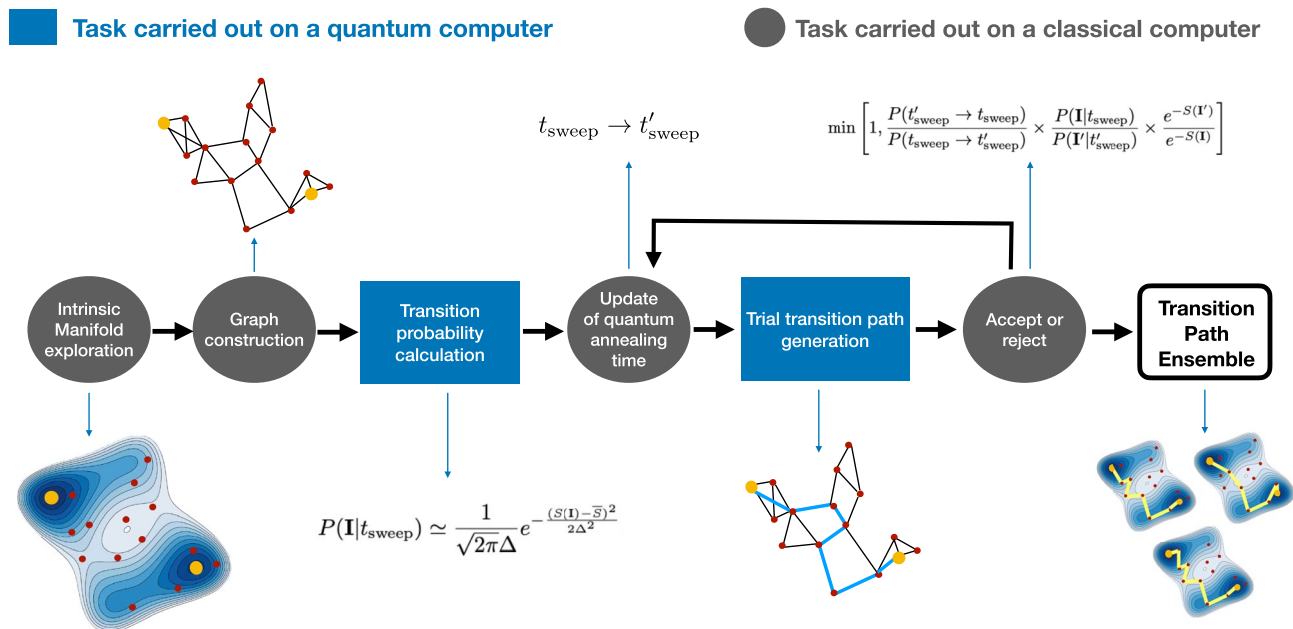
Figure 1. Schematic representation of the path sampling scheme introduced in this work, which combines ML and MD performed on a classical computer with QC performed on a quantum annealing machine. Our scheme samples the full transition path ensemble without any use of CVs or unphysical biases.

Even though promising advancements have recently been made by integrating molecular simulations with machine learning (ML) (see, e.g.,[12–18]), the quest for computationally affordable and accurate enhanced sampling of complex molecular systems remains open. In this endeavor, rapid advances in quantum computing provide new opportunities, as is illustrated in the context of quantum chemistry and biology by pioneering applications[19–26]. Over the last few years, quantum hardware has grown exponentially both in size and performance[27–29], to a point that it is now realistic to foresee the onset of a tangible quantum advantage in computational problems[30,31]. It is therefore both important and timely to address the question whether MD, ML, and Quantum Computing (QC) can join forces to tackle outstanding challenges of molecular simulations.

In this work, we integrate MD, ML, and QC to sample the transition path ensemble of thermally activated rare events without involving any unphysical bias or choice of CVs. The salient features of this scheme are illustrated in Fig. 1: First, ML and MD perform a preliminary uncharted exploration of the most visited regions of the configuration space[32]. These data are used to derive a general coarse-grained description of rare events based on Langevin dynamics. Then, QC on a quantum annealing machine[33–37] generates transition paths connecting the previously generated configurations. These paths are then accepted or rejected according to a Metropolis criterion implemented on a classical computer, which combines the statistical mechanics of the transition path ensemble with the internal physics of the quantum annealing machine, for which we used the D-Wave machine[38].

Importantly, at each iteration the quantum computer generates a new viable and uncorrelated transition path. In fact, at every step the quantum computer is re-initialized in an equal superposition of all computational basis states, which erases all memory of the previously sampled transition paths. In this way, we harness one of the defining features of quantum annealing to overcome a key limitation of path sampling algorithms. In addition, we exploit non-adiabatic effects in the annealing procedure (controlled by the sweep time) and the presence of classical fluctuations (due to intrinsic noise of the physical machine) to explore the transition path ensemble.

As a first illustrative application, we sample the transition path ensemble of a conformational transition in alanine dipeptide. Even though our approach is fully general and scalable, we have chosen this benchmark system as it is sufficiently small to enable us to encode and run our algorithm on presently existing D-Wave machines, yet it recapitulates the features of rare events in molecular systems. Our benchmark results agree well with those obtained by plain MD. We use an auto-correlation analysis to demonstrate that the quantum computer generates uncorrelated trial transition paths at every Monte Carlo step. Though existing quantum machines permit us to perform benchmarks only on proof-of-concept systems of limited complexity, the ongoing exponential growth in size and efficiency of quantum computing hardware suggests that, in the future, our approach might help us to investigate transitions that are currently challenging for state-of-the-art classical sampling methods.

The manuscript is organized as follows. In "Uncharted exploration of the intrinsic manifold" and "Coarse-grained representation of the dynamics on the intrinsic manifold" Section, we introduce the general theoretical framework, describe the algorithm used to perform the uncharted exploration of the intrinsic manifold, and our coarse-grained description of reactive processes. In 'Transition Path Sampling with a quantum annealer' Section, we discuss the encoding of the path sampling problem on a quantum annealing machine and derive our hybrid Monte Carlo scheme that combines classical and quantum computing. In 'Application to a molecular benchmark system' Section, we report on our illustrative application to alanine dipeptide. The main results are summarized and discussed in "Conclusions" Section.
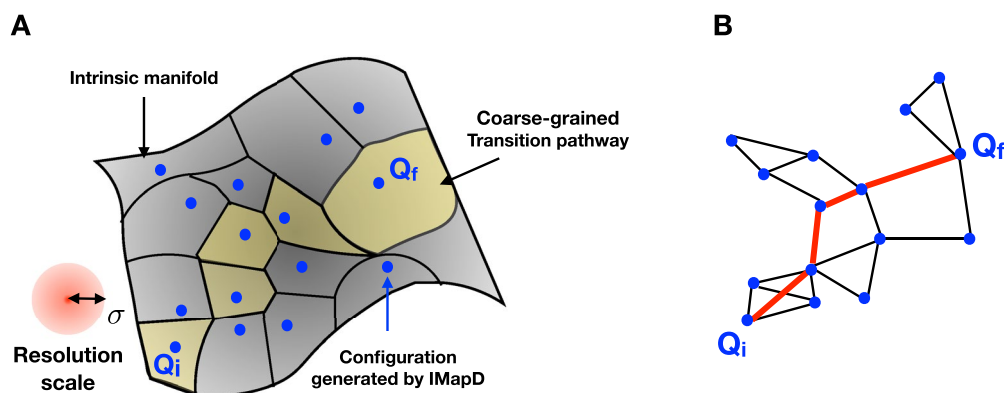
Figure 2. (**a**) Illustration of the coarse-grained representation of the Langevin dynamics. Dots represents configurations generated using iMapD, which lie on the intrinsic manifold. Each such configuration is regarded as a representative element of its Voronoi cell. Coarse-grained trajectories are identified by sequences of Voronoi cells, *I*. A typical transition path is highlighted in yellow. The intrinsic resolution scale σ of the effective theory is set by the average distance between the configurations generated by iMapD. (**b**) Graph representation of the coarse-grained Langevin dynamics. The red line denotes the transition path highlighted in yellow in panel (**a**).

Theoretical setup

In a molecular system at thermal equilibrium, the statistically relevant configurations accumulate in low-dimensional regions that define the so-called intrinsic manifold. Our path sampling algorithm exploits a recently developed scheme to efficiently explore this intrinsic manifold[32] (step 1 in Fig. 1). Then, it relies on a coarse-grained representation of the dynamics that is defined directly on this manifold, based on the configurations generated during the exploration, to define the input parameters for the quantum annealing part of our algorithm, discussed in "Transition Path Sampling with a quantum annealer" Section (step 2 in Fig. 1).

Uncharted exploration of the intrinsic manifold.. To efficiently explore the intrinsic manifold and sample relevant molecular configurations without any use of CVs or nonphysical bias, we rely on Intrinsic Map Dynamics (iMapD)[32]. In iMapD, a data-driven manifold learning technique—diffusion maps[39]—empowers the unbiased MD sampling. Given some local sampling, diffusion maps identify the boundary of the explored configuration space in a low-dimensional representation. New unexplored configurations in the vicinity beyond this boundary are identified, from which we start a further round of local unbiased sampling. By iterating between these two steps, iMapD rapidly explores the relevant parts of the configuration space. In the sections S1 and S3 of the Supplementary Information (SI), we summarize the key aspects of the theory and implementation of iMapD.

Using iMapD we obtain a sparse data set of configurations $\mathscr{C} = \{Q_k\}_{k=1,\dots,\nu}$, that by construction lie on the intrinsic manifold, and that were obtained at a much lower computational cost than by running equilibrium MD[32].

Coarse-grained representation of the dynamics on the intrinsic manifold. Once the data set $\mathscr{C}$ is established, we can use it to build a coarse-grained representation of the dynamics, defined directly on the intrinsic manifold explored with iMapD.

The sparse configuration data set obtained after the uncharted exploration defines a partition of the intrinsic manifold in finite sub-regions. The *i*-th region is identified with the neighborhood of configuration $Q_i \in \mathscr{C}$. For example, in a Voronoi tassellation, Q_i would represent the center of a cell whose boundaries lie midway to the neighboring sampled configurations (see Fig. 2).

The spatial resolution scale σ of this coarse-grained representation of the molecular dynamics is set by the average distance between neighboring configurations in the data set $\mathscr{C}$. The temporal resolution Δt is estimated by the average time the system takes to diffuse across neighboring regions. At this level of coarse-graining, transition pathways correspond to ordered sequences of visited sub-regions (see, e.g., the yellow regions in Fig. 2). Therefore, a transition path can be specified by an integer vector $\mathbf{I} = (i_1, \dots, i_{N_I})$, where i_k is the label pointing at the neighborhood of the configuration Q_{i_k}, which is visited at the *k*-th time step.

We develop a statistical mechanical formalism that enables us to describe the coarse-grained dynamics on the intrinsic manifold. This is necessary to correctly evaluate the probability of each coarse-grained transition path. We employ a powerful path-integral formalism combined with regularization and renormalization procedures that were originally developed in the framework of nuclear and subnuclear physics (see[40] for an enlightening pedagogical introduction).

Path integral formulation of stochastic dynamics Assuming a diffusive dynamics, we derived an expression for the probability of arbitrary coarse-grained paths $\mathbf{I}$ on the intrinsic manifold, in the form

$$P(\mathbf{I}) \propto e^{-S(\mathbf{I})}, \tag{1}$$

where the functional $S(\mathbf{I})$ is called the effective action of the coarse-grained path $\mathbf{I}$. Section S2 of the SI contains a self-contained derivation of the final expression for $S(\mathbf{I})$. What is important to note here, is that to spatially

3

coarse-grain the dynamics we "smear" the positions to the scale σ. This is equivalent to lowering the spatial resolution of the dynamics through a procedure that, in the framework of Renormalization Group theory, is commonly referred to as "position space regularization".

In practice, we associate each configuration $Q_i \in \mathscr{C}$ with a finite region of the configuration space. We then derive the expression of the probability of a coarse-grained transition path (Fig. 2) by using a Feynmann path-integral. The explicit evaluation of this integral is quite lengthy and is detailed in section S2 of the SI. The final result is:

$$K_{\text{cg}}(Q_f, t | Q_i, t_i) = \int_{Q_i}^{Q_f} \mathscr{D} \, Q \, e^{-\frac{1}{2mD_{\text{cg}}} \int_{t_i}^{t_f} d\tau \left(m\frac{\dot{Q}^2}{2} + V_{\text{cg}}[Q(\tau)] \right)}. \tag{2}$$

This is the probability in the coarse-grained dynamics to observe a transition between the regions Q_f and Q_i in a time $t - t_i$. $D_{\text{cg}} = \sigma^2/(2\Delta t)$ and $V_{\text{cg}}(Q)$ are the diffusion coefficient and effective potential of the coarse-grained theory, respectively. The definition of $V_{\text{cg}}(Q)$ and the numerical scheme we used to estimate it are also discussed in the SI.

Path probability in the graph representation Equation (2) leads to a closed expression for the coarse-grained effective path action entering Eq. (1) (for details, see section S5 in the SI):

$$S(\mathbf{I}) = \sum_k w_{i_{k+1} i_k}, \tag{3}$$

where

$$w_{ij} \equiv \frac{1}{2mD_{\text{cg}}} \left(\frac{m(Q_i - Q_j)^2}{2\Delta t} + V_{\text{cg}}(Q_i)\Delta t \right). \tag{4}$$

The exponential $e^{-w_{ij}}$ controls the probability of observing a transition between the regions centered around Q_i and Q_j of the intrinsic manifold, in an elementary (coarse-grained) time step Δt. We emphasize that, by definition of D_{cg} and V_{cg}, the weights given in Eq. (4) are small numbers. As a result, probabilities of different paths on the graph are comparable. Further details about explicit evaluation of the weights w_{ij} in a realistic application are reported in section S5 in the SI.

The present coarse-grained representation of the stochastic dynamics displays several analogies with Markov state modelling. A key difference is that, in our Renormalization Group inspired approach, the information about the kinetics is encoded in the renormalised effective potential, $V_{cg}(Q_i)$ and not in a stochastic transition matrix.

Transition Path Sampling with a quantum annealer. Designing sampling algorithms exploiting quantum annealers has become a highly active research field[41–45]. We leverage on this development by integrating sampling via a quantum annealer into our classical–hybrid scheme, in order to generate realistic ensembles of coarse-grained transition pathways $\mathbf{I}$. To this end, we need to sample from a path distribution $\propto e^{-S(\mathbf{I})}$. In principle, conventional stochastic algorithms on a classical computer could serve this purpose. However, their computational cost grows very rapidly with the number of configurations in the data set $\mathscr{C}$. Ultimately, classical Markov chain Monte Carlo path sampling algorithms are typically limited by long auto-correlation times in the chain. As we will show below, quantum computers can overcome this limitation: in our approach each Monte Carlo step performed on D-Wave can generate a new uncorrelated transition path. Note that we do not require a fully fair sampling of the space of possible paths, which is one of the challenges in quantum-annealer based sampling[46–48]. Employing a suitable reweighting procedure, it is sufficient for our algorithm if the exploration of the accessible space is sufficiently broad.

Quantum encoding of the transition path sampling problem The first step to derive our path sampling algorithm consists in introducing a graph representation of the path probability density defined in Eq. (1). We identify each configuration in the data set $\mathscr{C}$ with a node in the graph and define the topology of the graph so to ensure that connected neighboring nodes represent configurations that are both structurally and kinetically close (in Section S3 of the SI, we provide further details on how we enforce this condition in the application to alanine dipeptide). The weights w_{ij} of the edges in the graph are defined according to Eq. (4), thus ensuring that the sum of the weights along a given path $\mathbf{I}$ on the graph yields the path functional $S(\mathbf{I})$ entering Eq. (1).

The undirected graph representation enables us to map the sampling problem to a quantum annealing one. To this end, we introduce two sets of binary variables, $\Gamma_i^{(1)}$ and $\Gamma_{ij}^{(2)}$, where i and j run over the ν vertexes in the graph. If $\Gamma_i^{(1)} = 1$ ($\Gamma_i^{(1)} = 0$), then the i-th node is (is not) visited by the transition path on the graph (see red line in Fig. 2(b)). $\Gamma_{ij}^{(2)}$ is always 0 if the i and j are not adjacent in the graph. If i and j are adjacent, then $\Gamma_{ij}^{(2)} = 1$ when the path contains the $i \rightarrow j$ or $j \rightarrow i$ transition. We are specifically interested in configurations of the binary variables in which the set of non-vanishing entries of $\Gamma_i^{(1)}$ and $\Gamma_{ij}^{(2)}$ form a topologically connected path, i.e., a continuous line starting from the given initial node and terminating in the chosen final node.

To sample path configurations according to $e^{-S(\mathbf{I})}$, let us consider the following classical Hamiltonian of the binary variables:

$$H = \alpha H_{\text{C}} + H_{\text{T}}. \tag{5}$$

H_{C} is the constraint Hamiltonian, a positive-definite function that is zero only if the entries of the binary variables satisfy the path topology, $H_{\text{C}}(\Gamma^{(1)}, \Gamma^{(2)}) = 0$. This condition can be fulfilled by choosing[49]

$$H_C = H_s + H_t + H_r, \tag{6}$$

where

$$H_s = -\left(\Gamma_s^{(1)}\right)^2 + \left(\Gamma_s^{(1)} - \sum_i \Gamma_{si}^{(2)}\right)^2, \tag{7}$$

$$H_t = -\left(\Gamma_t^{(1)}\right)^2 + \left(\Gamma_t^{(1)} - \sum_i \Gamma_{ti}^{(2)}\right)^2, \tag{8}$$

$$H_r = \sum_{j \neq s,t} \left(2\Gamma_j^{(1)} - \sum_i \Gamma_{ji}^{(2)}\right)^2. \tag{9}$$

In this formulation, H_s and H_t introduce the condition that the path should start from the initial node s and end in the final node t, while H_r imposes the flux conservation at the remaining nodes.

$H_T = \sum_{ij} w_{ij} \Gamma_{ij}^{(2)}$ is the so-called target function. By definition, H_T yields the path action S whenever the configuration of the tensors $\Gamma^{(1)}$ and $\Gamma^{(2)}$ satisfy a path topology, that is, if $H_C(\Gamma^{(1)}, \Gamma^{(2)}) = 0$ and $\mathbf{I} = \mathbf{I}(\Gamma^{(1)}, \Gamma^{(2)})$ is the corresponding path, then $H_T(\Gamma^{(2)}) = S(\mathbf{I})$. The parameter α in Eq. (5) controls the relative strength of the constraint Hamiltonian, H_C. For $\alpha \gg 1$, all binary variables' configurations that violate the path topology correspond to very high excitations and are thus excluded from the sampling in the low-energy states that the quantum annealer performs.

Now, we are finally in a condition to tackle the problem of how to use a quantum annealer to sample path configurations with probability distribution $\propto e^{-S(\mathbf{I})}$. First, we use this machine to generate path-like binary variables' configurations according to a probability distribution that has a finite overlap with $e^{-S(\mathbf{I})}$. Then, a classical machine accepts or rejects the proposal, thus restoring the correct path probability distribution $e^{-S(\mathbf{I})}$ (see Fig. 1).

To implement this scheme, we switch to a generalized Ising Model formulation of our classical Hamiltonian, by means of a change of variables: $\sigma_i^z = 2\Gamma_i^{(1)} - 1$, $\sigma_{ij}^z = 2\Gamma_{ij}^{(2)} - 1$. Then, we promote the classical Eq. (5) to a quantum mechanical Hamiltonian $\hat{H}$, by substituting the classical Ising variables with Pauli z operators of a spin 1/2 algebra. Finally, the spin 1/2 states are encoded in the qubits of D-Wave.

In a standard quantum annealing process, the qubits are initialized in the ground-state of an easily solvable Hamiltonian[33–37], in our case

$$\hat{H}_{in} = -h_x \left(\sum_i \hat{\sigma}_i^x - \sum_{ij} \hat{\sigma}_{ij}^x \right), \tag{10}$$

where h_x is an arbitrary real constant. Then, the system is subjected to a time-dependent Hamiltonian

$$\hat{H}(t) = A(t)\,\hat{H}_{in} + B(t)\,\hat{H}, \tag{11}$$

with scheduling functions $A(t)$ and $B(t)$. These are chosen such that initially $A(0) = 1$ and $B(0) = 0$, while at the end of the protocol, i.e., at $t = t_{sweep}$, one has $A(t_{sweep}) = 0$ and $B(t_{sweep}) = 1$. That is, the sweep starts with $H(0) = H_{in}$ and ends in $H(t_{sweep}) = H$.

The spectrum of the quantum Hamiltonian $\hat{H}$ comprises all possible energy states of the classical Hamiltonian H defined in Eq. (5). Therefore, as long as $\alpha \gg 1$, the low-lying eigenstates of $\hat{H}$ represent path-like configurations $\mathbf{I}$ and their eigenvalues coincide with the path action $S(\mathbf{I})$. For a closed system, the adiabatic theorem implies that if the sweep is performed sufficiently slowly as compared to the minimal energy gap ΔE, i.e., for $t_{sweep} \gg \hbar/\Delta E$, then the system remains in its instantaneous ground-state, thus reaching the lowest energy solution at the end of the sweeping process. In this ideal condition, the annealing process would systematically return the least action path $\bar{\mathbf{I}}$[25]. Since the path probability in Langevin dynamics is given by $\sim e^{-S(\mathbf{I})}$, the least action path $\bar{\mathbf{I}}$ corresponds to the most probable transition path.

In realistic conditions, the probability of landing onto the ground-state remains < 1, even in the limit of very long sweeping times. Fluctuations in the results can be due to a combination of different factors, including the thermal coupling of the machine to its environment and other dissipation effects, non-adiabatic corrections in the sweeping procedure, or even specific bias introduced by the hardware layout, such as, e.g., those inherent with the limited topological connectivity of the quantum annealing machine. Achieving unbiased results in the presence of such errors in realistic machines is an active area of research[46–48,50]. In our work, we chose a different approach and used the machine itself to compute the probability distribution of intercepting specific solutions in the low-lying sector of the spectrum, using a cumulant expansion approximation. Then, we have corrected for such distribution by introducing a reweighing term in the acceptance/rejection formula. We have added a paragraph after to clarify this point.

This is because the coupling of the quantum annealing device to its environment induces decoherence and thermal relaxation[51]. It has been suggested that this coupling can be exploited to sample classical Boltzmann distributions[52,53]. However, in practice, the sampling can only be performed at some rescaled temperature that is very difficult to estimate *a priori*[53]. The reason is that, if the coupling $A(t)$ of the initial Hamiltonian H_{in} decays sufficiently fast, the thermal relaxation time may grow longer than the sweeping time t_{sweep}, and the relaxation

process freezes at some time $t_f < t_{\text{sweep}}$. In this case, the distribution of final energy states would be close to a modified Boltzmann distribution $e^{-B(t_f)S(\mathbf{I})}$, where t_f is the freezing time. It should be emphasized, however, that the existing quantum annealing machines such as D-Wave, are very often employed in hybrid optimization schemes that combine classical and quantum annealing. In this case, we do not expect the path probability should correspond to a Boltzmann distribution. In addition to fluctuations due to coupling to an environment, we exploit quantum fluctuations in the final state due to non-adiabaticity generated by relatively rapid sweeps. In the following, we only assume that there exists a regime of sweeping times for which the distribution of the paths generated by multiple hybrid energy minimization has a finite overlap with $e^{-S(\mathbf{I})}$.

A classical computer controlling a Metropolis scheme can exploit this overlap to yield the correct sampling of $e^{-S(\mathbf{I})}$ (Fig. 1). In general, this can be achieved by imposing the detailed balance condition $e^{-S(\mathbf{I})} T(\mathbf{I}'|\mathbf{I}) = e^{-S(\mathbf{I}')} T(\mathbf{I}|\mathbf{I}')$, where $T(\mathbf{I}'|\mathbf{I})$ is the transition probability from the path $\mathbf{I}$ to the path $\mathbf{I}'$ in the underlying stochastic process. We choose to generalize this dynamics to enable also the sweeping time t_{sweep} to vary along the Markov chain. We do so to ensure that D-Wave is mostly performing sweeps with a duration $t_{\text{sweep}} \sim t_0$, where t_0 is a tunable parameter representing a reasonable compromise between accuracy (slow sweeping) and efficiency (low consumption of quantum computing time). Upon enlarging the configuration space of the Monte Carlo dynamics to include t_{sweep}, the new detailed balance condition reads $\rho(\mathbf{I}, t_{\text{sweep}}) T(t'_{\text{sweep}}, \mathbf{I}'|t_{\text{sweep}}, \mathbf{I}) = \rho(\mathbf{I}', t'_{\text{sweep}}) T(t_{\text{sweep}}, \mathbf{I}|t'_{\text{sweep}}, \mathbf{I}')$, where $\rho(\mathbf{I}, t_{\text{sweep}})$ is the new equilibrium distribution. Our Monte Carlo dynamics must be defined in such a way to ensure that the equilibrium distribution is

$$\rho(t_{\text{sweep}}, \mathbf{I}) = p_0(t_{\text{sweep}}) \times e^{-S(\mathbf{I})}, \tag{12}$$

where $p_0(t_{\text{sweep}})$ is some arbitrary equilibrium distribution of the sweeping time, centered around t_0. Following the standard procedure to obtain the Metropolis acceptance/rejection criterium, we write the transition probability as a product of a trial move probability $\tau(\mathbf{I}', t'_{\text{sweep}}|\mathbf{I}, t_{\text{sweep}})$ and a corresponding acceptance probability $a(\mathbf{I}', t'_{\text{sweep}}|\mathbf{I}, t_{\text{sweep}})$. Since the sweeping time is allowed to vary along the chain, we factorize the trial move probability as

$$\tau(\mathbf{I}', t'_{\text{sweep}}|\mathbf{I}, t_{\text{sweep}}) = P(t'_{\text{sweep}}|t_{\text{sweep}}) P(\mathbf{I}'|t'_{\text{sweep}}), \tag{13}$$

where $P(t'_{\text{sweep}}|t_{\text{sweep}})$ is the probability for the sweeping time to go from t_{sweep} to t'_{sweep} in a Monte Carlo step, while $P(\mathbf{I}|t_{\text{sweep}})$ is the probability that a quantum annealing calculation lasting a time t_{sweep} yields the path $\mathbf{I}$. Combining all terms together, we obtain the following Metropolis acceptance rule:

$$\min\left[1, \frac{p_0(t'_{\text{sweep}}) P(t_{\text{sweep}}|t'_{\text{sweep}})}{p_0(t_{\text{sweep}}) P(t'_{\text{sweep}}|t_{\text{sweep}})} \frac{P(\mathbf{I}|t_{\text{sweep}})}{P(\mathbf{I}'|t'_{\text{sweep}})} \frac{e^{-S(\mathbf{I}')}}{e^{-S(\mathbf{I})}}\right] \tag{14}$$

In particular, in our simulations we chose to update t_{sweep} according to a Brownian dynamics with a harmonic drift term:

$$t_{\text{sweep}}^{i+1} = t_{\text{sweep}}^i - \delta t k(t_{\text{sweep}} - t_0) + \sqrt{2\delta t}\xi^i, \tag{15}$$

where ξ_i is a Gaussian distributed random variable of null mean and unitary variance and δt is an incremental sweeping time change.

The conditional probability $P(\mathbf{I}|t_{\text{sweep}})$ in Eq. (14) depends on the details of the quantum annealing machine and of the specific optimization algorithm. In general, computing $P(\mathbf{I}|t_{\text{sweep}})$ from a theoretical model of the annealing process can be very challenging. We overcome this problem and show how to estimate $P(\mathbf{I}|t_{\text{sweep}})$ by performing a moderate number of annealing processes, for each value of t_{sweep}. The spectrum of the target quantum Hamiltonian $\hat{H}$ is expected to be non-degenerate, since the weights in the graph w_{ij} are in general all different. In addition, for large values of the parameter α in Eq. (5), all low-lying states satisfy the constraints set by H_C, and thus correspond to path-like configurations $\mathbf{I}$. Therefore, each low-lying eigenvalue E of the quantum Hamiltonian $\hat{H}$ corresponds to the action of a single path, $E = S(\mathbf{I})$. Then, $P(\mathbf{I}|t_{\text{sweep}})$ can be directly inferred from a frequency histogram of the energies E obtained at the end of multiple annealing processes performed at fixed t_{sweep}, i.e., $P(\mathbf{I}|t_{\text{sweep}}) = P(E|t_{\text{sweep}})$. To minimize the consumption of quantum computing time, we can estimate $P(E|t_{\text{sweep}})$ by the lowest-order cumulant expansion as

$$P(\mathbf{I}|t_{\text{sweep}}) \simeq P(E|t_{\text{sweep}}) \simeq \frac{1}{\sqrt{2\pi}\Delta} e^{-\frac{(E(\mathbf{I})-\overline{E})^2}{2\Delta^2}}, \tag{16}$$

where $\overline{E}$ and Δ, respectively, are the average and the standard deviation of the energy obtained by many annealing processes at fixed sweeping time t_{sweep}. In principle, the estimated distribution can be improved systematically by including higher orders in the cumulant expansion, which may in particular become important in the presence of long tails. For our benchmark purposes, we find (16) to be sufficient.

Application to a molecular benchmark system

To illustrate our hybrid classical/quantum Monte Carlo scheme sketched in Fig. 1, we apply it to simulate the $C5 \rightarrow \alpha_R$ transition of alanine dipeptide. We have chosen this standard benchmark system, as it is sufficiently small to enable us to carry out the quantum computing calculations on existing D-Wave machines. This serves to illustrate all the relevant steps, which—thanks to the generality of our approach—can be applied to molecular systems of increased complexity in step with increasing qubit numbers in future generations of quantum annealers.
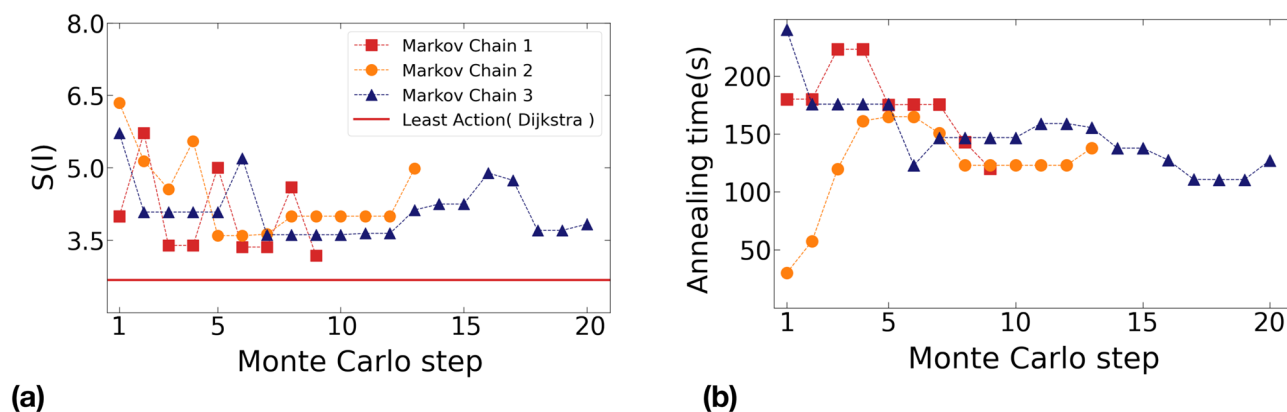
Figure 3. (a) Evolution of the path action $S(\mathbf{I})$ and (b) annealing time t_{sweep} along the Monte Carlo paths generated using the hybrid classical/quantum annealing implemented on D-Wave.

First, we use iMapD and our effective Langevin theory to construct the graph representation of the dynamics on the intrinsic manifold of this peptide (details on the implementation of iMapD and the calculation of the weights w_{ij} for this molecular system are provided in section and S5 of the SI). The results are shown in the Ramachandran plot reported in Fig. S2–S4 of the SI. The contour lines in the background represent the free energy surface, calculated from a frequency histogram of 1 μs of equilibrium MD at $T = 300K$, generated using OpenMM[54], in the AMBER99SB force field with explicit TIP3P water[55]. The spatial resolution of our effective theory is determined by the number of configurations v we keep to generate a sparse graph. With this choice, the average RMSD distance between neighbouring configurations in our network is $\delta_{RMSD} \simeq 0.5\,\text{Å}$. Then, $\sigma \simeq \delta_{RMSD}\sqrt{N_a}$, where $N_a = 22$ is the number of atoms in our molecule.

To implement our hybrid classical/quantum Monte Carlo scheme, we encode the quantum Hamiltonian $\hat{H}$ defined by the graph using the Ocean suite, operating on the D-Wave quantum annealer. Encoding our system on D-Wave requires 578 qbits, given by the sum of the number of nodes and edges of our network. To generate trial paths, we rely on the hybrid solver available on Leap, which combines quantum annealing with classical simulating annealing. In this case, t_{sweep} is identified with the total quantum and classical computing time employed by the solver. We estimate the resulting conditional probability $P(\mathbf{I}|t_{sweep})$ entering Eq. (14) by means of a direct calculation on D-Wave (Table S1 in SI), using Eq. (16). In Fig. S8 in the SI, we report the average value of the energy $\bar{E}$ and its standard deviation Δ, entering Eq. (16).

We initiated three independent Markov chains from arbitrary paths generated by a quantum annealing process at $t_{sweep} = 180\,s$, $30\,s$, and $240\,s$, corresponding to about $8.6\,s$, $1.4\,s$, and $11.4\,s$ of quantum annealing time, respectively (details on how we determine the initial and final nodes are outlined in Section S3 of the SI). We evolved t_{sweep} according to Eq. (15) with $k = 2 \times 10^{-4}\,\text{s}^{-1}$ and $t_0 = 150s$ and then accepted or rejected the new paths according to Eq. (14).

We recall that the α parameter determining the relative strength of the constraint and target Hamiltonian must be chosen in order to balance between two general requirements: (i) it needs to be as large as necessary to energetically separate configurations with false topology reliably from the correct low-energy manifold and (ii) it needs to be as small as possible, since the maximal achievable energy scale is limited by the hardware; thus, a larger value of α would in fact correspond to a decrease of all the other energy scales and hence to a worse energetic resolution of the configurations with correct topology. In this application, we heuristically set $\alpha = \sum_{ij} w_{ij}$, which is compliant with both the aforementioned requests. With this choice, on average, over 60% of the annealing sweeps led to configurations of binary variables $\Gamma^{(1)}$ and $\Gamma^{(2)}$ with a correct path topology (Table S2 in SI), thus providing viable trial transition paths.

In Fig. 3 we show the change in path action S (left panel) and the hybrid minimization time t_{sweep} (right panel), along our three Markov chains. As these results show, the Monte Carlo algorithm occasionally accepts trial moves with a higher action. They also show that longer annealing times do not always yield paths with lower actions. This is expected, since the $P(E|t_{sweep})$ distributions have significant overlap, as it can be inferred from Fig. S8 in the SI.

The transition paths generated by our scheme are consistent with the free energy landscape produced by equilibrium MD. Figure 4a shows the first and last accepted transition paths of one of the generated Markov chains. Both paths correctly connect the two meta-stable states, navigate the low-free energy regions of the surface, and cross the barrier at its lowest point.

The transition paths explore a region around the most probable path[56–58], which in Fig. 4b is shown based on calculations on a classical computer using the Dijkstra algorithm[59]. Figure 4b also reports how often the sampled transition paths pass the nodes of the network, i.e., the statistical weight of the corresponding configuration in the transition path ensemble. All transition paths go through the transition state. However, due to the presence of fluctuations, a finite probability is obtained also at configurations with relatively high free energy. The deterministic Dijkstra algorithm can only detect the global minimum of the functional $S(\mathbf{I})$. In contrast, our TPS algorithm accounts for fluctuations that lead to the full transition path ensemble.

The main strength of our hybrid classical/quantum scheme is that it allows us to efficiently obtain independent transition paths. At each adiabatic cycle, the quantum machine is reset to a quantum state in which all the spins
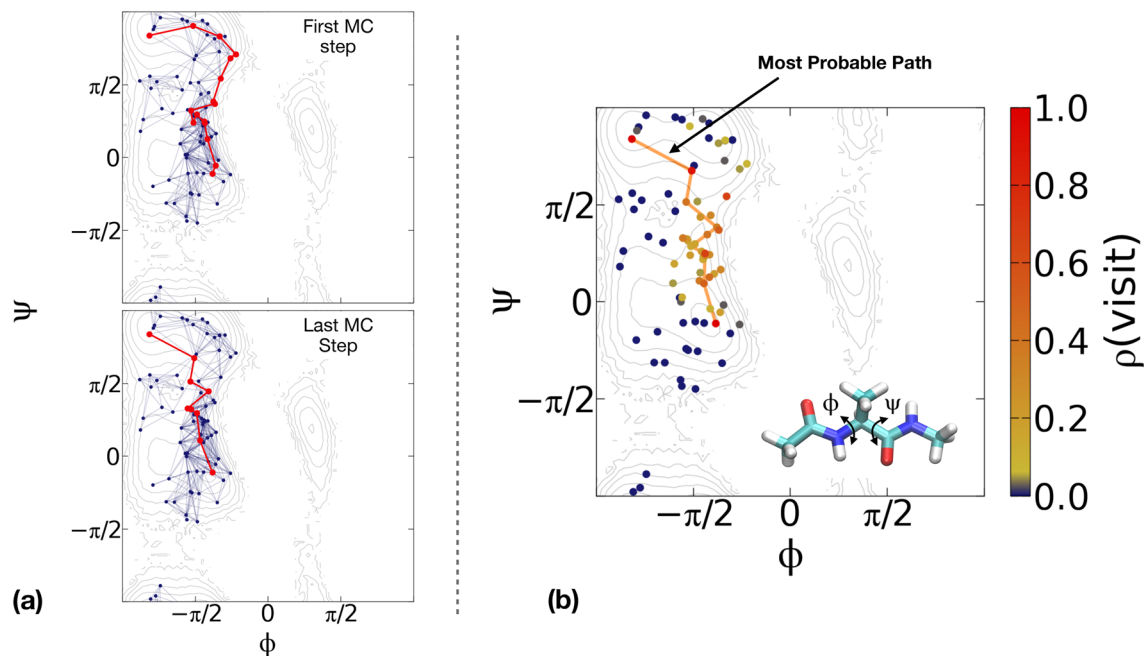
Figure 4. (**a**) Transition pathways for the $C5 \rightarrow \alpha_R$ transition of alanine dipeptide obtained from our Monte Carlo scheme. The red line denotes the first (top) and last (bottom) trajectory in a Markov chain. The points in the Ramachandran plots are obtained from projecting the configurations generated with iMapD. In the background is the free energy surface calculated from 1 μs of plain MD. (**b**) Transition path density on the Ramachandran plane, evaluated for the ensemble of trajectories calculated with our Monte Carlo scheme. The solid orange line is the most probable path, obtained using the Dijkstra algorithm[59] on a classical computer.
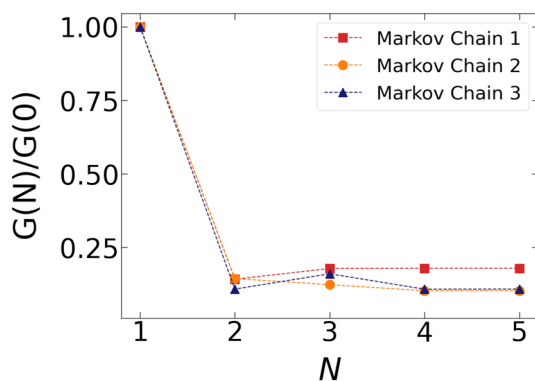
Figure 5. The ratio of auto-correlation function $G(N)/G(0)$ (see Eq. (S29) in SI) plotted as a function of Monte Carlo steps N for three independent Markov chains.

are aligned along the $\hat{x}$ direction. Consequently, each qubit has initially 50% probability to be in configuration $+1/2$ and $-1/2$. Since the only source of correlation is introduced by the Markovian stochastic evolution of the minimization time t_{sweep} (Eq. (15)), we expect each minimization procedure to yield a completely independent path (one-shot generation). To quantify the degree of correlation in the ensemble of trajectories sampled in three Markov chains, we consider the auto-correlation function $G(N)$ defined in the Section S6 of the SI, where N denotes distance in the Monte Carlo chain. In Fig. 5 we plot the behaviour of $G(N)$ (evaluated relatively to its initial value $G(0)$) for each independent Markov chain (see also discussion in Section S6 of the SI). These results clearly indicate that the correlation of the generated trajectories is suppressed after just a single Monte Carlo step.

Conclusions

In this work, we have established a novel computational framework to sample the transition path ensemble of molecular conformational transitions, which integrates a ML driven exploration with a hybrid Monte Carlo scheme that exploits the potential of QC. We have used the iMapD algorithm[32] to achieve an uncharted exploration of the molecular intrinsic manifold, without introducing any choice of CV, nor biasing force. These data enabled us to build a coarse-grained representation of the dynamics directly on the intrinsic manifold. To

construct this low-resolution theory, we adapted regularization and renormalization methods that were originally developed within the context of high energy physics[40], and which may also be useful in other applications in soft-condensed matter and biophysics[60,61]. We then encoded the path sampling problem in a form that enabled us to use a D-Wave quantum annealer to generate uncorrelated trial transition paths, thus enhancing the exploration of the transition path ensemble. Finally, by using the Metropolis criterion in Eq. (14), we made sure to account for each trajectory with its correct statistical weight in the transition path ensemble.

The algorithm we presented here is designed to sample the full transition path ensemble. This achievement represents a significant advancement with respect to previous attempts to compute the most probable transition paths on a quantum computer[25]. The transition path ensemble is often heterogeneous, displaying several alternative transition channels, corresponding to alternative molecular mechanisms. Even though in the proof-of-concept we discussed here we restricted our sampling to 23 transition paths, in general the number of trajectories is only limited by the available computational resources.

While significant effort has been made towards designing quantum algorithms for quantum many-body problems[19–24,62–66], only a few applications of quantum computing to classical molecular sampling problems have been reported to date[25,26,67–69]. Most of these attempts assume a simplified molecular representation, among which lattice discretization[26,67,68]. Unlike the method developed in[25], which was designed to return only the most probable path, here we sample the full transition path ensemble. In addition, we do not introduce any unphysical biasing force nor a choice of CVs to accelerate the exploration of configuration space. Finally, to the best of our knowledge, the present calculation represents the first successful application of a quantum computing machine to characterize a molecular transition using a state-of-the-art atomistic force field.

With the present quantum encoding, the size of the molecular systems that can currently be investigated is limited by the relatively small number of qubits that are available on the existing quantum annealing machines. The characterization of transitions with a comparable level of spatio-temporal resolution of much larger molecules (for example, the folding of a small protein) typically requires to generate at least $10^3 - 10^4$ points on the intrinsic manifold[25]. In this case, implementing our scheme on a quantum computer would require a number of qubits more than one order of magnitude larger than that of the most powerful existing quantum annealing device. However, if the size of quantum computing hardware continues to grow in size and performance according to the present exponential rate[27–29], we may hope this threshold to be reached within the foreseeable future. Due to the suppression of autocorrelation time, we expect that, for sufficiently large networks, our hybrid scheme may ultimately have an edge over classical stochastic methods.

In the future, heterogeneous platforms for high-performance computing might emerge that integrate CPUs and GPUs with quantum and machine learning cores. These new machines will require scientific software able to operate across the different parts, fully taking advantage of their strengths. Heterogeneous algorithms like the one we presented here will thus become increasingly important, with great potential for the computational molecular sciences.

Data availability
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

References
1. Dror, R. O., Dirks, R. M., Grossman, J., Xu, H. & Shaw, D. E. Biomolecular simulation: A computational microscope for molecular biology. *Annu. Rev. Biophys.* **41**, 429–452. https://doi.org/10.1146/annurev-biophys-042910-155245 (2012).
2. Peters, B. *Reaction Rate Theory and Rare Events* (Elsevier, 2017).
3. Yang, Y. I., Shao, Q., Zhang, J., Yang, L. & Gao, Y. Q. Enhanced sampling in molecular dynamics. *J. Chem. Phys.* **151**, 070902. https://doi.org/10.1063/1.5109531 (2019).
4. Laio, A. & Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci.* **99**, 12562–12566. https://doi.org/10.1073/pnas.202427399 (2002).
5. Cuendet, M. A. & Tuckerman, M. E. Free energy reconstruction from metadynamics or adiabatic free energy dynamics simulations. *J. Chem. Theory Comput.* **10**, 2975–2986. https://doi.org/10.1021/ct500012b (2014).
6. Maragliano, L. & Vanden-Eijnden, E. A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations. *Chem. Phys. Lett.* **426**, 168–175. https://doi.org/10.1016/j.cplett.2006.05.062 (2006).
7. a Beccara, S., Fant, L. & Faccioli, P. Variational scheme to compute protein reaction pathways using atomistic force fields with explicit solvent. *Phys. Rev. Lett.* **114**, https://doi.org/10.1103/physrevlett.114.098103 (2015).
8. Peters, B. Reaction coordinates and mechanistic hypothesis tests. *Annu. Rev. Phys. Chem.* **67**, 669–690. https://doi.org/10.1146/annurev-physchem-040215-112215 (2016).
9. Dellago, C., Bolhuis, P. G. & Chandler, D. On the calculation of reaction rate constants in the transition path ensemble. *J. Chem. Phys.* **110**, 6617–6625. https://doi.org/10.1063/1.478569 (1999).
10. Jung, H., ichi Okazaki, K. & Hummer, G. Transition path sampling of rare events by shooting from the top. *J. Chem. Phys.* **147**, 152716. https://doi.org/10.1063/1.4997378 (2017).
11. Bolhuis, P. G. & Swenson, D. W. H. Transition path sampling as markov chain monte carlo of trajectories: Recent algorithms, software, applications, and future outlook. *Adv. Theory Simul.* **4**, 2000237. https://doi.org/10.1002/adts.202000237 (2021).
12. Jung, H., Covino, R., Arjun, A., Bolhuis, P. G. & Hummer, G. Autonomous artificial intelligence discovers mechanisms of molecular self-organization in virtual experiments (2021).
13. Noé, F., Olsson, S., Köhler, J. & Wu, H. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science* **365**, 10.1126/science.aaw1147 (2019).
14. Bonati, L., Zhang, Y.-Y. & Parrinello, M. Neural networks-based variationally enhanced sampling. *Proc. Natl. Acad. Sci.* **116**, 17641–17647. https://doi.org/10.1073/pnas.1907975116 (2019).
15. Schneider, E., Dai, L., Topper, R. Q., Drechsel-Grau, C. & Tuckerman, M. E. Stochastic neural network approach for learning high-dimensional free energy surfaces. *Phys. Rev. Lett.* **119**, https://doi.org/10.1103/physrevlett.119.150601 (2017).

16. Wang, Y., Ribeiro, J. M. L. & Tiwary, P. Machine learning approaches for analyzing and enhancing molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **61**, 139–145. https://doi.org/10.1016/j.sbi.2019.12.016 (2020).

17. Noé, F., Tkatchenko, A., Müller, K.-R. & Clementi, C. Machine learning for molecular simulation. *Annu. Rev. Phys. Chem.* **71**, 361–390. https://doi.org/10.1146/annurev-physchem-042018-052331 (2020).

18. Sidky, H., Chen, W. & Ferguson, A. L. Machine learning for collective variable discovery and enhanced sampling in biomolecular simulation. *Mol. Phys.* **118**, https://doi.org/10.1080/00268976.2020.1737742 (2020).

19. Hempel, C. *et al.* Quantum chemistry calculations on a trapped-ion quantum simulator. *Phys. Rev. X* **8**, https://doi.org/10.1103/physrevx.8.031022 (2018).

20. Cao, Y. *et al.* Quantum chemistry in the age of quantum computing. *Chem. Rev.* **119**, 10856–10915. https://doi.org/10.1021/acs.chemrev.8b00803 (2019).

21. Genin, S. N., Ryabinkin, I. G. & Izmaylov, A. F. *Quantum Chemistry on Quantum Annealers* (2019).

22. Outeiral, C. *et al.* The prospects of quantum computing in computational molecular biology. *WIREs Comput. Mol. Sci.* **11**, https://doi.org/10.1002/wcms.1481 (2020).

23. McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C. & Yuan, X. Quantum computational chemistry. *Rev. Modern Phys.* **92**, https://doi.org/10.1103/revmodphys.92.015003 (2020).

24. Dumitrescu, E. *et al.* Cloud quantum computing of an atomic nucleus. *Phys. Rev. Lett.* **120**, https://doi.org/10.1103/physrevlett.120.210501 (2018).

25. Hauke, P., Mattiotti, G. & Faccioli, P. Dominant reaction pathways by quantum computing. *Phys. Rev. Lett.* **126**, https://doi.org/10.1103/physrevlett.126.028104 (2021).

26. Micheletti, C., Hauke, P. & Faccioli, P. Polymer physics by quantum computing. *Phys. Rev. Lett.* **127**, https://doi.org/10.1103/physrevlett.127.080501 (2021).

27. Scholl, P. *et al.* Quantum simulation of 2d antiferromagnets with hundreds of rydberg atoms. *Nature* **595**, 233–238. https://doi.org/10.1038/s41586-021-03585-1 (2021).

28. Pogorelov, I. *et al.* Compact ion-trap quantum computing demonstrator. *PRX Quant.* **2**, https://doi.org/10.1103/prxquantum.2.020343 (2021).

29. Ball, P. First quantum computer to pack 100 qubits enters crowded race. *Nature* **599**, 542–542. https://doi.org/10.1038/d41586-021-03476-5 (2021).

30. Arute, F. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510. https://doi.org/10.1038/s41586-019-1666-5 (2019).

31. Zhong, H.-S. *et al.* Quantum computational advantage using photons. *Science* **370**, 1460–1463. https://doi.org/10.1126/science.abe8770 (2020).

32. Chiavazzo, E. *et al.* Intrinsic map dynamics exploration for uncharted effective free-energy landscapes. *Proc. Natl. Acad. Sci.* **114**, E5494–E5503. https://doi.org/10.1073/pnas.1621481114 (2017).

33. Das, A. & Chakrabarti, B. K. (eds.) *Quantum Annealing and Related Optimization Methods*. Lecture Notes in Physics (Springer, Berlin, Germany, 2005), 2005 edn.

34. Das, A. & Chakrabarti, B. K. Colloquium: Quantum annealing and analog quantum computation. *Rev. Modern Phys.* **80**, 1061–1081. https://doi.org/10.1103/revmodphys.80.1061 (2008).

35. Albash, T. & Lidar, D. A. Adiabatic quantum computation. *Rev. Modern Phys.* **90**, https://doi.org/10.1103/revmodphys.90.015002 (2018).

36. Venegas-Andraca, S. E., Cruz-Santos, W., McGeoch, C. & Lanzagorta, M. A cross-disciplinary introduction to quantum annealing-based algorithms. *Contemp. Phys.* **59**, 174–197. https://doi.org/10.1080/00107514.2018.1450720 (2018).

37. Hauke, P., Katzgraber, H. G., Lechner, W., Nishimori, H. & Oliver, W. D. Perspectives of quantum annealing: Methods and implementations. *Rep. Progress Phys.* **83**, 054401. https://doi.org/10.1088/1361-6633/ab85b8 (2020).

38. Inc., D.-W. S. Leap webpage (2022).

39. Coifman, R. R. & Lafon, S. Diffusion maps. *Appl. Comput. Harmonic Anal.* **21**, 5–30. https://doi.org/10.1016/j.acha.2006.04.006 (2006).

40. Lepage, P. How to renormalize the schrodinger equation (1997).

41. Adachi, S. H. & Henderson, M. P. Application of quantum annealing to training of deep neural networks (2015).

42. Chancellor, N., Szoke, S., Vinci, W., Aeppli, G. & Warburton, P. A. Maximum-entropy inference with a programmable annealer. *Sci. Rep.* **6**, https://doi.org/10.1038/srep22318 (2016).

43. Benedetti, M., Realpe-Gómez, J., Biswas, R. & Perdomo-Ortiz, A. Quantum-assisted learning of hardware-embedded probabilistic graphical models. *Phys. Rev. X* **7**, https://doi.org/10.1103/physrevx.7.041052 (2017).

44. Winci, W. *et al.* A path towards quantum advantage in training deep generative models with quantum annealers. *Mach. Learn. Sci. Technol.* **1**, 045028. https://doi.org/10.1088/2632-2153/aba220 (2020).

45. Sieberer, L. M. & Lechner, W. Programmable superpositions of ising configurations. *Phys. Rev. A* **97**, https://doi.org/10.1103/physreva.97.052329 (2018).

46. Könz, M. S., Mazzola, G., Ochoa, A. J., Katzgraber, H. G. & Troyer, M. Uncertain fate of fair sampling in quantum annealing. *Phys. Rev. A* **100**, https://doi.org/10.1103/physreva.100.030303 (2019).

47. Yamamoto, M., Ohzeki, M. & Tanaka, K. Fair sampling by simulated annealing on quantum annealer. *J. Phys. Soc. Japan* **89**, 025002. https://doi.org/10.7566/jpsj.89.025002 (2020).

48. Kumar, V., Tomlin, C., Nehrkorn, C., O'Malley, D. & au2, J. D. I. Achieving fair sampling in quantum annealing (2020).

49. Krauss, T. & McCollum, J. Solving the network shortest path problem on a quantum annealer. *IEEE Trans. Quant. Eng.* **1**, 1–12. https://doi.org/10.1109/tqe.2020.3021921 (2020).

50. Golden, J., Baertschi, A., O'Malley, D. & Eidenbenz, S. Fair sampling error analysis on nisq devices. *ACM Trans. Quant. Comput.* **3**, 1–23. https://doi.org/10.1145/3510857 (2022).

51. Amin, M. H. S., Truncik, C. J. S. & Averin, D. V. Role of single-qubit decoherence time in adiabatic quantum computation. *Phys. Rev. A* **80**, https://doi.org/10.1103/physreva.80.022303 (2009).

52. Amin, M. H. Searching for quantum speedup in quasistatic quantum annealers. *Phys. Rev. A* **92**, https://doi.org/10.1103/physreva.92.052323 (2015).

53. Benedetti, M., Realpe-Gómez, J., Biswas, R. & Perdomo-Ortiz, A. Estimation of effective temperatures in quantum annealers for sampling applications: A case study with possible applications in deep learning. *Phys. Rev. A* **94**, https://doi.org/10.1103/physreva.94.022308 (2016).

54. Eastman, P. *et al.* OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Comput. Biol.* **13**, e1005659. https://doi.org/10.1371/journal.pcbi.1005659 (2017).

55. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins Struct. Funct. Bioinform.* **78**, 1950–1958. https://doi.org/10.1002/prot.22711 (2010).

56. Faccioli, P., Sega, M., Pederiva, F. & Orland, H. Dominant pathways in protein folding. *Phys. Rev. Lett.* **97**, https://doi.org/10.1103/physrevlett.97.108101 (2006).

57. Sega, M., Faccioli, P., Pederiva, F., Garberoglio, G. & Orland, H. Quantitative protein dynamics from dominant folding pathways. *Phys. Rev. Lett.* **99**, https://doi.org/10.1103/physrevlett.99.118102 (2007).

58. Elber, R. & Shalloway, D. Temperature dependent reaction coordinates. *J. Chem. Phys.* **112**, 5539–5545. https://doi.org/10.1063/1.481131 (2000).
59. Sniedovich, M. Dijkstra's algorithm revisited: the dynamic programming connexion. *Control Cybern.* **35** (2006).
60. Corradini, O., Faccioli, P. & Orland, H. Simulating stochastic dynamics using large time steps. *Phys. Rev. E* **80**, https://doi.org/10.1103/physreve.80.061112 (2009).
61. Orioli, S. & Faccioli, P. Dimensional reduction of markov state models from renormalization group theory. *J. Chem. Phys.* **145**, 124120. https://doi.org/10.1063/1.4963196 (2016).
62. Zohar, E., Cirac, J. I. & Reznik, B. Quantum simulations of lattice gauge theories using ultracold atoms in optical lattices. *Rep. Progress Phys.* **79**, 014401. https://doi.org/10.1088/0034-4885/79/1/014401 (2015).
63. Dalmonte, M. & Montangero, S. Lattice gauge theory simulations in the quantum information era. *Contemp. Phys.* **57**, 388–412. https://doi.org/10.1080/00107514.2016.1151199 (2016).
64. Bañuls, M. C. *et al.* Simulating lattice gauge theories within quantum technologies. *Eur. Phys. J. D* **74**, https://doi.org/10.1140/epjd/e2020-100571-8 (2020).
65. Hauke, P., Cucchietti, F. M., Tagliacozzo, L., Deutsch, I. & Lewenstein, M. Can one trust quantum simulators?. *Rep. Progress Phys.* **75**, 082401. https://doi.org/10.1088/0034-4885/75/8/082401 (2012).
66. Cirac, J. I. & Zoller, P. Goals and opportunities in quantum simulation. *Nat. Phys.* **8**, 264–266. https://doi.org/10.1038/nphys2275 (2012).
67. Perdomo-Ortiz, A., Dickson, N., Drew-Brook, M., Rose, G. & Aspuru-Guzik, A. Finding low-energy conformations of lattice protein models by quantum annealing. *Sci. Rep.* **2**, https://doi.org/10.1038/srep00571 (2012).
68. Lu, L.-H. & Li, Y.-Q. Quantum approach to fast protein-folding time. *Chin. Phys. Lett.* **36**, 080305. https://doi.org/10.1088/0256-307x/36/8/080305 (2019).
69. Mazzola, G. Sampling, rates, and reaction currents through reverse stochastic quantization on quantum computers. *Phys. Rev. A* **104**, https://doi.org/10.1103/physreva.104.022431 (2021).

Author contributions

D.G. and P.F. performed calculations, D.G. analyzed the data, R.C. and P.F. supervised the research, all authors contributed to developing the method and writing the manuscript.

Competing interests

P.F. is co-founder of Sibylla Biotech SRL, a company exploiting molecular simulations to perform early stage drug discovery. The rest of authors declare no competing interest.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-20032-x.

Correspondence and requests for materials should be addressed to R.C. or P.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.