



Cite this: *Digital Discovery*, 2026, 5, 64

Received 20th August 2025

Accepted 13th November 2025

DOI: 10.1039/d5dd00377f

[rsc.li/digitaldiscovery](http://rsc.li/digitaldiscovery)

Simulating chemical systems is highly sought after and computationally challenging, as the number of degrees of freedom increases exponentially with the size of the system. Quantum computers have been proposed as a computational means to overcome this bottleneck, thanks to their capability of representing this amount of information efficiently. Most efforts so far have been centered around determining the ground states of chemical systems. However, hardness results and the lack of theoretical guarantees for efficient heuristics for initial-state generation shed doubt on the feasibility. Here, we propose a heuristically guided approach that is based on inherently efficient routines to solve chemical simulation problems, requiring quantum circuits of size scaling polynomially in relevant system parameters. If a set of assumptions can be satisfied, our approach finds good initial states for dynamics simulation by assembling them in a scattering tree. In particular, we investigate a scattering-based state preparation approach within the context of mergo-association. We discuss a variety of quantities of chemical interest that can be measured after the quantum simulation of a process, e.g., a reaction, following its corresponding initial state preparation.

## Chemically motivated simulation problems are efficiently solvable on a quantum computer

Philipp Schleich,<sup>1,2\*</sup> Lasse Bjørn Kristensen,<sup>1,3</sup> Jorge A. Campos-Gonzalez-Angulo,<sup>1,2</sup> Abdulrahman Aldossary,<sup>1,2</sup> Davide Avagliano,<sup>1,2</sup> Mohsen Bagherimehrab,<sup>1,2</sup> Christoph Gorgulla,<sup>1,2</sup> Joe Fitzsimons<sup>1,2</sup> and Alán Aspuru-Guzik<sup>1,2,3,4,5,6,7,8,9,10,11,12,13</sup>

### 1. Introduction

The idea of using quantum computers for the simulation of quantum systems goes back to the first proposals of quantum computing by Benioff, Feynman, and Manin.<sup>1-3</sup> This idea has generated substantial effort toward applying quantum computing to chemical problems, as conventional quantum many-body simulations for chemistry are inherently limited by the curse of dimensionality and constitute a significant portion of current supercomputing usage. This effort has largely focused on two problems: quantum simulation of dynamics (also known simply as Hamiltonian simulation or quantum simulation in the field of quantum computing) and the ground-state preparation problem. Quantum simulation<sup>4-8</sup> describes the problem of time-evolving an initial state according to the Schrödinger equation under a Hamiltonian. The relevant Hamiltonians for chemical and most physical processes can be efficiently computed, and the corresponding time evolution, governed by the Schrödinger equation, is probably within the computational complexity class BQP (bounded-error quantum polynomial-time). This class encompasses decision problems that a quantum computer can solve in polynomial time with a bounded error probability.<sup>6,9</sup> In contrast, the problem of determining the ground state – formulated as the local Hamiltonian problem in quantum complexity theory – is complete for the class Quantum Merlin-Arthur (QMA),<sup>9-11</sup> a larger class sometimes compared to the classical complexity class NP. More recently, it was shown that simulating Schrödinger operators of the form  $-\Delta + V$  with some restrictions regarding smoothness and boundedness on  $V$  and without particle exchange symmetries, a slight restriction compared to general local Hamiltonians, is also BQP-complete, and the ground-state problem for such operators is StoqMA-hard,<sup>12</sup> namely Merlin-Arthur for Hamiltonians without a sign problem (stoquastic Hamiltonians).<sup>13</sup> As a result, it is not clear whether ground state problems can be solved efficiently, even on a quantum computer.

Ground-state search and Hamiltonian simulation are often used jointly. Simulation algorithms serve as a subroutine to find

<sup>1</sup>Department of Computer Science, University of Toronto, Toronto, ON M5S 3G4, Canada. E-mail: [aspuru@utoronto.ca](mailto:aspuru@utoronto.ca); [philipps@cs.toronto.edu](mailto:philipps@cs.toronto.edu)

<sup>2</sup>Vector Institute of Artificial Intelligence, Toronto, ON M5G 1M1, Canada

<sup>3</sup>Department of Chemistry, University of Toronto, Toronto, ON M5S 3H6, Canada

<sup>4</sup>Department of Computer Science, University of Copenhagen, Copenhagen, DK-2100, Denmark. E-mail: [lakr@di.ku.dk](mailto:lakr@di.ku.dk)

<sup>5</sup>Chimie ParisTech, PSL University, CNRS, Institute for Life and Health Sciences (iCLéHS UMR 8060), 75005 Paris, France

<sup>6</sup>Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA

<sup>7</sup>Department of Physics, Harvard University, Cambridge, MA 02138, USA

<sup>8</sup>Horizon Quantum Computing, Singapore, 138565

<sup>9</sup>Department of Chemical Engineering & Applied Chemistry, University of Toronto, Toronto, ON M5S 3E5, Canada

<sup>10</sup>Department of Materials Science & Engineering, University of Toronto, Toronto, ON M5S 3E4, Canada

<sup>11</sup>Acceleration Consortium, Toronto, ON M5S 3H6, Canada

<sup>12</sup>Canadian Institute for Advanced Research, Toronto, ON M5G 1Z8, Canada

<sup>13</sup>NVIDIA, 431 King St W #6th, Toronto, ON M5V 1K4, Canada



the ground-state energy (*e.g.*, *via* phase estimation<sup>10,14–16</sup>), in the elucidation of reaction mechanisms described in ref. 17 or for free energy calculations in ref. 18, while ground states can be input states for performing quantum dynamics. State-of-the-art implementations of Hamiltonian simulation can be categorized into several classes. One is the class of algorithms based on Trotter–Suzuki formulae<sup>9,19</sup> which split the exponential of a sum into products of exponentials. Another class is based on “qubitization”,<sup>20,21</sup> which makes use of quantum signal processing<sup>22</sup> to encode the simulation in a quantum walk. Additionally, there are randomized algorithms such as QDRIFT<sup>23</sup> and its extensions.<sup>24,25</sup> Each of these approaches has its advantages in different scenarios,<sup>26,27</sup> and methods based on product formulae in particular can make the evolution of perturbative systems, as often present in chemistry and physics, even more efficient.<sup>28,29</sup> However, as they all lead to polynomially sized circuits for Hamiltonians of constant sparsity, we refrain from going into details here and refer to the relevant literature. A good overview can be found in ref. 30, with additional recent theoretical progress reported in ref. 31 and 32 and experimental progress reported in ref. 33 and 34.

We propose a departure from the current mainstream approach of the quantum computing community to chemistry, moving beyond the 1950s’ computational chemist’s way of thinking,<sup>33,36</sup> which has been shaped by the limitations of classical computing to use the ground-state as the starting point of computations, to a new era. With the advent of fault-tolerant quantum computers, dynamical simulation of quantities that a practicing chemist might care about is within reach. Specifically, we argue that a wide set of relevant quantum chemistry problems are inherently addressable through dynamical evolution alone, leading to efficient quantum algorithms for these problems. This holds, *e.g.*, for the determination of reaction rates, photochemistry, or spectroscopy. Examples of such problems and strategies for how to extract relevant quantities are summarized in Section 1.2.

A central requirement for such dynamics-based approaches is having techniques for preparing chemically relevant initial states. For this state preparation, our proposed framework employs a limited set of attainable atomic initial states and then dynamically prepares input states for a reaction of interest through a scattering process. Particularly, our main contribution is the approach to prepare molecular states by hierarchically ‘combining’ atomic ones – here specifically by scattering them. This allows for a heuristic state preparation that is guided by chemical intuition. Then, another dynamical evolution embodies a reaction, and a broad set of relevant quantities can be measured. Under certain assumptions on what constitutes relevant initial states, the ensuing algorithm is not limited by the QMA-hardness of preparing ground states and thermal states. Similarly, the focus on less restrictive initial states and more general observables means we can circumvent an orthogonality catastrophe, namely, a vanishing success probability to retrieve a ground state. As has been recently shown,<sup>37</sup> this orthogonality catastrophe is why exponential speedups for ground-state energy estimation of molecular Hamiltonians on

quantum computers may be hard to attain as the (state-of-the-art) methods considered for state preparation fail to produce reliable overlaps for molecules of increasing size. However, a distinction to make is that our methodology is based on heuristic physical intuition, which makes it not directly comparable to methods like QPE from a complexity-theoretic perspective. Furthermore, a method was proposed recently that is able to address the orthogonality catastrophe in some cases with a divide and conquer approach.<sup>38</sup> Similar to our ‘molecule factory’ in Fig. 3 where we split the problem into fragments based on chemical intuition (which can *e.g.* be guided by knowledge from retrosynthesis in certain cases), they use a tree decomposition of the Hamiltonian.

Two observations are at the foundation of our approach. First, the simulation of  $k$ -local Hamiltonians is achievable by polynomial-size quantum circuits. Hamiltonians that stem from chemical problems are 2-local due to the nature of the Coulomb interaction and, thus, also have finite locality when represented as quantum operations on qubits. Including photons to simulate light–matter interactions in the computation increases the maximum support of operations but not beyond a constant factor. The second observation is that we can aim to simulate processes corresponding to experiments that can be performed in a lab in finite time, *i.e.*, problems that are, in some sense, experimentally verifiable. Molecular ground states, as viewed from the perspective of computational chemistry, which implies a frozen molecular geometry at absolute-zero temperature, do not necessarily represent a system’s thermal equilibrium state and hence do not belong to the class of problems corresponding to experiments that can be carried out in finite time. Specifically, this encompasses situations when the Born–Oppenheimer approximation fails, and when effects such as nonadiabatic coupling and anharmonicity are not negligible, or the problem requires treatment through thermal states. We propose to simulate the process of producing reactants with a hierarchical scattering process. First, we prepare the ground states of atoms by quantum phase estimation – this does not fall under the orthogonality catastrophe, as atoms are finite-sized and state preparation feasible as long as we stay within the regime where efficient heuristics for initial state preparation are available – and then use a simple scattering process to produce molecular reactants. Our method integrates artificial potentials and photonic fields to induce the success of scatterings, leading to a lower-bounded probability of success. Specifically, mergo-association, as discussed in Section 3, is a promising avenue to realize this. This lower bound means that a fixed number of repetitions of a weak measurement scheme to herald success – see Section 3.2 – will suffice to ensure the success of intermediary scattering processes. Thus, meaningful molecular input states, which do not need to be ground states, are efficiently prepared. Our framework facilitates the modeling of complex chemical reactions by hierarchically operating the scattering with  $N$  atoms to create  $M$  reactants, which can then undergo a quantum simulation corresponding to a reaction, followed by measurements of reaction rates and time correlation functions.



### 1.1. Hamiltonian simulation and complexity considerations

We continue by discussing relevant complexity classes for the problems we consider and the efficiency of their solutions. The complexity class BQP (bounded-error quantum polynomial-time) is oftentimes considered the quantum generalization of the complexity class P (polynomial time), or, more precisely, its probabilistic extension BPP (bounded-error probabilistic polynomial-time). Polynomial complexity is usually considered efficient as the increase in cost when scaling relevant parameters is somewhat moderate. Problems from QMA (Quantum Merlin-Arthur), in contrast, can be considered “hard”; QMA is the quantum analog of NP (non-deterministic polynomial time) or, more precisely, the probabilistic class MA (Merlin-Arthur). QMA describes promise problems with solutions that can be verified in polynomial time with bounded error probability. However, there is no guarantee of efficiency in finding their solution.<sup>39–41</sup>

Our argument is that a large class of chemically relevant phenomena can be addressed by algorithms solving efficiently solvable problems, which is grounded on the following conjectural, intuitive argumentation: namely, that phenomena that occur in finite time in a chemical laboratory can be simulated in finite time, and, moreover, that the gap between time in nature and time on a quantum computer is only polynomially sized.

Our argument is that a large class of chemically relevant phenomena can be addressed by algorithms for efficiently solvable problems. More precisely, our central hypothesis is that phenomena that occur in finite time in a chemical laboratory can be simulated in finite time, while the gap between time in nature and time on a quantum computer is only polynomially-sized. In other words, we conjecture that the relevant system sizes and time scales for observable chemistry phenomena often grow at most polynomially with system size, allowing for in principle efficient quantum simulations. For instance, while undergoing chemical reactions, systems such as those discussed in Section 1.2 do not typically cool down to their electronic ground state, making thermal states or efficiently reachable metastable states<sup>42</sup> more relevant.

Here, we can draw a distinction between problems of polynomial complexity, such as those from BQP, and QMA-hard problems, such as the local Hamiltonian problem. We note that the QMA-hardness of the local Hamiltonian problem is not sufficient to infer the same complexity for a molecular Hamiltonian with the continuous two-body Coulomb interaction term, although folklore oftentimes considers this to be the case. Therefore, we envision a fundamental change in the way chemistry problems on quantum computers are typically approached, and we display this in Fig. 1. We denote by COMP-CHEM the set of computational problems of chemical relevance – including phenomena like ground states. Note that these are not necessarily decision problems; our argumentation is based on circuit fragments that stem from known complexity results rather than decision problems and the language of complexity theory. The set that we call COMP-CHEMPOLY represents the focal set of problems within this work; namely, problems of relevance

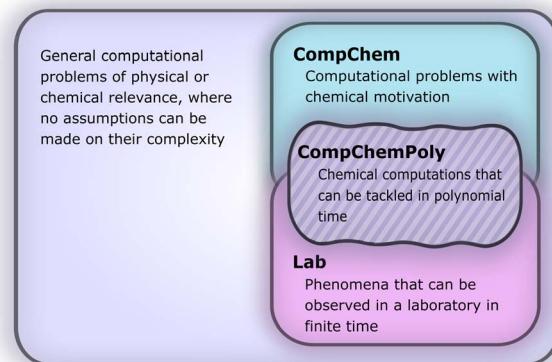


Fig. 1 Complexity of solving chemical problems. We target the set of computational problems, COMP-CHEMPOLY, consisting of problems within chemistry with polynomial complexity when solved on a quantum computer. These problems have unknown overlap with those that correspond to observables in a laboratory; we conjecture this overlap between LAB and COMP-CHEMPOLY to be non-trivial.

to chemistry and that have polynomial complexity. Our conjecture here is thus that chemical problems that are feasible in a laboratory, LAB in Fig. 1, have non-trivial overlap with chemical simulation problems that can be simulated efficiently, COMP-CHEMPOLY. Our present study highlights a wide set of chemically interesting problems involving simulation as a subroutine yet avoiding subroutines and decision problems which are known to be hard. Many approaches in chemistry that involve searching for the ground state are related to problems which are QMA-, StoqMA or NP-hard;<sup>12,43–45</sup> examples of tasks that fall into this category are finding the universal functional in density functional theory or the Hartree–Fock problem. Significant progress was made in solving the latter numerically, to the point that it is mostly seen as “solved” despite the formal hardness. This view results in hope about tackling the ground-state problem as well. Yet, at this point, we seem far from similar success in chemistry ground-state search, and hope for solving practical chemical problems may thus lie in dynamics.

To represent chemical systems, we choose molecular Hamiltonians that are not necessarily restricted to the Born–Oppenheimer (BO) approximation. Non-BO was previously explored in the split-operator formalism on quantum computers, where kinetic and potential dynamics are factored.<sup>46,47</sup> While it was argued in ref. 46 that the non-BO approach is more efficient than implementing the BO procedure, there have been recent advances that render BO more efficient thanks to a fully coherent implementation.<sup>48</sup> Hamiltonians occurring in chemistry are composed of the operators laid out in Table 1, from which locality is an obvious consequence, as  $k$ -locality follows from the 2-body nature of the interactions.

Our framework is independent of the choice of basis used to represent the Hamiltonian numerically and the choice between the first and second quantization. Asymptotically, a first-quantized representation tends to be more efficient for fault-tolerant quantum algorithms with an abundant number of



**Table 1** Hamiltonian components for molecular systems interacting with photons. For a given particle with mass  $m$  and charge  $q$ ,  $\mathbf{p}$  and  $\mathbf{r}$  are, respectively, the corresponding momentum and position operators. For a photon with frequency  $\omega$ ,  $a$  and  $a^\dagger$  are the corresponding annihilation and creation operators.  $A(\mathbf{r}) \propto \mathbf{c}(\mathbf{k}, \mathbf{r})a + \mathbf{c}(\mathbf{k}, \mathbf{r}) \times a^\dagger$ , where  $\mathbf{k}$  is the photon wave-vector, and  $\mathbf{c}(\mathbf{k}, \mathbf{r})$  is a polarization vector, is the vector potential corresponding to the photon field

Charged particle kinetic energy	$\frac{\mathbf{p}^2}{2m}$
Photon energy	$\omega \left( a^\dagger a + \frac{1}{2} \right)$
Interparticle potential	$\frac{qq'}{ \mathbf{r} - \mathbf{r}' }$
Photon-particle interaction	$\frac{q}{m} \mathbf{A}(\mathbf{r}) \left[ \frac{q}{2m} \mathbf{A}(\mathbf{r}) - \mathbf{p} \right]$

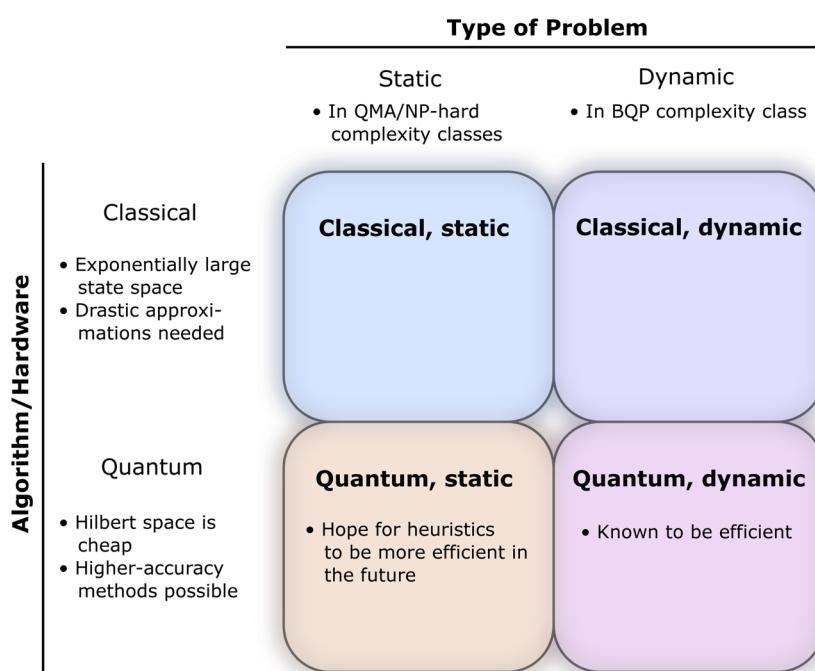
logical qubits, as the number of required qubits grows linearly with the number of particles and logarithmically with the number of basis functions (since one stores the basis information for every particle), as opposed to the linear dependence in the number of basis functions for second quantization, where one stores occupation for each basis function.<sup>30,31,49–53</sup> In this work, we restrict ourselves to the first-quantized representation for concrete examples. For a circuit to have polynomially-sized complexity, it is further necessary that the operator norm of the simulated Hamiltonian, or rather, its individual terms acting on at most  $k$  qubits, grow at most polynomially in the number of qubits used to represent them.<sup>7–9</sup> Energy is an extensive thermodynamic property, so it grows roughly linearly

by increasing the system size (*i.e.*, the number of particles). Therefore, the thermodynamic relation between the amount of matter and internal energy is linear.<sup>†</sup>

## 1.2. Measuring dynamical quantities of interest and a review of exemplary applications

While more details about the framework will be provided below, here, we first provide a brief motivation and review of some of the potential applications. Simulating the dynamics of molecular processes provides access to meaningful information about the rearrangements of atomic nuclei and electrons and their interactions with electromagnetic fields as they unfold. Already with classical resources, time-dependent approaches offer numerous advantages over time-independent ones, as the former are more amenable to handling the continuum portion of the spectrum and grant access to the relevant elements of the scattering matrix<sup>‡</sup> over a meaningful range of energies. We provide an overview of chemical problems, with time-dependent *versus* stationary quantities, solved on different hardware, in Fig. 2.

Measuring observables in a dynamical picture requires considering the time evolution of a wave packet, *i.e.*, a superposition of solutions of the time-dependent Schrödinger equation (TDSE). Most measurements of dynamical quantities can be phrased in terms of wave packet correlation functions, whose calculation fits perfectly into our framework. Transition amplitudes are measured, *e.g.*, using a Hadamard test.<sup>46,54,55</sup> We can follow the scheme introduced in ref. 56, which is capable of



**Fig. 2** Classifying chemical problems related to their hardness and space complexity. Dynamical properties are quantumly efficient, whereas static properties are generally hard. As quantum computers do not suffer from the curse of dimensionality, one can expect the sweet spot of quantum simulations, up to constant factors in the cost, to lie in the evaluation of dynamic properties, assuming suitable initial states can be prepared.



obtaining two-point correlation functions, or the cumulative correlation function method in ref. 57. This scheme is extendable to the  $S$ -matrix and to  $n$ -point correlators by additional time-evolutions and block encodings, similarly to the linear response framework in ref. 55. If results at more than two times are required, we can use a history state encoding following the conditional time evolutions of the quantum circuit of quantum phase estimation, similar to the construction in ref. 58, which also considers observables measured across various time-steps.

To that end, we add a clock register  $\sum_{\ell=0}^{n_t} |\ell\Delta t\rangle$  so that  $n_t\Delta t = t$ .

Then, instead of a direct time evolution of the overall system, we split the evolution into chunks of  $\Delta t$  and condition on the clock register to produce a superposition of the state at different evolution time steps. Furthermore, if needed, applying a quantum Fourier transform before measurement is straightforward and does not compromise the overall efficiency of the algorithm. While this framework could be interpreted as purely theoretical, we can easily show that this approach covers a vast range of contexts, extending from nature to chemical laboratories.

**1.2.1 Reaction rates.** Chemists are commonly interested in observed kinetics, *i.e.*, reaction rates, a topic inherently suitable for quantum dynamics simulation. For instance, dynamical simulation for nonadiabatic processes such as charge transfer reactions falls into the class of quantum circuits of polynomial complexity.<sup>59</sup> Our framework allows using the measurement schemes proposed in ref. 46 in which the reaction rate can be computed from the degree of localization of the wave packet in nuclear configuration space corresponding to the formed products. In addition to making the simulation itself achievable, the quantum approach reduces the measurement to a binary search instead of the more complicated evaluation of a time correlation function.<sup>60</sup> If needed, the latter can also be computed from the dynamics *via* a rate constant calculated as a function of the scattering cross-sections.

**1.2.2 Photochemistry and photophysics.** Photochemical processes are triggered when a molecule absorbs a photon. In these transformations, electronically excited states become populated, giving access to reactive channels that are thermally unachievable.<sup>61</sup> The interaction between a molecule and an external photonic field alters the potential energy surface on which wave packets propagate. Therefore, the nuclear and electronic degrees of freedom cannot be decoupled, and the Born–Oppenheimer approximation breaks down.<sup>62</sup> The dynamics simulation of these systems on classical computers becomes prohibitive after including very few degrees of freedom. In contrast, the inclusion of photonic fields in the quantum simulation is straightforward by explicitly adding a register for photonic degrees, an explicit bath, or a modified Hamiltonian as in Section 3.1. Further photophysical processes that can be simulated are, for instance, when spectroscopic measurements do not suffice, ultrafast spectroscopy, which tracks the system during a light-initiated transformation. Other examples are light-harvesting complexes, molecular machines, and photovoltaics; recent work in ref. 63 also proposed

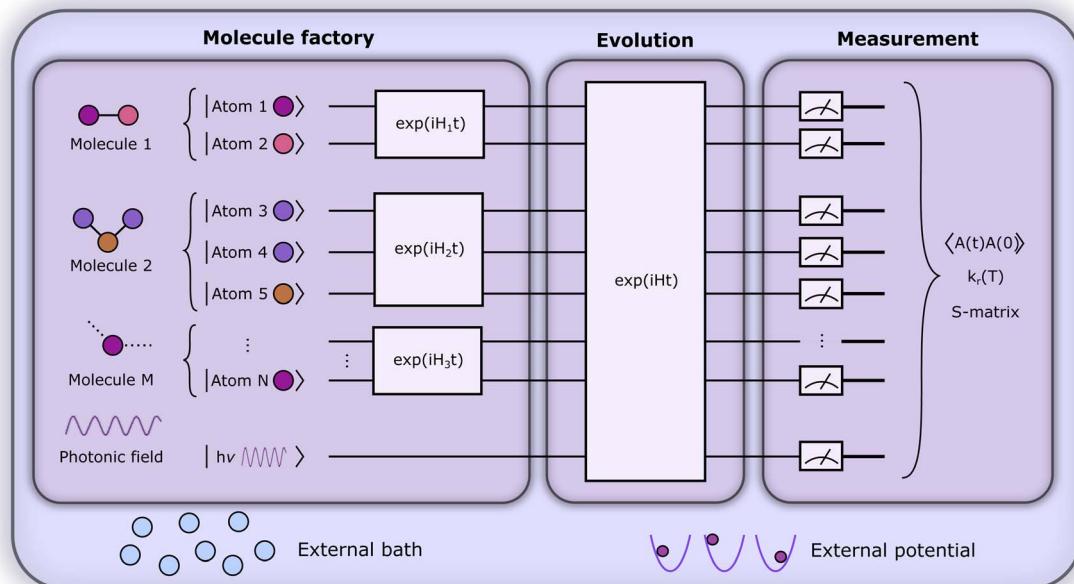
a dynamics-based algorithm for singlet-fission solar cell design. A dynamics-focused approach would make it possible for quantum computers to simulate quantum systems, *e.g.*, optoelectronic devices, where the quantum dynamics are far more important than the eigenstates of the Hamiltonian. In these devices, the relaxation pathways of the excitons are exploited for light generation and harvesting. Classical simulations suffer from the large space of excitons and phonon coupling, making current simulations beyond hopeless. Conversely, simulating the time evolution of thermal quantum states inherently captures all the necessary behavior.

### 1.2.3 Linear and non-linear molecular spectroscopy.

Another class of problems that dynamics can solve is the in-laboratory characterization of molecular structures and properties. The absorption spectrum of a molecule is given by the Fourier transform of the wave packet autocorrelation function.<sup>64</sup> By judiciously selecting an initial state, measuring the autocorrelation function can provide us with different spectra, including electronic, vibrational, and rotational.<sup>65</sup> Through the computation of emission and absorption spectra, our approach accommodates the simulation of fluorescent systems, such as those used in biomolecule marking or thermally activated delayed fluorescence, where forbidden relaxations are assisted by thermal coupling to the environment. Additionally,  $n$ -time correlation functions allow the exploration of linear-response molecular spectroscopy beyond UV-Vis and fluorescence, such as rotational or vibrational spectroscopy, in which contributions from excited states are typically significant at room temperature. Simulating spectroscopic measurements can be used not only to reproduce experiments but also to probe the simulated quantum system, *e.g.*, the presence of an IR signal may indicate the formation of a molecule, allowing reactions to be flagged in the molecular factory approach presented below and depicted in Fig. 3. Two-dimensional spectroscopy could also benefit enormously from extracting  $n$ -time correlation functions from dynamics by following a dynamical approach. For instance, two-dimensional infrared (2D IR) spectroscopy reveals second-order vibrational couplings, which characterize molecular interactions.<sup>66,67</sup> Classical simulations of this state-of-the-art technique typically exclude anharmonicity and produce errors associated with the BO approximation and vibrational population transfer. In this example and many others, finding the ground state of the system is far removed from reproducing the experimental spectrum.

**1.2.4 Free energy simulations.** Free energies play a role in many naturally occurring physical processes, as they determine whether a process occurs spontaneously, such as whether a ligand binds to a protein, whether a material such as salt dissolves in water, or into what shape a protein folds. The type of free energy relevant to a specific system depends on the nature of the system. In the case of an isothermal, isobaric system that only allows volumetric work, the relevant free energy is the Gibbs free energy. In applications, one is primarily interested in free-energy differences between two different states of the system, which can be characterized by two Hamiltonians,  $H_1$  and  $H_2$ . There exist multiple methods for calculating the free-energy difference between these two states.





**Fig. 3** Computational framework. Our simulation framework can be separated into a state preparation procedure ("molecule factory"), the evolution of the reaction of interest, and a measurement step that extracts useable information. The molecule factory prepares a set of molecular input states for the reaction, which may resemble thermal or ground states. These states are produced in a tree-like fashion equipped with a weak measurement scheme to ensure that the target states are prepared with sufficient probability in a heralded way while ensuring the correct spin-symmetry throughout the process, if needed. A photonic field serves as a source of energy for reactions, and an external bath, either explicit or implicit, serves as an energy sink. Furthermore, we utilize external potentials in the spirit of optical tweezers tailored to the different Hamiltonians throughout the procedure to ensure sufficient success probability and to control positions in space.

One class of methods uses fluctuation relations, such as the Jarzinsky equation, that require ensembles of dynamics simulation of the systems of interest.<sup>68</sup> To estimate the free-energy difference, these methods require evaluating the total energy of each trajectory for both the initial and final states. A time-dependent Hamiltonian  $H(\lambda(t))$  is used to alter the system from state 1 to state 2, where  $\lambda$  is an externally controlled parameter such that  $H(0) = H_1$  and  $H(1) = H_2$  (the transition is generally non-adiabatic, and states do not have to be or remain in the ground state such as in adiabatic quantum computing). The number of energy values required depends on the type of problem considered, the speed at which the Hamiltonian is transformed, and the desired accuracy. There are two straightforward ways to calculate free energies in alignment with our framework. First, one can create a sufficiently large number of identical systems and perform separate quantum simulations before measuring the total energies of the initial and final states. Alternatively, one can use the clock-register approach from above and simulate a single trajectory. To obtain a sufficiently large number of statistically independent energy values of both initial and final states, one can simulate each state long enough to take measurements with sufficiently large time intervals in between. Once individual energies are measured, the averages and final calculation of the free energy difference can be carried out on a classical computer. This approach is expected to be preferred in terms of quantum and classical resources compared to representing a complete ensemble as

long as the variance of single-trajectory estimates is not significantly larger than the joint estimate.

**1.2.5 Quantum machine learning.** Instead of directly measuring interesting quantities, the output of the dynamical quantum simulation can be processed by quantum machine learning. This framework can be envisioned as machine learning with quantum input data, with possible classical or quantum outputs.<sup>69,70</sup> In particular, recent results indicate that there is a provable advantage in the efficiency of extracting information when given access to multiple copies of a state in a format that a quantum computer can manipulate compared to having access to only measurements performed on the state.<sup>71</sup> Thus, certain chemically relevant properties of states and quantum evolutions may be learnable more efficiently in our quantum computing framework than in a conventional experiment. Alternatively, it is also possible to measure spectroscopic quantities and replace physical experiments in machine learning pipelines that operate on spectroscopic results, potentially aiding further developments in molecular design.<sup>72,73</sup>

## 2. Computational framework

In what follows, we describe the computational framework in more detail – *i.e.*, a dynamics-based state-preparation scheme that serves as input for a “main” quantum dynamics routine, preceding measurement. Our overall idea is based on the

experiment in ref. 74, where molecules were “built” using two atoms with finite success probability, as well as a mergo-association scheme that merges atoms, *e.g.* Rb and Cs, using a trap potential.<sup>75–77</sup> The chemical dynamics we aim to simulate to extract chemical properties requires a sensible initial state for the time evolution. This state must reproduce a natural one to faithfully extract the desired properties. Depending on the process under consideration, good candidates could be found among ground states, or perturbed ground states, such as those obtained by laser excitation. However, as previously argued, general molecular ground states are hard to access. Nevertheless, preparing the ground states of atoms is feasible – the lighter the atom, the easier – and can be done efficiently as constant overlap for heuristic input states is expected.<sup>37</sup> For this reason, we propose to follow a hierarchical approach, as outlined in Fig. 3. All processes generating the reactants and products are carried out through Hamiltonian simulation with local Hamiltonians, including molecular and external field components. After using this technique to generate initial states, we aim to measure observables, *e.g.*, reaction rates according to the scheme in ref. 46 and 54 or auto-correlation functions,<sup>55</sup> as outlined in the previous section.

We start with a state representing  $N$  atoms, all of which are assumed to be in their ground states or another state of interest  $\rho_i^{(\text{atom})}$ , such as a thermal state. The preparation of these states can be achieved by existing algorithms for ground-state<sup>10,78–81</sup> and thermal-state preparation<sup>82,83</sup> followed by amplitude amplification to enhance the probability of obtaining the desired state. For generality, we represent both pure and mixed states using density-matrix notation. Since the atoms are all finitely sized, and we can prepare an initial state with a significant overlap with the desired state for each atom, any amplitude amplification costs are constant with respect to overall system size. The overall cost to prepare the atomic initial states is  $O\left(N \text{poly}\left(\frac{1}{\varepsilon}\right)\right)$ , with  $\varepsilon$  being the accuracy in preparing the atomic states. Using these atoms encoded as a state  $\rho_1^{(\text{atom})} \rho_2^{(\text{atom})} \dots \rho_N^{(\text{atom})}$ , we seek to create  $M$  molecular reactants  $\rho_1^{(\text{mol})} \rho_2^{(\text{mol})} \dots \rho_M^{(\text{mol})}$ . For this step to produce a molecular state from atomic states, appropriate (anti)symmetrization is essential too; see also the discussion in Appendix A.2.2.

Based on these initial states, we next prepare the reactants by a scattering process mimicking a physical experiment. In other words, we jointly evolve a set of atoms meant to form a reactant molecule. We discuss the modeling and treatment of a bath that allows exchanging energy with an environment further below.

Consider one of the reactant molecules to be prepared. We combine the constituent atoms by successively scattering in a tree-like manner, as in Fig. 3, until the desired molecular state is attained. In the worst case, each molecule is generated by combining only two participants at a time, leading to a binary scattering tree, and thus an overall number of scatterings that is quasi-linear in the number of input atoms. It is essential to ensure a high overlap with the desired intermediate state at each level in this procedure. Otherwise, the overlap would decrease at each combination step. For instance, with an initial

overlap of  $(1 - \delta)$ , the overlap would drop exponentially to  $(1 - \delta)^n$  after  $n$  steps. A seemingly obvious choice for mitigating this problem would be to use oblivious amplitude amplification.<sup>84</sup> However, in our case, the open-system character of the simulation and the fact that the abstract angle to be amplified is not independent of the (unknown) input state are major roadblocks for this approach, which we leave up to future research. Instead, we propose the following approach towards bounded success probability. Inspired by the nanoreactor approach in ref. 85 and the mergo-association scheme from ref. 75–77, we introduce artificial potentials that confine the products to be combined in each scattering step, as well as an additional photonic field as an energy source and a bath as an energy sink. Following the procedure we propose in Section 3.1, mergo-association shows more promise in terms of suitability for simulation on the molecular scale at this point as we can confine individual, small systems without the need for large ensembles in high-pressure and high-temperature regimes. Using this framework, we show that, heuristically, we can expect a constant lower bound  $P$  on the probability of success for certain types of bonds for the chemical formation of reactants. Such a lower bound is an important assumption in the procedure and also highlights its heuristic nature. For the example of mergo-association, we show that the existence of a constant lower bound on the success probability for certain types of bonds, such as covalent bonds, is reasonable; *cf.* Section 3.1. The intuition we observe there is that given a certain order of magnitude of bonding energy (say, the typical regime of covalent bonds of approximately 30–180 kJ mol<sup>−1</sup>), a requirement of simulating potentials of similar magnitude may lead to a substantial constant factor in the simulation cost, but the scaling with respect to system size of this cost would heuristically be only polynomial in the combined nuclear masses, not exponential.

Suppose the scattering is organized as a binary (or higher-order, ternary, ...) tree. For each node in the tree, there is a simulation channel  $\mathcal{E}(\rho)$ , parametrized by the subsystem size, dissipation model (bath), and confinement properties (artificial potentials). Then, we may assume that this channel produces a state of the kind

$$\mathcal{E}(\rho_0) = p_0 \rho_{\text{suc}} + (1 - p_0) \rho_{\text{nsuc}} + C, \quad (1)$$

with given input  $\rho_0$  and probability of success  $p_0$ ;  $C$  describes any potential coherence between the subspace of successful and unsuccessful scatterings and comes from overlap terms such as  $\text{tr}[\rho_{\text{suc}}^\dagger \rho_{\text{nsuc}}]$ . We may similarly assume we have an observable  $\mathbb{O}_{\text{suc}}$  that allows us to distinguish between the subspaces spanned by  $\rho_{\text{suc}}$  and  $\rho_{\text{nsuc}}$  by weak measurements, which enables the heralding of the scattering success and projecting the state into the successful subspace. The presence of a coherence  $C$  does not influence measurement outcomes of  $\text{tr}[\mathbb{O}_{\text{suc}} \mathcal{E}(\rho_0)]$ ; see Appendix A.2.3. Using these two building blocks, we will now outline the simulation strategy of each node, as depicted in Fig. 4. As depicted in this figure, we propose alternating simulations and weak measurements of  $\mathbb{O}_{\text{suc}}$ , terminating on the heralding of a successful reaction and the corresponding projection on  $\rho_{\text{suc}}$ . Note that said measurement of  $\mathbb{O}_{\text{suc}}$ , even if carried out weakly, could disturb the state



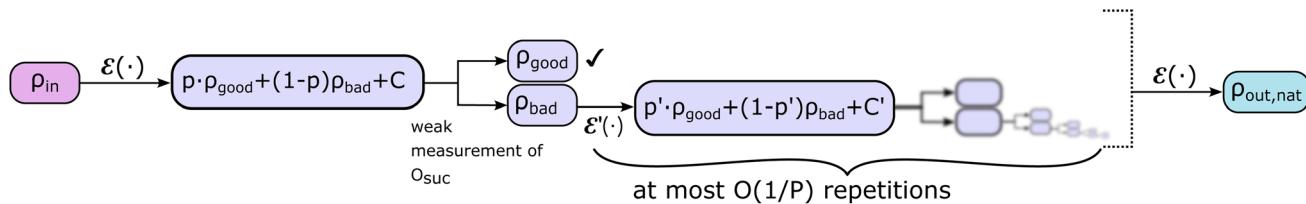


Fig. 4 Procedure of a single scattering step. Evolution through the simulation channel  $\mathcal{E}(\cdot)$  produces states overlapping successful and unsuccessful spaces. To project the state onto one of these subspaces, a weak measurement is performed, yielding either success or failure. In the case of success, proceed and potentially apply another step of time evolution to ensure the state represents a natural state. If unsuccessful, perform an additional, possibly modified, time evolution, which again produces an overlap in the successful subspace, then measure again. Repeat this until success, the expected number of repetitions scaling inversely with the lower bound  $P$  on the success probability per step. Success can be quantified by a weak measurement of an observable  $\mathbb{O}_{\text{suc}}$  that, e.g., signifies the success of forming a bond by capturing spatial proximity. For details on the weak measurement procedure and measurement oracle construction as well as antisymmetrization, see Appendix A.2.

such that it does not describe a realistic state encountered in nature any more after the measurement, especially in the case of the successful outcome. However, we assume that the simulation channel  $\mathcal{E}(\cdot)$ , such as the one devised below through mergo-association, does resemble a process in nature. Thus, it tends to make states decay towards such physically meaningful states, and another application of the channel will, therefore, map us back into a state resembling those in nature,

$$\mathcal{E}(\rho_{\text{suc}}) = p_{\text{suc}}\rho_{\text{suc}}^{\text{nat}} + p_{\text{not-suc}}\rho_{\text{not-suc}}^{\text{nat}} + C^{\text{nat}}. \quad (2)$$

If the molecule is unstable, but the reacted state is desired, we may simply skip this reapplication and proceed. Hence, if measuring  $\mathbb{O}_{\text{suc}}$  yields a positive outcome, generating meaningful states with a high success rate seems plausible. Further, we can use the following strategy to ensure success even if  $\mathbb{O}_{\text{suc}}$  shows that scattering was unsuccessful. Although the state is projected towards the unsuccessful subspace (see Appendix A.2.3 for details), failure is heralded. Therefore, we can apply another simulation channel  $\mathcal{E}'(\cdot)$ , which may be slightly modified from  $\mathcal{E}(\cdot)$  (e.g., stronger confinement or dissipation), leading to

$$\mathcal{E}(\rho_{\text{post-meas}}) = p_1\rho_{\text{suc}}' + (1-p_1)\rho_{\text{not-suc}}' + C'. \quad (3)$$

This assumption is similar to that in ref. 79 in that we expect a redistribution of probability toward the state of interest. We can then redo the measurement and iterate between measurements of  $\mathbb{O}_{\text{suc}}$  and (progressively more) modified simulation channels until success occurs.

To summarize, we can use the above strategy to create a tree-like sequence of scatterings so that, at each step, the success probability at each of the nodes in the tree is bounded by, say,  $1 \geq P > 0$ . As we can herald success, at most  $O(1/P)$  repetitions are required per node. Since  $P$  is assumed to be fixed, the number of repetitions per node is  $O(1)$ . Furthermore, since failure does not require a restart from the leaves of the tree, the complexities of individual nodes straightforwardly add up. Hence, the number of repetitions grows linearly with the number of nodes in the tree, which, with  $N$  initial constituents, goes at most as  $O(N \log(N))$ . In Section 3.2 and Appendix A.2, we provide a more

detailed discussion of the construction of the weak measurements.

We leverage parallels to optical tweezers<sup>86,87</sup> or molecular beams<sup>88,89</sup> that are used in physical experiments to engineer the aforementioned artificial potentials to boost the success probability of scatterings. Before we go into details regarding a specific instance in terms of mergo-association in Section 3.1, we first outline more general approaches here. Molecular beams offer well-established success rates, though they require an abundance of particles and, hence, an abundance of available qubits. On the other hand, traps *via* artificial potentials – e.g., adding to the Hamiltonian a harmonic potential term confined to a specific region as done in eqn (4) – can considerably boost success without high additional space requirement. Many of the applications we will discuss benefit from reactant states in a form resembling nuclear wave packets. To that end, consider the input states for the scattering process to be prepared as such. Then, the said process supported by artificial potentials would, in most physical cases, largely preserve the locality in configuration space and maintain the wave-packet-like nature of the states. At this point, one may choose to induce a reaction by explicitly modeling a photonic reservoir to excite the reactants.<sup>90</sup> Additionally, an explicit register for a bath can be used to absorb excess energy once the reaction has occurred, allowing the products to relax. Alternatively, instead of explicitly tracing out a subsystem here, post-selection on a ‘successful’ reaction, as described later in Section 3.2, can serve to model energy moving out of the system. Finally, energy dissipation can be implemented using Markovian open-system simulation methods, as discussed in more detail below.

Beyond the scattering approach, one could think of using a molecular Hartree–Fock (HF) state as an alternative heuristic for an initial state for the reaction dynamics under investigation. One can expect the efficacy of this approach to be limited to cases where the Hartree–Fock state as input to a simulation channel is close enough to the manifold of physically meaningful states so that convergence to a desired state as input to a reaction happens in controllable time, such as in systems with low static correlation. While such systems tend to be amenable to classical treatment, they may be candidates for early experiments of our approach on quantum devices, as it is likely that



using an HF initial state has considerably lower constant factors than the scattering approach. In contrast, the scattering approach will apply to more general states (such as those where Hartree–Fock does not provide a sufficient heuristic).

As mentioned above, the embedding of processes into a larger environment play an important role in the framework during the molecular preparation stage. The ability to dissipate excess energy is essential for both the probability of successfully forming stable bound states and the ability of the dynamics to emulate the open-system evolution of chemical experiments. Recently proposed methods for simulating the weak coupling to a large, memory-less (*i.e.*, Markovian) environment modelled by a Lindbladian operator<sup>91,92</sup> can be used to model the presence of such a dissipative bath.<sup>82,83,93–95</sup> This simulation is efficient in the size of the system and for some methods is shown to converge to a thermal state.<sup>82,83</sup> The only parameter of this procedure that does not scale polynomially with system size is the thermalization time, which is difficult to predict and can, in principle, grow exponentially with system size. However, our observation is that slowly thermalizing systems in our simulation correspond to systems that also thermalize slowly in nature. Thus, for physically motivated open-system models, we would expect to produce either thermalized, metastable, or slowly thermalizing states, depending on which ones are prevalent in nature. A good example of these types of systems would be a glassy molecular mixture. From this perspective, we conjecture that polynomial-time simulations are sufficient to reach all chemically relevant states.

One thing to note is that the system on which the readout is to be performed may need to include some degrees of freedom surrounding the molecule, *e.g.*, if solvent effects, photon or phonon couplings, or non-Markovian dynamics<sup>96</sup> are important. The preparation of this more explicit bath follows the same framework as the main molecular degrees of freedom.

### 3. A scattering-based state-preparation step

With a conceptual framework in mind, in this section, we discuss a specific instance of a single scattering step. The approach is based on an external-potential assisted merging of molecular fragments and the construction of a success-heralding measurement oracle. In the first part, we discuss the coherent part of the simulation of the merging. Part of the open-system character of the approach in Fig. 3 is brought into play by means of the measurement oracle outlined in the second part of this section.

#### 3.1. Molecular states by mergo-association

A promising approach to realize the assemblage of molecular wavefunctions in Fig. 3 is mergo-association, as considered in ref. 75–77. Here, optical tweezers confine two fragments – *e.g.*, two hydrogen atoms – to form a bond.

To mimic this, we propose to carry out the simulation in the following way. We consider a real-space first-quantized representation the most convenient with oracles for the repeat-until

success procedure. We closely follow the implementation put forward in ref. 30 to represent the Hamiltonian. However, note that the core building block is open-system simulation, *e.g.*, implemented by combining the methods of ref. 30 with open-system techniques, as proposed in ref. 82. Then, we place the two participating systems at a reasonable estimate for a bonding distance. For dihydrogen, this is well known; for more involved systems, a heuristic guess needs to be found. Initially, we implement their dynamics according to two independent Hamiltonians,  $H_A$  on subsystem A and  $H_B$  on subsystem B. Then, we slowly turn on inter-system Coulomb interactions  $H_{AB} = V^{\text{int}}$  together with trapping potentials  $V^{\text{trap}}$ , according to

$$H(s) = H_A + H_B + f(s)H_{AB} + g(s)V^{\text{trap}}. \quad (4)$$

For the example of two nuclei, modeling the trap by a harmonic potential acting on nuclear coordinates  $R_1, R_2$ , its functional form is given simply by<sup>75–77</sup>

$$V^{\text{trap}}(R_1, R_2) = V_1^{\text{trap}}(R_1) + V_2^{\text{trap}}(R_2), \quad (5)$$

$$V_j^{\text{trap}}(R_j) = \frac{m_j}{2}(R_j - R_{0,j})^T \omega_j^2 (R_j - R_{0,j}), \quad j = 1, 2. \quad (6)$$

More details on this are provided in Appendix A.1.1. A qualitative choice of  $f, g$  is displayed in Fig. 5. After the merging ( $s \geq s_0$ ), the trap is re-released ( $s_0 \leq s \leq s_1$ ), while the interaction stays on. With a certain probability of success, the state will not undergo a diabatic transition into a higher-lying excited state, meaning a successful merger has occurred. The specific design of scheduling functions is left up to further research. However, in order to mimic mergo-association, we propose the use of an  $f(s)$  that schedules the inter-species interactions to resemble the trajectory of a Coulomb potential. More details on a quantum encoding of the trap potential can be found in Appendices A.1 and A.1.1.

Equivalently to the adiabatic evolution in  $s$ , one may put system A and system B at initial distance  $z^*$  and move their centers of masses together at a predetermined rate  $\frac{d\Delta z}{dt}$  until the bonding length is reached. Hence, the scheduling functions  $f(s)$ ,  $g(s)$  are not present in the Hamiltonian anymore, and naturally the strength of the interactions is implicitly determined by distance, recovering the procedure of mergo-association.

The mergo-association scheme we follow relies on assumptions from scattering theory. The setup in scattering theory is that two scattered objects, situated beyond a certain scattering length, can be modeled as free particles. Within the scattering length, the scattered product is considered as ‘one’. Conceptually, this is analogous to the ‘molecule factory’ in Fig. 3. The scattering length, or harmonic length, is defined as  $\beta = \sqrt{\hbar/(\mu\omega)}$  with a frequency  $\omega$  and relative mass  $\mu = m_1m_2/(m_1 + m_2)$  and induces a notion of a bonding energy *via*  $\hbar\omega$ . As we are interested in chemical bonds, a proxy for this scattering frequency can be the most weakly bound state or the highest-lying bonded vibrational excited state.



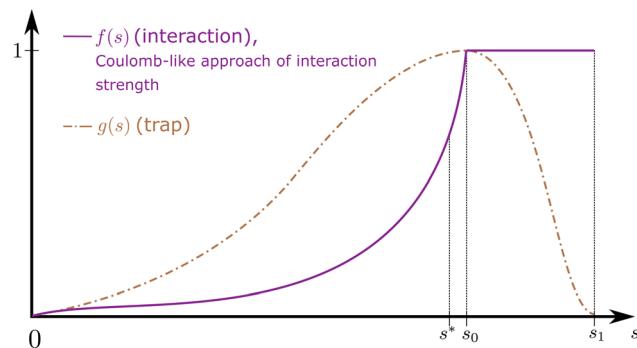


Fig. 5 Scheduling of interactions. Inter-species interaction is described by  $f(s)$  so that  $f(0) = 0$ ,  $f(s \geq s_0) = 1$ , and  $f$  is monotonically increasing until  $s_0$  and then constant. The harmonic trap follows the schedule  $g(s)$  with  $g(0) = 0$ ,  $g(s_0) = 1$ , and  $g(s_1) = 0$ . It is monotonically increasing until  $s_0$  and then decreasing until  $s_1$ . Following adiabatic evolution, there is a 'point of contact' of states at  $s^*$ , a little earlier than  $s_0$ , which is the evolution parameter (with corresponding effective distance) used to evaluate the diabatic transition probability. Here, we assume the scenario in which the constituents are already placed at close distance and we 'slowly turn on' the Coulomb interaction, where with the trajectory of  $f(s)$ , we aim to mimic an evolution of interaction strength that is Coulomb-like if they were to approach each other. To that end, let  $z_{0,1}$  and  $z_{0,2}$  be the centers of the traps. Then, in order to follow  $V^{\text{int}}(R_1(s), R_2(s)) = \frac{q_1 q_2}{|\Delta R_{12}(s)|}$  with nuclear charge  $q_j$ , let the implemented interaction be  $V^{\text{int}}(s) = f(s) H_{AB} = f(s) \frac{q_1 q_2}{|z_{0,1} - z_{0,2}|}$ . Via  $f(s) = \frac{|z_{0,1} - z_{0,2}|}{|\Delta R_{12}(s)|}$ , the desired motion of the nuclei can then be emulated. Towards  $s \rightarrow s_0$ , the potential that is fixed at the strength corresponding to the trap centers needs to be replaced by the actual strengths to account for fluctuations in the positions which will matter at that stage.

The key aspect of such a mergo-association is the choice of trap frequency  $\omega$  and the consequences on the diabatic transition probability. Here, we work with an isotropic trap for simplicity, hence eqn (6) is described by a single trap frequency  $\omega$ . Successful, in our case, means that a (e.g., covalent) bond has been formed and that there is no diabatic transition beyond a certain vibrational frequency of the trapped system, say  $\omega_a$ . Then, it is possible to approximate the probability of success by using the Landau-Zener probability,<sup>97</sup>

$$p_{\text{succ}} \geq 1 - p_{\text{LZ}}; \quad p_{\text{LZ}} \approx \exp(-2\pi\Gamma) \quad \text{with} \\ \Gamma = \frac{\omega_{\text{eff}}^2/\hbar}{\left| \frac{d}{ds}(E_{\text{mol}} - E_{\text{atom}}) \right|}. \quad (7)$$

Here,  $\omega_{\text{eff}}^2$  is the effective potential strength of the trapped system (see Appendix A.1.2). Then, we use the fact that the scheduling function  $f(s)$  can be related to a corresponding inter-nuclear motion  $R(s)$  in a physical mergo-association experiment. Thus, we can rewrite the expression in terms of derivatives of this equivalent position,  $\frac{dE}{ds} = \frac{dE}{dR} \frac{dR}{ds}$ . More details are given in the caption of Fig. 5. Defining the velocity  $v = \frac{dR}{ds}$  and

using the shorthand  $\partial E_{\text{mol}}$ ,  $\partial E_{\text{atom}}$  for the energy derivatives, this gives

$$p_{\text{LZ}} \approx \exp\left(-\frac{2\pi}{\hbar} \frac{\omega_{\text{eff}}^2}{|\partial E_{\text{mol}} - \partial E_{\text{atom}}|v}\right). \quad (8)$$

A more detailed discussion can be found in Appendix A.1.2, and we continue by presenting a simplified final expression. We start by the observation in [ref. 76, eqn (54)] that

$$\omega_{\text{eff}}^2 \approx \omega^2 \tilde{E}_a^{1/2} \exp(-\tilde{E}_a) \quad (9)$$

with  $\tilde{E}_a = \frac{\hbar\omega_a}{\hbar\omega}$  as a 'relative bonding energy'. Together with expressions for energy derivatives from eqn (A16) and an approximation to eqn (9) we can obtain a bound on  $p_{\text{LZ}}$ ,

$$p_{\text{LZ}} \lesssim \exp\left(-4\left(\frac{\pi}{\mu}\right)^{1/2} \left(\hbar\omega \frac{\tilde{E}_a}{3 + \tilde{E}_a}\right)^{1/2} \exp\left(-\frac{1}{2}\tilde{E}_a - \frac{3}{2}\right) \frac{v}{\tilde{E}_a}\right). \quad (10)$$

Ideally, we want  $p_{\text{LZ}}$  to be close to zero so that  $p_{\text{succ}}$  is close to one. Then, up to the speed of evolution  $v$  and effective mass  $\mu$ , this depends on the ratio of binding energy compared to trap energy,  $\tilde{E}_a \sim \omega_a/\omega$ . The inner exponential is closer to one (which means smaller  $p_{\text{LZ}}$ ) for small  $\tilde{E}_a$  and thus for  $\omega$  as large as possible relative to  $\omega_a$ . Beyond this, for fixed  $\hbar\omega_a$ , the other term depending on frequencies,  $\left(\hbar\omega \frac{\tilde{E}_a}{3 + \tilde{E}_a}\right)$ , is also maximal for large trap strength  $\omega$ . However, there is a trade-off in maximizing this probability of success *versus* minimizing the cost factor of the block-encoding of the operator. The latter represents the cost of representing the potential in the block-encoding access model, and is quantified by the operator one-norm of the potential,  $\alpha_{\text{trap}}$ . Generally, we would aim to keep the frequency no larger than necessary to keep the block-encoding cost low. Additionally, we can look at the system of interest in ref. 75–77, namely Rb and Cs. They use  $\omega = 150$  kHz for all of their simulations and experiments. This order of magnitude is about the same as the appearing  $\omega_a$ , which is roughly 110 kHz; the scattering length (regarding binding energy) for this system in a weakly bonded regime is approximately  $645a_0$ .<sup>98</sup> This setup allows<sup>75</sup> high probabilities of success to be observed (see [ref. 75, Fig. 4], close to 80%).

### 3.1.1 Cost estimate of simulating the trap potential.

Typical covalent bonds are of the order of 100 kcal mol<sup>-1</sup>, which amounts to approximately a 35 000 cm<sup>-1</sup> frequency for  $\omega_a$  (approx. 10<sup>12</sup> kHz). Therefore, we expect the required  $\omega$  to be, in general, vastly larger than 150 kHz; however, the hope is that the remaining behavior will carry over, and we only need to scale it according to the relative difference in bonding energy. Then, in fact,  $E_a = \hbar\omega_a$  can remain bounded in the two-atom/molecule scatterings and effectively act as a constant in an asymptotic sense. This means for  $p_{\text{LZ}}$  to be bounded as well,  $\omega$  incurs a (likely substantial) constant factor through  $E_a$  and then scales as  $O(\mu)$  for  $\mu = \frac{m_1 m_2}{m_1 + m_2}$ . Therefore, since at the last scattering



stage in the “molecule factory”,  $\mu \sim N_{\text{nuc}}$ , we can roughly estimate  $\omega$  in the expression of  $\alpha_{\text{trap}}$  by a linear factor  $N_{\text{nuc}}$  and a system-dependent constant factor for the binding energy of covalent bonds. Using this in our asymptotic expression for the block-encoding factor of the trap,

$$\alpha_{\text{trap}} \approx O(N_{\text{nuc}}^3 \mathcal{Q}_{\text{trap}}^{2/3}). \quad (11)$$

We want to compare this result with the corresponding sub-normalization factors of the kinetic energy,  $\sim \frac{N_{\text{el}} N^{2/3}}{\mathcal{Q}^{2/3}}$ , where  $N$  is the number of basis functions across three dimensions and  $\mathcal{Q}$  the volume of the simulation box, and the Coulomb energies,  $\sim \frac{N_{\text{el}}^2 N^{1/3}}{\mathcal{Q}^{1/3}}$ .<sup>30</sup> If the density of grid points grows linearly with the number of particles, then the encoding costs of these operators increase roughly with power  $\frac{5}{3}$  or  $\frac{7}{3}$  of the system size, whereas the cost of the trap scales cubically. It also grows with the volume  $\mathcal{Q}^{2/3}$  rather than with the inverse grid density. A more accurate representation will need a finer grid spacing, increasing the cost of representation for the Coulomb potentials. For the trap potential, the box size matters, which will need to increase for larger systems, although there is no increase in cost for increasing accuracy through a finer grid spacing.

**3.1.2 Additional overheads incurred due to non-BO simulation.** The merge-association approach generally requires simulation of both electrons and nuclei which are confined by the trap. Therefore, non-Born–Oppenheimer simulation is most appropriate, even though approaches based on the techniques in ref. 18 and 48 where nuclei are propagated classically but information is available in a quantum register could also be considered. Due to the higher degree of localization of the heavier nuclei, the grid spacing for the nuclear degrees of freedom will have to be chosen finer than for the electrons. This manifests, *e.g.*, through the de Broglie wavelength of the particles. For an electron in the lowest-energy orbital of hydrogen, the de Broglie wavelength is roughly  $\lambda_e \approx 333$  pm, whereas for a proton it is  $\lambda_e \approx 52$  pm using a velocity of  $7.6 \times 10^3$  m s<sup>-1</sup> approximated through the zero-point vibrational mode of the H<sub>2</sub> molecule. Therefore, the localization of a proton is approximately 6.38 times higher, meaning we may estimate the necessary number of grid-points for the nuclear resolution to be approximately  $6.38^3 \approx 260$  times the one for electronic resolution in three dimensions.

This necessary increase in the number of basis functions has direct consequences on the implementation cost, through query complexity as outlined in the paragraphs above, but also through constant factors.

The constant factor Toffoli complexity regarding the implementation of phase estimation of a Born–Oppenheimer Hamiltonian is reported in [ref. 30, theorems 4 and 5]. Parameters affected by discretization in this expression are the block-encoding factor (in our work  $\alpha$ , in ref. 30  $\lambda$ ) and the number of bits used to store the number of plane-waves used for discretization  $n_p$ .

We first discuss the impact by  $n_p$ . It appears with terms growing at most quadratically; the same scaling would hold when using a grid. Using the fact from above that a grid for protons would need to be roughly 260 times as dense as one for electrons,  $n_p(\text{proton}) \approx n_p(\text{electron}) + [\log_2(260)] \lesssim n_p(\text{electron}) + 9$ , with the  $n_p$  parameter appearing in linear and quadratic order in the cost for both qubitization and interaction picture simulation. Therefore, we can expect the Toffoli count following [ref. 30, theorems 4 and 5] to only grow modestly even though the grid needs to be significantly denser. Additionally, the subnormalization factor of the block-encoding of the Hamiltonian terms may depend on the number of grid-points or basis functions  $N$  that is chosen; see eqn (11) and below it. These terms are linear (potential energy) and quadratic (kinetic energy) in the number of grid points per dimension; a roughly 6.38-fold increase as discussed above would thus lead to a 6.38 times higher subnormalization of the potential energy terms as well as a roughly 40.7 times higher subnormalization for the kinetic energy. As we point out in the previous paragraph, the simulation of the trap itself does not depend on the grid density. Hence overall, we expect a rather low impact of non-BO simulation on the constant factors in the simulation beyond the subnormalization factor; the latter grows linear/quadratic with finer discretization and polynomial in the particle number. Gate counts related to creating superpositions over the grid, *etc.* depend poly-logarithmically on the chosen discretization.

The non-BO simulation approach would be necessary in order to properly represent the merge-association approach. Our state-preparation therefore would incur slightly higher simulation costs for systems that are sufficiently well represented by the BO approximation. Note that following the algorithms in ref. 18 and 48 as an alternative pathway to implement a merge-association-like process, *i.e.*, resolving the nuclear dynamics classically and only the electrons quantumly, may incur similar block-encoding costs that would similarly need to be considered.

### 3.2. Measurement oracles for success heralding

Suppose now that after a scattering step, we aim to herald success according to eqn (2) and (3). To that end, we need a measurement operator  $\mathbb{O}_{\text{succ}}$  able to give the desired partitioning. A more detailed discussion of the following outline can be found in Appendix A.2.

**3.2.1 Success of merging.** First, we start by briefly outlining the success heralding *via* weak measurements. We are given a quantum state that results from a previous ‘scattering’ event, such as by simulating the Hamiltonian in Section 3.1, which is supported by a trap potential. Such a state will be, following the notion of success that we used before, in a superposition of states that underwent a diabatic transition into higher-lying vibrational states or not. This notion of success does not correspond with the one that we can directly observe in a non-destructive manner. Thus, we develop another notion of ‘success’ that we can use for efficient success heralding to ensure a successful merging. The state we assume is in a first-quantized real-space discretization; this may not be an optimal basis but is most convenient for our discussion, and



efficient transformations to other bases can be thought of as well.<sup>51,99</sup> Then, success can be formulated most straightforwardly using proximity as a geometrical criterion. *E.g.*, suppose we collect a set of locations of interest where we want to test proximity across the nuclei,  $\mathcal{I} \subseteq [N_{\text{nuc}}] \times [N_{\text{nuc}}]$ , then

$$\tilde{\mathcal{C}}_{\text{geom},\Delta}(\{r, R\}, \mathcal{I}) = \begin{cases} 1 & \|R_j - R_k\| \leq \Delta_{jk} \quad \forall (j, k) \in \mathcal{I} \\ 0 & \text{else.} \end{cases} \quad (12)$$

with  $\Delta_{jk} > 0$ . We mark success if the criterion  $\mathcal{C}$  returns 1. Note that we choose to restrict the evaluation to nuclear coordinates as the relative position of nuclei within a bonded state is essentially stationary compared to the electrons. The definition in eqn (12) is classical and assumes direct access to  $\{R_j\}_{j \in [N_{\text{nuc}}]}$ , whereas types of states we encode are superpositions of tensor products across basis state labels,  $|r_1, r_2, \dots, r_{N_{\text{el}}}; R_1, \dots, R_{N_{\text{nuc}}}\rangle$ . More generally, we can think of what happens to a mixed state when going through the success heralding. Then, we consider states of the form

$$\begin{aligned} \rho = \sum_{p,q} \rho_{p,q} & \left| \psi(\{r, R\}_p) \right\rangle \left\langle \psi(\{r, R\}_q) \right|, \left| \psi(\{r, R\}_p) \right\rangle \\ & \sim \left| r_1^{(p)}, r_2^{(p)}, \dots, r_{N_{\text{el}}}^{(p)}; R_1^{(p)}, \dots, R_{N_{\text{nuc}}}^{(p)} \right\rangle. \end{aligned} \quad (13)$$

The set  $\{r, R\} = \{r_j\}_{j=1}^{N_{\text{el}}} \cup \{R_j\}_{j=1}^{N_{\text{nuc}}}$  denotes a nuclear and electronic configuration, and  $\{r, R\}_j$  denotes an instance as the  $p$ th term of the superposition. Configurations include a spin degree of freedom. Hence the range of  $p, q$  is the number of possible basis states labeling the grid points.

Now, we assume that there is a quantum circuit that can achieve the action of  $\mathcal{C}$  efficiently, namely a  $U_C$  that takes the state  $\rho$  plus an ancilla and stores success in the ancilla, according to the following sets:

$$\begin{aligned} A &= \left\{ j : \mathcal{C}(\{r, R\}_j) = 1 \right\} \\ B &= \left\{ j : \mathcal{C}(\{r, R\}_j) = 0 \right\}. \end{aligned} \quad (14)$$

The crucial part here is that such a criterion induces a bi-partition of the Hilbert space, splitting the set of states into two groups. Specifically, every state of the form  $|\psi(\{r, R\}_j)\rangle$  either corresponds to a configuration with  $\mathcal{C}(\{r, R\}_j) = 1$  or a configuration with  $\mathcal{C}(\{r, R\}_j) = 0$ . The set  $A$  enumerates the states the oracle accepts as merged, and  $B$  enumerates the states that the oracle rejects. For convenience we will denote  $|\psi_j\rangle := |\psi(\{r, R\}_j)\rangle$  from now on. With this bi-partition, a general input state can now be written as

$$\begin{aligned} \rho = \sum_{j,k \in A} \rho_{j,k} & |\psi_j\rangle \langle \psi_k| + \sum_{j,k \in B} \rho_{j,k} |\psi_j\rangle \langle \psi_k| + \sum_{j \in A, k \in B} (\rho_{j,k} |\psi_j\rangle \langle \psi_k| \\ & + \rho_{k,j} |\psi_k\rangle \langle \psi_j|) \end{aligned} \quad (15)$$

Weak measurement of this oracle and thus flagging of a successful merging can be implemented following the subsequent operations based on work in ref. 100–103:

(1) Append an ancilla qubit,  $q_{a,1}$ , and perform the unitary  $U_C$  on the joint system. This stores the ‘success value’ in the ancilla.

(2) Append a second ancilla,  $q_{a,2}$ , and perform a controlled rotation  $\text{CR}_y(\delta)$  gate on  $q_{a,2}$  conditioned on  $q_{a,1}$ .

(3) Perform  $U_C$  on the input state and  $q_{a,1}$ , resetting  $q_{a,1}$ .

(4) Measure  $q_{a,2}$  in the computational basis and reset it for later use. The measurement on  $q_{a,2}$  is then used as a result flag of a successful merger.

Finally, we can conclude that the measurement oracle  $\mathbb{O}_{\text{suc}}$  is composed of the circuit  $U_C$  and the weak measurement scheme above. More details are given in Appendix A.2.3.

An important aspect to consider in constructing these measurements is that the wavefunctions encoded, even when expressed as a mixed state, need to satisfy the necessary exchange symmetry (antisymmetric for fermions, symmetric for bosons). Ongoing measurement induces the risk of compromising such symmetries. Thus, the measurement needs to be adapted to satisfy that. A detailed discussion is provided in Appendix A.2.2.

**3.2.2 Spin.** It is straightforward to append  $\mathbb{O}_{\text{suc}}$  by a check of whether the obtained state is in the correct spin state. Suppose we have access to an implementation of the time-evolution of the (normalized) spin operator  $S^2$ . Then, one step of quantum phase estimation allows storing in an ancilla qubit whether there is a singlet (0) or triplet (1) state. Augmentation to higher-order spin states (such as doublets, quadruplets, etc.) follows simply by augmenting the ancilla register of the QPE circuit to represent the appropriate spin numbers. Based on this result, we can accept or reject according to the obtained outcome and thus effectively have a scheme that projects into the desired spin state. Therefore, this yields an implementation of a projection into the correct spin state; this requires nontrivial knowledge about the expected spin state. One can draw from ideas in ref. 104 to deal with more complicated ensembles of spin states. A detailed procedure in case of lack of knowledge or intuition about the expected spin state is subject to further research.

We remark that construction of the  $S^2$  operator is simple given an encoding as in eqn (13), as then it simply translates to its form as Pauli operators acting on the spin degree of freedom.

## 4. Conclusion and outlook

We have provided an algorithmic framework that, in principle, can solve a broad set of chemically relevant problems using inherently efficient building blocks. To that end, we considered that, while general ground states are hard to obtain, we may assume we are able to prepare atomic ones as a single-cost effort that can enable a building block library. A scattering process, implemented by simulating dynamics and boosted by artificial potentials, can produce a molecular input state for a subsequent dynamical simulation, which is then followed by the measurement of dynamical quantities. Preparing molecular states *via* merger-association is a particularly promising candidate here, which we discussed in detail. We provide examples of applications from, *e.g.*, spectroscopy, photochemistry, and beyond.

Future work remains to build upon this approach and provide a more general and detailed analysis of merger-



association and numerical experiments of this approach, *e.g.*, to gauge the feasibility of the procedure with respect to constant factors and to investigate the more precise costs arising from choosing specific problem instances and methodologies. Another interesting extension would be the inclusion of other modeling tools used in classical simulation, such as the Nosé–Hoover thermostat.<sup>105</sup> Additionally, an interesting avenue would be to consider additional classical dynamical simulation such as more molecular dynamics (*e.g.*, studied in the context of protein modeling in ref. 106), in particular in the case when conservation laws (such as energy conservation) allow direct mappings to Hamiltonian simulation.<sup>107</sup> One can also think of incorporating different notions of how to distinguish bonded states in molecules, *e.g.*, grounding on the notion of molecular structure in ref. 108.

A common focal point in quantum chemistry is finding ground state energies, a problem known to be QMA-complete for local Hamiltonians like those seen in molecular systems. We should not restrict ourselves to that perspective, which focuses on a problem known to be hard, and looks for paths that allow using dynamics more directly. Additionally, for situations when the ground state is of interest, it may prove useful to give up the search for exact solutions to a hard problem and look at heuristics. The Hartree–Fock problem is known to be NP-hard yet practically efficient, thanks to approximations.<sup>44</sup> Attempts based on open-system dynamics such as in ref. 79 and 109 may be a promising path towards quantum heuristics for ground states. Nevertheless, we call upon what dynamical quantum simulation offers for chemistry.

## Author contributions

Conceptualization: AAG, JF, PS; methodology: PS, LBK, AA, JCGA, DA, MB, CG; investigation: PS, LBK, AA, JCGA, DA, MB; visualization: DA, PS, CG, MB; funding acquisition: AAG; project administration: PS; supervision: AAG, JF; writing – original draft: PS, LBK; writing – review & editing: PS, AAG, AA, LBK, JCGA, DA, MB, CG, JF.

## Conflicts of interest

There are no conflicts to declare.

## Data availability

No software or code has been included, and no new data were generated or analysed as part of this study. All results are part of the document.

## A. Appendices

### A.1. Scattering molecules *via* mergo-association

This section includes more details on merging molecules by mergo-association as described in Section 3.1.

**A.1.1 Outline of quantum computational encoding.** First, we describe the quantum computational encoding of the approach. A grid  $G$  of  $N$  points labels an integer lattice, where

for convenience, we assume that  $N$  has an integer cubic root so that  $\exists m \in \mathbb{N}, m^3 = N$ ,

$$G = \left[ -\frac{N^{1/3} - 1}{2}, \frac{N^{1/3} - 1}{2} \right]^3. \quad (\text{A1})$$

The associated volume box  $\mathcal{Q} := \left[ -\frac{L}{2}, \frac{L}{2} \right]^3$ . Exemplarity for dihydrogen with  $N_{\text{el}} = 2$ ,  $N_{\text{nuc}} = 2$ , and an estimated volume of  $|\mathcal{Q}| = L^3 = (10a_0)^3$  is expected to suffice, where  $a_0$  is the Bohr radius. With a grid spacing of  $\sim 0.1a_0$ , this volume would lead to  $N \sim 10^6$ . Then, upon defining the size of the simulation box, it is straightforward to map every  $p \in G$  to a coordinate,  $r_p \in \mathcal{Q}$  for electrons and  $R_p \in \mathcal{Q}$  for nuclei.

The following is needed to describe the desired dynamics: Hamiltonians describing the isolated hydrogen atoms, interactions between the hydrogen atoms, and a ‘trap potential’ modeled by a harmonic oscillator potential. The system Hamiltonian without a trap is given by

$$H_{\text{sys}} = T_{\text{el}} + T_{\text{nuc}} + U_{\text{nuc-el}} + V_{\text{ee}} + V_{\text{nn}}. \quad (\text{A2})$$

As the initial state to the simulation, we assume access to a product state of the ground states of the hydrogen atom, respectively,

$$|\psi_{0,\text{H}}\rangle |\psi_{0,\text{H}}\rangle. \quad (\text{A3})$$

From ref. 30, we know the sub-normalization factors (which we will repeat further below) to encode these in real space, together with estimates for the cost of Hamiltonian simulation algorithms. The placement of the individual  $|\psi_{0,\text{H}}\rangle$  should follow a guess for the desired molecular distance so that the relaxation through the trap potential induces a more accurate placement for the bonded state.

The harmonic trap potential from ref. 76 and 77 is

$$V^{\text{trap}}(R_1, R_2) = V_1^{\text{trap}}(R_1) + V_2^{\text{trap}}(R_2), \quad (\text{A4})$$

which confines the nuclear motion per atom  $j$ :

$$V_j^{\text{trap}}(R_j) = \frac{m_j}{2} (R_j - R_{0,j})^T \omega_j^2 (R_j - R_{0,j}). \quad (\text{A5})$$

We may neglect electronic mass and coordinates for the trap parameters given the large discrepancy between electronic and nuclear mass; the confinement is a heuristic approach nonetheless. The positions  $R_{0,j}$  define the center of the two traps. For single-particle nuclei, such as dihydrogen, the determination of trap centers is straightforward; in contrast, for larger systems, center-of-mass and relative coordinates need to be considered.<sup>77</sup> The frequencies  $\omega_j^2$ , in general, make up positive, diagonal matrices; the expression above describes the strength of the trap, which may not be spherically symmetric. For the sake of our discussion here, we will assume the trap to be isotropic, and then the trap frequency can be described by a scalar  $\omega_j$ . The role of these frequencies and relevant energy scales will be discussed further below in Appendix A.1.2.



Next, we discuss a quantum implementation of the trap potential. We reformulate eqn (A5) to

$$V^{\text{trap}}(R_1, \dots, R_{N_{\text{nuc}}}) = \sum_{j \in \text{nuclei}} \frac{m_j}{2} \sum_{p \in G} \sum_{w \in \{x, y, z\}} \omega_{j,w}^2 (R_{p_j,w} - R_{0,w})^2 |p\rangle\langle p|_j, \quad (\text{A6})$$

acting on the position labeled by  $p$ . The state we encode is a superposition over  $|p\rangle = |p_1\rangle \dots |p_{N_{\text{el}}+N_{\text{nuc}}}\rangle$ , which hold grid labels for each particle in a register, plus one extra qubit per particle for spin if desired. Synonymously in this work, we use typewriter-font to denote the grid labels,  $|r_1\rangle \dots |r_{N_{\text{el}}}\rangle |R_1\rangle \dots |R_{N_{\text{nuc}}}\rangle$ . A state called  $|r_j\rangle$  or  $|R_k\rangle$  will hold the explicit coordinate information. This requires  $\log(N)$  qubits per particle to represent its position on the grid. For dihydrogen, overall  $N_{\text{nuc}} = N_{\text{el}} = 2$ , so we need  $4 \times \log_2(\text{number of grid-points})$  qubits to encode the grid; with the estimated  $10^6$  grid points from above, that makes roughly 80 qubits for the  $|p\rangle$  register.

The procedure that is simulated is described using scheduling functions as mentioned above in eqn (4), so that  $f(s)$  introduces the inter-atomic Coulomb interactions and  $g(s)$  guides the strength of the trap. As a consequence, implementing this scattering follows evolution across  $\mathcal{T} e^{-i \int_0^{s_1} H(s) ds}$ , with  $H(s)$  from eqn (4). The truncated Dyson series algorithm in ref. 110 can serve as a way to implement this. A necessary requirement is then to have access to the potentials as a HAM-T oracle, which we outline for  $V^{\text{trap}}$ :

$$\sum_{s=0}^{n_t} \frac{V^{\text{trap}} \left( s \frac{s_1}{n_t} \right) \otimes |s\rangle\langle s|}{\alpha_{\text{trap}}}. \quad (\text{A7})$$

Using this, we can simulate over  $s$  across the scheduling functions so that the inter-species interactions remain and the trap has fully decayed, *i.e.*, until  $s_1$ . This operation can be constructed easily given the LCU outlined in eqn (A8) and (A10) together with oracular access to the scheduling functions from eqn (A9) and an appropriate superposition over  $|s\rangle$ ,  $s = 0, \dots, n_t$ . To construct superpositions over where  $n_t$  is not a power of two, see *e.g.* [ref. 30, Appendix J]. The equivalent construction is necessary for the interaction term together with  $f(s)$ ; see eqn (4), and encodings for the Coulomb-term in the literature [ref. 30, Appendix K]. There is an argument to make here about the (a) diabaticity of this evolution, which will be part of Appendix A.1.2.

We continue by outlining an LCU encoding of  $V^{\text{trap}}(s')$  for a specific  $0 \leq s' \leq s_1$ . Coordinate directions are denoted by  $\{x, y, z\} \sim \{0, 1, 2\}$  and we use the usual LCU notation of PREP and SEL operations so that  $(|0\rangle_{\text{ancilla}} \otimes \mathbb{1}_{\text{system}}) \text{PREP}^\dagger \cdot \text{SEL} \cdot \text{PREP}(|0\rangle_{\text{ancilla}} \otimes \mathbb{1}_{\text{system}})$  block-encodes the desired operator. The state we are acting on is assumed to be a superposition over  $|p\rangle|s'\rangle$ , with the state  $|p\rangle$  representing a spatial part and  $|s'\rangle$  the current step in the evolution of the merging. The PREP step involves preparing a superposition across nuclei and coordinate directions, weighted by the square root of the interaction strength at that point:

$$|0\rangle|p\rangle|s\rangle \mapsto \sum_{j: \text{nuclei}} \sum_{s=0}^{n_t} \sum_{w=0}^2 \frac{\sqrt{\frac{m_j}{2}} \omega_{j,w}}{\sum_{j',w'} \sqrt{\frac{m_{j'}}{2}} \omega_{j',w'}} |j\rangle|w\rangle|p\rangle|s\rangle \quad (\text{A8})$$

Next, we illustrate the SEL operation, where we omit amplitudes and explicit sums for clarity. The goal here is to compute state-dependent interaction strengths on the fly. Namely, we aim to prepare a register that contains  $(R_{p_{j,w}} - R_{0,w})^2$ , and another containing  $g(s)$  to multiply the latter, so that we can use *e.g.* ref. 111 to move information from the state into the amplitude. Consequently, access to oracles  $O_f$  and  $O_g$  that represent the scheduling is required,

$$O_{flg} : |0\rangle|s\rangle \rightarrow |flg(s)\rangle|s\rangle. \quad (\text{A9})$$

The oracular access to the scheduling function and the label-to-coordinate mapping may be realized with classical data encodings, *e.g.*, a QROM or other techniques.<sup>51,112,113</sup> Using a number of ancillas given by  $n_{\text{anc}}$  initialized in zero, and assuming a state that encodes the relative center of the trap  $R_0$ ,

$$\begin{aligned} & |j\rangle|w\rangle|s\rangle|0\rangle_a |p_1 \dots p_{N_{\text{part}}}\rangle|R_{0,w}\rangle \\ & \mapsto |j\rangle|w\rangle|s\rangle|0\rangle_a |g(s)\rangle|R_{p_{j,w}}\rangle|p\rangle|R_{0,w}\rangle \\ & \mapsto |j\rangle|w\rangle|s\rangle |g(s)(R_{p_{j,w}} - R_{0,w})^2\rangle |g(s)\rangle|R_{p_{j,w}}\rangle|p\rangle|R_{0,a,w}\rangle \\ & \mapsto |j\rangle|w\rangle|s\rangle|0\rangle_{a''} |g(s)|R_{p_j} - R_0|^2\rangle|p\rangle|R_{0,w}\rangle \end{aligned} \quad (\text{A10})$$

The size of the necessary ancillary register  $|0\rangle_a$  that is used to store intermediary and final results scales linearly in the bits of precision used for numerical representation (to store  $g(s)$ ,  $R_{p_{j,w}}$  and perform multiplication and sum-of-squares). As mentioned above, to finalize the SEL operation, the state information in  $|g(s)|R_{p_j} - R_0|^2\rangle$  needs to be moved into the amplitude. There are multiple ways to do so, such as the inequality-testing approach in ref. 111, or, the conceptually simplest would be controlled rotations, conditioned on  $|g(s)|R_{p_j} - R_0|^2\rangle$  and applied to the system register.

We refrain from analyzing exact Toffoli or gate counts as done in ref. 30, expecting that the addition of the trap potentials would not add a significant change here per single query to involved oracles and circuits as the appearing arithmetic operations are comparable to the usual potentials in the Hamiltonian. The factor that may make the trap potentials more costly and needs to be studied more closely is its sub-normalization factor. This factor captures the number of necessary queries for a successful block-encoding. We recall the sub-normalization factors from [ref. 30, Appendix K],

$$\alpha_T \in O\left(\frac{N_{\text{el}} N^{2/3}}{|\mathcal{Q}|^{2/3}}\right), \quad \alpha_V \in O\left(\frac{N_{\text{el}}^2 N^{1/3}}{|\mathcal{Q}|^{1/3}}\right), \quad \alpha_U \in O\left(\frac{N_{\text{el}}^2 N^{1/3}}{|\mathcal{Q}|^{1/3}}\right) \quad (\text{A11})$$

with  $T$  representing kinetic energy,  $V$  electron-electron interactions and  $U$  nuclear-electron interactions.



The block-encoding factor of the trap potential may be upper-bounded by  $\alpha_{\text{trap}} \leq \left\| \sum_{j \in [N_{\text{nuc}}]} V_j^{\text{trap}} \right\|$ . Then,

$\alpha_{\text{trap}} \leq N_{\text{nuc}} m_{\text{max}} \omega_{\text{max}}^2 \max_{r, r'} |r - r'|^2$ . In the present discretization,

$\max_{r, r' \in \Omega_{\text{trap}}} |r - r'|^2 \in O(|\Omega_{\text{trap}}|^{2/3})$ , where  $\Omega_{\text{trap}} \subseteq \Omega$  is the part of the domain the trap potential is defined on. Most of the mass of particles under such a potential will be incentivized to be close to the center; therefore, an upper bound that puts the maximum mass at a maximum distance will not be very tight in most scenarios, and a smaller  $\alpha_{\text{trap}}$  may be sufficient. One could further think about designing the shape of the trap potential to be decaying for larger distances to reduce the encoding cost or oscillatory as in optical tweezers, such that the maximum amplitude is always within reach – we leave concrete specifications up to future work. The maximum nuclear mass  $m_{\text{max}} = \max_{1 \leq j \leq N_{\text{nuc}}} m_j$  can be treated as a constant factor attached to the nuclei. Then, we have that

$$\alpha_{\text{trap}} \in O(\omega_{\text{max}}^2 N_{\text{nuc}} |\Omega_{\text{trap}}|^{2/3}). \quad (\text{A12})$$

Up to the frequency  $\omega_{\text{max}}$ , the sub-normalizations of the present procedure in eqn (A11) compared to the trap potential inhibit a few key differences. Because the strength of the trap depends on a maximum distance, there is no notion of ‘grid density’ in the cost. Therefore, the factors in eqn (A11) increase when a lower target accuracy is desired. The trap encoding is not directly dependent on accuracy, though it is more pronounced on system size. The system size of the trap may be upper-bounded with the overall system studied, as present at the last scattering stage before ‘evolution’ in Fig. 3.

Therefore, it is key to understand the role of  $\omega_{\text{max}}$  in the scattering and whether there is a way to relate this quantity to the system size.

**A.1.2 Probability of success of mergo-association.** To address the choice of trap frequency, we consider that the probability of a diabatic transition into  $|a\rangle$  through the mergo-association from ref. 76 and 77 can be approximated by the Landau–Zener rule and is proportional to

$$p_{\text{LZ}} \approx \exp \left( -\frac{2\pi}{\hbar} \frac{\omega_{\text{eff}}^2}{\left| \frac{\partial}{\partial s} (E_{\text{mol}} - E_{\text{atom}}) \right|} \right) \quad (\text{A13})$$

Here,  $\omega_{\text{eff}}$  is an effective frequency we explain further below that models the strength of the potential. Then, we express the rate of change of the energy,  $\frac{\partial}{\partial s} (E_{\text{mol}} - \partial E_{\text{atom}})$ , as a consequence of the chain rule, as  $|\partial E_{\text{mol}} - \partial E_{\text{atom}}|v$ . Here,  $\partial E$  are energy derivatives near the initial configuration with respect to the internuclear distance and  $v$ , the speed of the evolution along the scheduling function.

$$p_{\text{LZ}} \approx \exp \left( -\frac{2\pi}{\hbar} \frac{\omega_{\text{eff}}^2}{|\partial E_{\text{mol}} - \partial E_{\text{atom}}|v} \right) \quad (\text{A14})$$

Thus, the success probability of the mergo-association, so that the merging happens without transitioning into a higher vibrational state of the trapped system as desired, is

$$p_{\text{suc}} \sim 1 - p_{\text{LZ}}. \quad (\text{A15})$$

As we pointed out in the previous section, this means that the type of evolution we call successful here is adiabatic. It is important to note that this is not adiabatic ground-state preparation, which is explicitly beyond the scope of our work. Following another discussion in the preceding section, the threshold to success here considers molecular bound states, so the ‘upper limit’ for adiabaticity would be the highest-lying vibrational bound state. One way or another, a practical implementation will need to consider this. This still poses limits on the speed of evolution (given by  $v$  in eqn (A14)). Then,  $v = \max \left\{ \left| \frac{dz}{df} \frac{d}{ds} f(s^*) \right|, \left| \frac{dz}{dg} \frac{d}{ds} g(s^*) \right| \right\}$ , where  $z$  is the considered internuclear distance.

Looking at eqn (A14), we can identify that the stronger (“steeper”) the potential as described by the effective frequency, the higher the success, and the faster the merging proceeds, the less likely. This conclusion is consistent with intuition.

We continue by breaking down the quantities that appear. Following ref. 76,  $\partial E_{\text{atom}} \approx 0$ , as it is reasonable to assume the energy surface for separated atoms is flat. Taking a harmonic oscillator approximation to the relative motion of the nuclei, [ref. 76, eqn (37)] obtains an approximation to the gradient,  $\partial E_{\text{mol}} \approx \mu \omega^2 z^*$ , where  $z^*$  is the point of contact of the avoided crossing between the two states. In terms of the scheduling functions that we envision, this corresponds to the state at  $s = s^*$ ,  $0 < s^* < s_0 < s_1$  (Fig. 5): briefly before both interaction and trap are ‘fully acting’. Assuming an isotropic potential, the spatial coordinate is approximated as  $z^* = \left( \frac{\hbar}{\mu \omega} \right)^{1/2} \left( 3 + \frac{\omega_a}{\omega} \right)^{1/2}$  so that [ref. 76, eqn (35)]

$$\partial E_{\text{mol}} \approx (\hbar \mu)^{1/2} \omega (3\omega + \omega_a)^{1/2}, \quad (\text{A16})$$

where  $\omega_a$  is defined as the frequency associated with the bonded vibrational excited state.

To estimate  $\omega_{\text{eff}}$ , we can also follow [ref. 76, eqn (52)]:

$$\begin{aligned} \omega_{\text{eff}}^2 &\approx \left( \frac{\langle a | (1 - |000\rangle\langle 000|) V^{\text{int}} | 000 \rangle}{1 - |\langle a | 000 \rangle|^2} \right)^2 \\ &\approx \langle a | V^{\text{int}} | 000 \rangle^2 = \frac{2\hbar^2}{\sqrt{\pi}} \omega_a^{1/2} \omega^{3/2} \exp \left[ -\frac{1}{2} \left( 3 + \frac{\omega_a}{\omega} \right) \right] \end{aligned} \quad (\text{A17})$$

States “000” describe the vibrational ground-state, “a” an excited vibrational state, and  $\omega$  is the frequency of the harmonic oscillator trap. The approximation to only look at the transition element itself is investigated in [ref. 76, Fig. 6] – for a merely qualitative argument, this is a sufficient choice. The choice for  $\omega_a$  in our case is an estimate for an upper threshold of the



respective bonded vibrational energy subspace. Introducing a ‘binding energy’-quantity  $E_a = \hbar\omega_a$ , we can rewrite  $\omega_{\text{eff}}^2$  to be approximately proportional to [ref. 76, eqn (54)],

$$\omega_{\text{eff}}^2 \approx \omega^{3/2} (E_a/\hbar)^{1/2} \exp\left(-\frac{E_a}{\hbar\omega}\right) = \omega^2 \tilde{E}_a^{1/2} \exp(-\tilde{E}_a). \quad (\text{A18})$$

In the latter equation, we introduce the relative binding energy  $\tilde{E}_a = E_a/(\hbar\omega)$ , defined in relation to the trap energy. Then,

$$\begin{aligned} p_{\text{LZ}} &\approx \exp\left(-4\left(\frac{\hbar\pi}{\mu}\right)^{1/2} \left(\frac{\omega\omega_a}{3\omega + \omega_a}\right)^{1/2} \frac{e^{-\frac{1}{2}\left(3 + \frac{\omega_a}{\omega}\right)}}{v}\right) \\ &= \exp\left(-4\left(\frac{\pi}{\mu}\right)^{1/2} \left(\frac{\hbar\omega E_a}{3\hbar\omega + E_a}\right)^{1/2} \frac{\exp\left(-\frac{1}{2}\frac{E_a}{\hbar\omega} - \frac{3}{2}\right)}{v}\right) \\ &= \exp\left(-4\left(\frac{\pi}{\mu}\right)^{1/2} \left(\hbar\omega \frac{\tilde{E}_a}{3 + \tilde{E}_a}\right)^{1/2} \frac{\exp\left(-\frac{1}{2}\tilde{E}_a - \frac{3}{2}\right)}{v}\right). \end{aligned} \quad (\text{A19})$$

## A.2. Construction of measurement oracles for certifying reactions

For this oracle construction, we assume a first-quantized real space implementation. The choice of representation in first-quantization on real space makes the construction of the following oracle easier; of course, plane-wave or plane-wave dual bases<sup>52</sup> are possible too and can be translated to with appropriate transformations. The encoded quantum states are linear combinations of wavefunctions with fermionic (anti-symmetric) and bosonic (symmetric) parts. Within such a combination, a single component looks like a tensor product over electronic  $\{r\} \sim \{r_j\}_{j=1}^{N_{\text{el}}}$  and nuclear  $\{R\} \sim \{R_k\}_{k=1}^{N_{\text{nuc}}}$  grid labels (represented on the grid eqn (A1)),

$$|\psi(\{r, R\})\rangle = |r_1 \dots r_{N_{\text{el}}}\rangle_{\text{el}} |r_1 \dots r_{N_{\text{nuc}}}\rangle_{\text{nuc}}; \quad (\text{B1})$$

We further remember that the above collapsed notation that also contains spin information in an extra qubit, *i.e.*,  $|r_j\rangle \sim |r_j\rangle \otimes |\sigma_j\rangle$  with  $\sigma_j \in \{\uparrow, \downarrow\}$ , and analogous for nuclei if desired. Then, a density matrix formulation of linear combinations of such a state is

$$\tilde{\mathcal{C}}_{\text{geom},\Delta}(\{r, R\}, I) = \begin{cases} 1, & \|R_j - R_k\| \leq \Delta_{jk} \quad \forall (j, k) \in I \subseteq [N_{\text{nuc}}] \times [N_{\text{nuc}}] \\ 0, & \text{else.} \end{cases} \quad (\text{B4})$$

$$\rho = \sum_{\{r, R\}, \{r', R'\}} \rho_{\{r, R\}, \{r', R'\}} |\psi(\{r, R\})\rangle \langle \psi(\{r', R'\})|, \quad (\text{B2})$$

with the entries of the density matrix in the position basis,  $\rho_{\{r, R\}, \{r', R'\}} \in \mathbb{C}$ . The goal of this section is to construct an oracle that can distinguish the ‘good’, reacted components of such a state from the ‘bad’, non-reacted components. Given a state representing two molecular fragments that have evolved for some time, we want to distinguish the parts of the wavefunction that correspond to the fragments having reacted to form a single molecule from those where they are still separate unbound fragments.

As we consider linear transformations, it suffices to study the effect on the individual components from eqn (B1) to draw conclusions for a general state in eqn (B2). Therefore, we first focus on a single position-basis state  $|\psi(\{r, R\})\rangle$  in the upcoming Appendix A.2.1 before we consider exchange symmetry in Appendix A.2.2 and more general superpositions in Appendix A.2.3.

**A.2.1 Geometrical criteria.** We now discuss an approach to test whether or not a molecular bond has formed for a state of the form  $|\psi(\{r, R\})\rangle$ . When atomic nuclei can be precisely located, a reasonable description for the molecular structure can be given in terms of a set of inter-nuclear distances  $\{R_{jk}^{\text{react}}\}$ . Using this information, a corresponding criterion  $\mathcal{C}$  for whether a given configuration corresponds to a desired molecule can then be constructed,

$$\mathcal{C}_{\text{geom},\varepsilon}(\{r, R\}) = \begin{cases} 1, & \left| \|R_j - R_k\| - R_{jk}^{\text{react}} \right| \leq \varepsilon \quad \forall j, k \\ 0, & \text{else.} \end{cases} \quad (\text{B3})$$

Note that evaluating this criterion is classically efficient with respect to the number of bits devoted to the constituent quantities, scaling no worse than  $O(N_{\text{nuc}}^2)$  in the number of nuclei and no worse than quadratically in the bit-precision due to a product and a square root in the evaluation of the norm. In other words, evaluating this function is classically efficient in the size of the molecule as long as the bit-precision only grows polynomially, corresponding to the requirement that the number of grid points in the simulation should not grow faster than exponentially in the involved particle number.

In some cases, it may be possible to devise more accurate criteria, such as by comparing the angles between inter-nuclear distances to desired bond angles. On the other hand, there may be cases where less is known about the target structure. Then, the best that one can hope for is to check that the two fragments are at least spatially adjacent, corresponding to the looser requirement

for some suitable  $\Delta_{jk} > 0$  and a set of locations of interest  $I$ , in which we only check a subset of the encoded locations

regarding proximity. There are many ways this type of geometric requirement could be mixed and matched, *e.g.*, utilizing different levels of knowledge or different precisions  $\varepsilon$  for



different groups of atoms. Either way, common to these approaches is the fact that the classical evaluation of the criterion is efficient under the mild restrictions to the grid resolution outlined above.

A result of this classical efficiency is that a corresponding unitary to perform the computation can also be implemented efficiently on a quantum computer using reversible logic. Specifically, for a given efficient criterion  $\mathcal{C}$  of the form considered here, the following unitary can be implemented:

$$U_{\mathcal{C}} |p_{N_{\text{el}}} \dots p_{N_{\text{el}}+N_{\text{nuc}}} \rangle_{\text{nuc}} |0\rangle = |p_{N_{\text{el}}} \dots p_{N_{\text{el}}+N_{\text{nuc}}} \rangle_{\text{nuc}} |\mathcal{C}(\{r, R\})\rangle. \quad (\text{B5})$$

Thus, extracting information about the criterion into an ancilla qubit for easy access is possible. Using a measurement of this ancilla then allows for a projection onto the good ( $\mathcal{C} = 1$ ) subspace, heralded by an outcome “1” of the measurement. However, the outcome “0” similarly fully projects the state into the bad subspace ( $\mathcal{C} = 0$ ). Furthermore, care needs to be taken to avoid either kind of projection destroying physically important symmetries of the state, including fermionic and bosonic particle-exchange symmetries. Mitigating each of these potential problems will be the topics of the following two sections.

**A.2.2 Preserving exchange symmetry throughout measurement.** So far, we have been primarily considering a single component of the wavefunction,  $|\psi(\{r, R\})\rangle$ , in which the positions of all particles are fully specified. However, we typically deal with a collection of nuclei and electrons modeled through wavefunctions that need to be of the correct symmetry with respect to the exchange of particles. We focus on preserving exchange symmetry compared to other symmetries present in the system that may also be broken in the real physical processes. Hence, the goal is to evaluate reaction criteria in a manner that does not violate the symmetries of present states.

**A.2.2.1 Symmetrization of the wavefunction.** Recall that the wavefunctions we encode are given as superpositions of position-labelling basis states,

$$(|r_1\rangle |r_2\rangle \dots |r_{N_{\text{el}}}\rangle) (|R_1\rangle |R_2\rangle \dots |R_{N_{\text{nuc}}}\rangle) \quad (\text{B6})$$

We repeat that by  $|r_j\rangle, |R_k\rangle$  we mean the labels for grid points and spin rather than the resolved coordinates and choose this notation as it is more illustrative in the context of (anti)symmetrization of the wavefunctions. For fermionic (indistinguishable) particles, the wavefunction needs to be anti-symmetrized with respect to the indistinguishable degrees of freedom. Formally, this can be done by applying the anti-symmetrization operator  $\mathcal{A}$ . Using the electronic degrees of freedom as an example, and with  $\sigma$  a permutation from the permutation group over  $N_{\text{el}}$  elements,  $S_{N_{\text{el}}}$ :

$$\mathcal{A} (|r_1\rangle |r_2\rangle \dots |r_{N_{\text{el}}}\rangle) = \frac{1}{\sqrt{N_{\text{el}}!}} \sum_{\sigma \in S_{N_{\text{el}}}} \text{sgn}(\sigma) |r_{\sigma(1)}\rangle |r_{\sigma(2)}\rangle \dots |r_{\sigma(N_{\text{el}})}\rangle \quad (\text{B7})$$

Similarly, bosonic degrees of freedom will be symmetric with respect to exchange, described by symmetrization  $\mathcal{S}$  of the general form

$$\mathcal{S} (|R_1\rangle |R_2\rangle \dots |R_{N_{\text{nuc}}}\rangle) = \frac{1}{\sqrt{k!}} \sum_{\sigma \in S_{N_{\text{nuc}}}} |R_{\sigma(1)}\rangle |R_{\sigma(2)}\rangle \dots |R_{\sigma(N_{\text{nuc}})}\rangle \quad (\text{B8})$$

In general, a set of nuclei to be modeled as indistinguishable will correspond to a subset of nuclear registers to be (anti)symmetrized. Thus, the desired symmetry properties are characterized by sets of registers to be symmetrized,  $\{B_i\}_{i=1}^{N_{\text{B}}}$ , and sets of registers to be anti-symmetrized,  $\{F_i\}_{i=1}^{N_{\text{F}}}$ . Denoting which basis states an operator acts on by a subscript, one gets the following anti-symmetrized state:

$$\mathcal{A} (|r_1\rangle |r_2\rangle \dots |r_{N_{\text{el}}}\rangle) \otimes \left( \prod_{i=1}^{N_{\text{F}}} \mathcal{A}_{F_i} \prod_{j=1}^{N_{\text{B}}} \mathcal{S}_{B_j} \right) (|R_1\rangle |R_2\rangle \dots |R_{N_{\text{nuc}}}\rangle). \quad (\text{B9})$$

**A.2.2.2 Illustrative example:  $\text{H}_2\text{O}_2$ .** To elucidate our notation, consider the molecule  $\text{H}_2\text{O}_2$ , and assume that the non-symmetrized positions are encoded as

$$|R_{\text{O},1}\rangle |R_{\text{O},2}\rangle |R_{\text{H},1}\rangle |R_{\text{H},2}\rangle. \quad (\text{B10})$$

In this case, the oxygen nuclei are spin-0 nuclei, meaning they should be symmetrized. On the other hand, the hydrogen nuclei are spin-1/2 particles, meaning they should be anti-symmetrized. We can represent this by the following two sets:

$$B_1 = \{1, 2\}, \quad F_1 = \{3, 4\}. \quad (\text{B11})$$

The first tells us to symmetrize the oxygen registers (1 and 2), and the second tells us to anti-symmetrize the hydrogen registers (3 and 4). The state after symmetrization is therefore

$$\begin{aligned} \mathcal{A}_{F_1} \mathcal{S}_{B_1} (|R_{\text{O},1}\rangle |R_{\text{O},2}\rangle |R_{\text{H},1}\rangle |R_{\text{H},2}\rangle) = \frac{1}{2} ( & |R_{\text{O},1}\rangle |R_{\text{O},2}\rangle \\ & + |R_{\text{O},2}\rangle |R_{\text{O},1}\rangle ) \otimes ( |R_{\text{H},1}\rangle |R_{\text{H},2}\rangle - |R_{\text{H},2}\rangle |R_{\text{H},1}\rangle ). \end{aligned} \quad (\text{B12})$$

The encoded molecular state will generally be a superposition of states of form in eqn (B9), *i.e.*, the tensor products of an antisymmetric wavefunction, a symmetric wavefunction, and possibly a non-symmetrized component. Then, for the remainder, it suffices to discuss the oracle's effect on one term of the superposition. To see what symmetry implies for the construction of oracles, consider again the example of  $\text{H}_2\text{O}_2$ . Looking at the form of the wavefunction in eqn (B10) and the equilibrium geometry of the molecule, one might be tempted to define a criterion of the form

$$\begin{aligned} \mathcal{C}_\varepsilon (\{R\}) = ( & (\|R_1 - R_3\| - 95 \text{ pm}) < \varepsilon ) \wedge (\|R_2 - R_4\| - 95 \text{ pm}) < \varepsilon \\ & \wedge (\|R_1 - R_2\| - 147 \text{ pm}) < \varepsilon \end{aligned} \quad (\text{B13})$$

in suitable numerical units and with subscripts denoting the ordering of the registers. After all, if we achieve the equilibrium



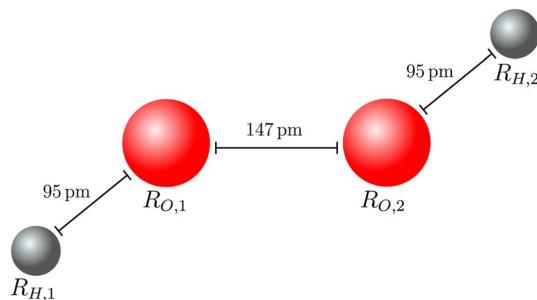


Fig. 6 Equilibrium configuration of a  $\text{H}_2\text{O}_2$  molecule, with oxygen nuclei marked in red and hydrogen nuclei marked in grey.

configuration (see Fig. 6), an oracle based on this criterion will, by construction, certify the non-symmetrized state as reacted:

$$\begin{aligned} U_C |R_{O,1}\rangle|R_{O,2}\rangle|R_{H,1}\rangle|R_{H,2}\rangle|0\rangle &= |R_{O,1}\rangle|R_{O,2}\rangle|R_{H,1}\rangle|R_{H,2}\rangle \\ &\otimes |C_\epsilon(\{R_{O,1}, R_{O,2}, R_{H,1}, R_{H,2}\})\rangle \\ &= |R_{O,1}\rangle|R_{O,2}\rangle|R_{H,1}\rangle|R_{H,2}\rangle|1\rangle \end{aligned} \quad (\text{B14})$$

However, the same cannot be said for states where the positions are permuted. For instance, swapping the oxygen positions, the evaluation of the oracle reads

$$\begin{aligned} C_\epsilon(\{R_{O,2}, R_{O,1}, R_{H,1}, R_{H,2}\}) &= (\|R_{O,2} - R_{H,1}\| - 95 \text{ pm} < \epsilon) \\ &\wedge (\|R_{O,1} - R_{H,2}\| - 95 \text{ pm} < \epsilon) \\ &\wedge (\|R_{O,2} - R_{O,1}\| - 147 \text{ pm} < \epsilon) \end{aligned} \quad (\text{B15})$$

and consulting the geometric configuration in Fig. 6 shows that this Boolean formula evaluates to zero. Thus,

$$U_C |R_{O,2}\rangle|R_{O,1}\rangle|R_{H,1}\rangle|R_{H,2}\rangle|0\rangle = |R_{O,2}\rangle|R_{O,1}\rangle|R_{H,1}\rangle|R_{H,2}\rangle|0\rangle. \quad (\text{B16})$$

Considering the full symmetrized wavefunction, one finds

$$\begin{aligned} U_C \left[ \frac{1}{2} (|R_{O,1}\rangle|R_{O,2}\rangle + |R_{O,2}\rangle|R_{O,1}\rangle) \otimes (|R_{H,1}\rangle|R_{H,2}\rangle - |R_{H,2}\rangle|R_{H,1}\rangle) \right] |0\rangle \\ = \frac{1}{2} [|R_{O,1}\rangle|R_{O,2}\rangle|R_{H,1}\rangle|R_{H,2}\rangle - |R_{O,1}\rangle|R_{O,2}\rangle|R_{H,2}\rangle|R_{H,1}\rangle]|1\rangle \\ + \frac{1}{2} [|R_{O,2}\rangle|R_{O,1}\rangle|R_{H,1}\rangle|R_{H,2}\rangle - |R_{O,1}\rangle|R_{O,2}\rangle|R_{H,2}\rangle|R_{H,1}\rangle]|0\rangle \end{aligned} \quad (\text{B17})$$

Despite all four components corresponding to the same molecule (just with particle labeling altered), only two of them have been certified by the oracle. Furthermore, if the ancilla is measured, the resulting projected state is no longer of the symmetric form described in eqn (B9), and direct inspection shows that it is neither symmetric nor anti-symmetric upon permutations among the nuclear coordinates.

One possible response to this problem might be simply reapplying an (anti-)symmetrization step whenever a projection occurs. In principle, this operation scales linearly (up to logarithmic factors) in the system size,<sup>114,115</sup> thus, it would not change the polynomial runtime of the algorithm. However, existing algorithms require a particular structure on the input state, a requirement that is linked to unitarity<sup>114</sup> and thus likely challenging to lift. Furthermore, even with a hypothetical generalization of the algorithm, the present work proposes a quasi-continuous weak-measurement strategy for monitoring the reaction criteria. Since this entails frequent (weak) measurements during the simulation run, it is likely to lead to a high total number of measurements, scaling extensively with the simulation time. Restoring symmetry after each measurement could, therefore, add significant complexity to the algorithm, even if a fast and applicable symmetrization algorithm is available. Therefore, we next discuss a more comprehensive approach to address these shortcomings.

**A.2.2.3 Symmetry of states.** To build certification oracles that are valid under the correct exchange symmetry, we will first formalize the symmetrization requirement for the states of the system. Consider first a single set of identical particles, represented by a set of indices  $\mathcal{X}$ . The group of permutations of this set is characterized by the symmetric group,  $S_{|\mathcal{X}|}$ , and for each permutation  $\sigma \in S_{|\mathcal{X}|}$ , we can define a unitary consisting of SWAP gates that implements the permutation of our registers:

$$U_\sigma |R_1\rangle|R_2\rangle\dots|R_{|\mathcal{X}|}\rangle = |R_{\sigma(1)}\rangle|R_{\sigma(2)}\rangle\dots|R_{\sigma(|\mathcal{X}|)}\rangle, \quad (\text{B18})$$

or, more compactly

$$U_\sigma |\psi\{r, R\}\rangle = |\psi[\sigma\{r, R\}]\rangle \quad (\text{B19})$$

$\sigma\{r, R\}$  denotes the (ordered) set of positions after the permutation has been applied. In a group-theory language, this corresponds to a representation of the group  $S_{|\mathcal{X}|}$  onto the space of states. In terms of these constructions, the symmetrization requirement is

$$U_\sigma |\psi\rangle = \begin{cases} \text{sgn}(\sigma) |\psi\rangle & \text{Fermions} \\ |\psi\rangle & \text{Bosons} \end{cases} \quad \forall \sigma \in S_{|\mathcal{X}|} \quad (\text{B20})$$

where  $\text{sgn}(\sigma)$  is the sign of the permutation; the expression depends on whether the identical particles represented by the indices in  $\mathcal{X}$  are fermions or bosons.

Generalizing this construction to multiple sets of bosonic and fermionic particles, the symmetry group takes the form of a product group of the individual bosonic and fermionic symmetries,

$$S = S_B \otimes S_F \quad (B21)$$

$$S_B = \otimes_{j=1}^{N_B} S_{|B_j|}, \quad S_F = \otimes_{i=1}^{N_F} S_{|F_i|}, \quad (B22)$$

with the group representation taking a similar product form. Note that the tensor product structure implies that every element in  $S$  can be decomposed into the product of a fermionic and bosonic component,  $\sigma = \sigma_B \otimes \sigma_F$ . Using this notation, the symmetry requirement becomes

$$U_\sigma |\psi\rangle = \text{sgn}(\sigma_F) |\psi\rangle \quad \forall \sigma \in S. \quad (B23)$$

**A.2.2.4 Symmetry of oracles.** Using the notation introduced above, we can now state the requirement that a reaction criterion  $\mathcal{C}$  needs to fulfill to respect the symmetry defined in the previous section. Specifically, we will say that a criterion  $\mathcal{C}$  respects symmetry when a projective measurement of  $\mathcal{C}$  on a symmetrized state produces post-measurement states that are still correctly symmetrized. Below, we will show that this property holds if and only if

$$\mathcal{C}(\sigma[\{r, R\}]) = \mathcal{C}(\{r, R\}) \quad \forall \sigma \in S, \quad \forall \{r, R\}. \quad (B24)$$

To see that this is sufficient criteria, consider the projection operators related to the readout of  $\mathcal{C}$  using an ancilla qubit:

$$\Pi_\pm^C = \frac{1}{2} U_C^\dagger (\mathbb{1} \otimes (\mathbb{1} \pm Z)_{\text{anc}}) U_C \quad (B25)$$

$$\tilde{\Pi}_\pm^C = (\mathbb{1} \otimes \langle 0|_{\text{anc}}) \Pi_\pm^C (\mathbb{1} \otimes |0\rangle_{\text{anc}}) \quad (B26)$$

Looking at the constituent operators, we can note that

$$\begin{aligned} U_C (U_\sigma \otimes \mathbb{1}) |\psi(\{r, R\})\rangle |0\rangle &= U_C |\psi(\sigma[\{r, R\}])\rangle |0\rangle \\ &= |\psi(\sigma[\{r, R\}])\rangle |\mathcal{C}(\sigma[\{r, R\}])\rangle \end{aligned} \quad (B27)$$

$$\begin{aligned} (U_\sigma \otimes \mathbb{1}) U_C |\psi(\{r, R\})\rangle |0\rangle &= U_\sigma |\psi(\{r, R\})\rangle |\mathcal{C}(\{r, R\})\rangle \\ &= |\psi(\sigma[\{r, R\}])\rangle |\mathcal{C}(\{r, R\})\rangle \end{aligned} \quad (B28)$$

Thus, if the relation in eqn (B24) holds, it implies that

$$[U_C, (U_\sigma \otimes \mathbb{1})] = \left[ U_C^\dagger, (U_\sigma \otimes \mathbb{1}) \right] = 0 \quad \forall \sigma \in S \quad (B29)$$

Therefore, the projectors also commute with the permutations,

$$[\Pi_\pm^C, (U_\sigma \otimes \mathbb{1})] = [\tilde{\Pi}_\pm^C, U_\sigma] = 0 \quad \forall \sigma \in S \quad (B30)$$

which in turn means the post-projection states of (anti)symmetrized states are also (anti)symmetrized

$$\begin{aligned} U_\sigma \left( \tilde{\Pi}_\pm^C |\psi\rangle \right) &= \tilde{\Pi}_\pm^C U_\sigma |\psi\rangle \\ &= \tilde{\Pi}_\pm^C \text{sgn}(\sigma_F) |\psi\rangle \\ &= \text{sgn}(\sigma_F) \left( \tilde{\Pi}_\pm^C |\psi\rangle \right). \end{aligned} \quad (B31)$$

Hence, adhering to eqn (B24) is sufficient for the measurements of the criteria to preserve (anti)symmetry. To see that it is also a necessary condition, assume that a configuration  $\{r_0, R_0\}$  and permutation  $\sigma_0$  exist so that

$$\mathcal{C}(\sigma[\{r_0, R_0\}]) \neq \mathcal{C}(\{r_0, R_0\}). \quad (B32)$$

However, suppose for contradiction that the projectors related to  $\mathcal{C}$  still map (anti)symmetrized states to (anti)symmetrized states. Under this assumption, consider a state  $|\psi\rangle$  formed by the (anti)symmetrization of  $|\psi(\{r_0, R_0\})\rangle$ . By definition, this state has a non-zero overlap with  $|\psi(\{r_0, R_0\})\rangle$ . Writing the decomposition of the state in terms of positions,

$$|\psi\rangle = \sum_{r, R} \alpha_{\{r, R\}} |\psi(\{r, R\})\rangle, \quad (B33)$$

this is equivalent to the statement that  $|\alpha_{\{r_0, R_0\}}|^2 > 0$ .

Assume now without loss of generality that  $\mathcal{C}(\{r_0, R_0\}) = 0$ , and consider

$$U_{\sigma_0} \tilde{\Pi}_+^C |\psi\rangle = \text{sgn}(\sigma_{0,F}) \tilde{\Pi}_+^C |\psi\rangle \quad (B34)$$

which is fulfilled by the assumption that the projectors preserve (anti)symmetry. Since  $(\tilde{\Pi}_+^C)^2 = \tilde{\Pi}_+^C$ , this implies

$$\tilde{\Pi}_+^C U_{\sigma_0} \tilde{\Pi}_+^C |\psi\rangle = \text{sgn}(\sigma_{0,F}) \tilde{\Pi}_+^C |\psi\rangle. \quad (B35)$$

Therefore, the norms of the left-hand and right-hand sides above are also equal. Using the following sets

$$A = \{\{r, R\} : \mathcal{C}(\{r, R\}) = 0\} \quad (B36)$$

$$B_{\sigma_0} = \{\{r, R\} : \mathcal{C}(\sigma_0[\{r, R\}]) = 0\} \quad (B37)$$

this equality of norms can be written as

$$\sum_{\{r, R\} \in A \cap B_{\sigma_0}} |\alpha_{\{r, R\}}|^2 = \sum_{\{r, R\} \in A} |\alpha_{\{r, R\}}|^2 \quad (B38)$$

Note that basic set theory implies that  $A \cap B_{\sigma_0} \subseteq A$ , meaning every term in the left-hand sum is also present in the right-hand sum. Subtracting these common terms, this implies

$$\sum_{\{r, R\} \in A \setminus B_{\sigma_0}} |\alpha_{\{r, R\}}|^2 = 0 \quad (B39)$$

There can be no nonzero amplitude for any configurations in  $A \setminus B_{\sigma_0}$ . However, by assumption  $\mathcal{C}(\{r_0, R_0\}) = 0$ , meaning  $\{r_0, R_0\}$



ε A. Furthermore, eqn (B32) implies that  $C(\sigma_0[\{r, R\}]) \neq 0$ , meaning  $\{r_0, R_0\} \notin B_{\sigma_0}$ . Thus,  $\{r_0, R_0\} \in A \setminus B_{\sigma_0}$  and  $|\alpha_{\{r_0, R_0\}}|^2 > 0$ . We have, in other words, arrived at a contradiction. For a criterion  $C$  that does not fulfill eqn (B24), configurations necessarily exist where measuring  $C$  breaks symmetrization.

Note that all arguments above apply equally to other groups of spatial transformations under which the wave function should be invariant (up to a phase), for instance, SO(3). They also may be extended to other operators that commute with the Hamiltonian, such as total spin. In general, one should construct criteria that are maximally invariant under trivial transformations. One will otherwise discard components that correspond to desired configurations, yielding artificially low success probabilities (e.g., 50% in the case of fully formed  $\text{H}_2\text{O}_2$ ).

Then, the input state from above can be rewritten to account for the ‘successful’ and ‘unsuccessful’ subspaces as

$$\rho = \sum_{j,k \in A} \rho_{j,k} |\psi_j\rangle\langle\psi_k| + \sum_{j,k \in B} \rho_{j,k} |\psi_j\rangle\langle\psi_k| + \sum_{j \in A, k \in B} (\rho_{j,k} |\psi_j\rangle\langle\psi_k| + \rho_{k,j} |\psi_k\rangle\langle\psi_j|). \quad (\text{B42})$$

Below, we will go through each step outlined in the main text in Section 3.2.1, discussing motivation and relevant considerations.

**A.2.3.1 Step 1: first application of  $U_C$ .** This step aims to extract the information represented by the oracle into an ancilla for easy access. Using the definition of the sets  $A$  and  $B$ , the state resulting from this step can be explicitly written as

$$\begin{aligned} U_C(\rho \otimes |0\rangle\langle 0|) U_C^\dagger = & \sum_{j,k \in A} \rho_{j,k} |\psi_j\rangle\langle\psi_k| \otimes |1\rangle\langle 1| + \sum_{j,k \in B} \rho_{j,k} |\psi_j\rangle\langle\psi_k| \otimes |0\rangle\langle 0| \\ & + \sum_{j \in A, k \in B} (\rho_{j,k} |\psi_j\rangle\langle\psi_k| \otimes |1\rangle\langle 0| + \rho_{k,j} |\psi_k\rangle\langle\psi_j| \otimes |0\rangle\langle 1|). \end{aligned} \quad (\text{B43})$$

**A.2.3 Effect of oracle measurement for general superpositions.** In this section, we cover the effect of weak measurement of an oracle on a general mixed state. The goal is to elucidate the measurement procedure, the effects of the measurements, and the role of the parameter  $\delta$  describing the strength of the measurement.

We recall the form of a general input state,

$$\rho = \sum_{j,k} \rho_{j,k} |\psi(\{r, R\}_j)\rangle\langle\psi(\{r, R\}_k)| \quad (\text{B40})$$

as well as access to a classical criterion  $C$  and a unitary implementation  $U_C$ , as discussed in Section 1.2.1. Further, we recall the definition of the set of ‘successful states’ marked by  $A$  and ‘unsuccessful states’ marked by  $B$  from the main text as:

$$\begin{aligned} A &= \left\{ j : C(\{r, R\}_j) = 1 \right\} \\ B &= \left\{ j : C(\{r, R\}_j) = 0 \right\} \\ \langle\psi_j| &:= |\psi(\{r, R\}_j)\rangle \end{aligned} \quad (\text{B41})$$

At this stage, a measurement of the ancilla qubit would project the state onto either the first or the second term in this sum. In the case of projection onto the first term (i.e., a measurement outcome of ‘1’), this projection would correspond to a heralded projection onto the desired set of states, as identified by the oracle. However, such a measurement would come with a risk of a projection onto the space that does not correspond to success. Thus, a full measurement risks projecting the state out of the desired space, and repeated measurements would risk a Zeno-effect-like freezing of the dynamics in the  $B$  subspace.

**A.2.3.2 Steps 2 + 3: preparation for weak measurement.** To avoid the problems discussed above, this step partially extracts the information stored in the first ancilla into a second ancilla, then resets the first ancilla using a second application of  $U_C$ . For ease of notation, we combine these two steps here and omit the first ancilla (now unentangled and in the state  $|0\rangle\langle 0|$ ) from the expression:

$$\begin{aligned} \rho_\delta = & \left( \cos(\delta)^2 \sum_{j,k \in A} \rho_{j,k} |\psi_j\rangle\langle\psi_k| + \sum_{j,k \in B} \rho_{j,k} |\psi_j\rangle\langle\psi_k| + \cos(\delta) \sum_{j \in A, k \in B} (\rho_{j,k} |\psi_j\rangle\langle\psi_k| + \rho_{k,j} |\psi_k\rangle\langle\psi_j|) \right) \otimes |0\rangle\langle 0| \\ & + \sin(\delta) \cos(\delta) \sum_{j,k \in A} \rho_{j,k} |\psi_j\rangle\langle\psi_k| \otimes (|0\rangle\langle 1| + |1\rangle\langle 0|) \\ & + \sin(\delta) \sum_{j \in A, k \in B} (\rho_{j,k} |\psi_j\rangle\langle\psi_k| \otimes |1\rangle\langle 0| + \rho_{k,j} |\psi_k\rangle\langle\psi_j| \otimes |0\rangle\langle 1|) \\ & + \sin(\delta)^2 \sum_{j,k \in A} \rho_{j,k} |\psi_j\rangle\langle\psi_k| \otimes |1\rangle\langle 1|. \end{aligned} \quad (\text{B44})$$



Note that only the first and last of these expressions will be relevant once the ancilla is measured.

*A.2.3.3 Step 4: measurement and heralding.* We are now ready to see the effect of a measurement of the second ancilla and to understand the role played by the parameter  $\delta$ . Consider first the projectors corresponding to the two measurement outcomes of the projective measurement:

$$\Pi_A = \mathbb{1} \otimes \left( \frac{\mathbb{1} - Z}{2} \right) \quad (\text{B45})$$

$$\Pi_B = \mathbb{1} \otimes \left( \frac{\mathbb{1} + Z}{2} \right) \quad (\text{B46})$$

Assume now that the measurement yields a ‘1’ outcome. The probability of this event is

$$\begin{aligned} p_1 &= \text{Tr}(\rho_\delta \Pi_A) = \text{Tr}(\Pi_A \rho_\delta \Pi_A) \\ &= \sin(\delta)^2 \text{Tr} \left( \sum_{j,k \in A} \rho_{j,k} |\psi_j\rangle\langle\psi_k| \right) \\ &= \sin(\delta)^2 \sum_{j \in A} \rho_{j,j} \\ &= \sin(\delta)^2 p_{\text{suc}}, \end{aligned} \quad (\text{B47})$$

where  $p_{\text{suc}} = \sum_{j \in A} \rho_{j,j}$  is the probability of the system having transitioned to a state labeled by the criteria as successfully merged. Assuming  $p_1$  is nonzero, the state after the measurement is given by

$$\begin{aligned} \rho_1 &= \frac{1}{p_1} \Pi_A \rho_\delta \Pi_A \\ &= \frac{1}{p_{\text{suc}}} \sum_{j,k \in A} \rho_{j,k} |\psi_j\rangle\langle\psi_k|. \end{aligned} \quad (\text{B48})$$

Thus, this measurement outcome is the desirable one: it heralds that the state has been projected onto the desired subspace. Furthermore, the probability of getting this outcome depends both on the overlap of the input state with the desired subspace and the parameter  $\delta$ , with values of  $\delta \approx \frac{\pi}{2}$  maximizing the probability.

Consider now instead the ‘0’ outcome. The probability for this outcome is

$$p_0 = 1 - p_1 = 1 - \sin(\delta)^2 p_{\text{suc}} \quad (\text{B49})$$

and the post-measurement state is

$$\begin{aligned} \rho_0 &= \frac{1}{p_0} \Pi_B \rho_\delta \Pi_B \\ &= \frac{1}{p_0} \left( \cos(\delta)^2 \sum_{j,k \in A} \rho_{j,k} |\psi_j\rangle\langle\psi_k| + \sum_{j,k \in B} \rho_{j,k} |\psi_j\rangle\langle\psi_k| + \cos(\delta) \sum_{j \in A, k \in B} (\rho_{j,k} |\psi_j\rangle\langle\psi_k| + \rho_{k,j} |\psi_k\rangle\langle\psi_j|) \right). \end{aligned} \quad (\text{B50})$$

Note that this state closely resembles the input state from eqn (B42), except the components related to the  $A$  subspace have decreased in magnitude by a factor of  $\cos(\delta)$  and a renormalization by  $1/p_0$  has occurred. To better understand this measurement-induced perturbation, we can define the following normalized states and coefficients

$$\rho_A = \rho_1 = \frac{1}{p_{\text{suc}}} \sum_{j,k \in A} \rho_{j,k} |\psi_j\rangle\langle\psi_k| \quad \rho_B = \frac{1}{1 - p_{\text{suc}}} \sum_{j,k \in B} \rho_{j,k} |\psi_j\rangle\langle\psi_k| \quad (\text{B51})$$

$$\Lambda_A(\delta, p_{\text{suc}}) = \frac{\sin(\delta)^2 (1 - p_{\text{suc}}) p_{\text{suc}}}{(1 - p_{\text{suc}}) + \cos(\delta)^2 p_{\text{suc}}} \quad (\text{B52})$$

$$\Lambda_B(\delta, p_{\text{suc}}) = \frac{\sin(\delta)^2 (1 - p_{\text{suc}}) p_{\text{suc}}}{1 - \sin(\delta)^2 p_{\text{suc}}} \quad (\text{B53})$$

$$\Lambda_C(\delta, p_{\text{suc}}) = \frac{1 - \cos(\delta) - \sin(\delta)^2 p_{\text{suc}}}{1 - \sin(\delta)^2 p_{\text{suc}}} \quad (\text{B54})$$

to rewrite the state as

$$\begin{aligned} \rho_0 &= \rho - \Lambda_A(\delta, p_{\text{suc}}) \rho_A + \Lambda_B(\delta, p_{\text{suc}}) \rho_B \\ &\quad - \Lambda_C(\delta, p_{\text{suc}}) \sum_{j \in A, k \in B} (\rho_{j,k} |\psi_j\rangle\langle\psi_k| + \rho_{k,j} |\psi_k\rangle\langle\psi_j|) \end{aligned} \quad (\text{B55})$$

$$\begin{aligned} &\simeq \rho - \delta^2 (1 - p_{\text{suc}}) p_{\text{suc}} \rho_A + \delta^2 (1 - p_{\text{suc}}) p_{\text{suc}} \rho_B \\ &\quad - \delta^2 \left( \frac{1}{2} - p_{\text{suc}} \right) \sum_{j \in A, k \in B} (\rho_{j,k} |\psi_j\rangle\langle\psi_k| + \rho_{k,j} |\psi_k\rangle\langle\psi_j|) + O(\delta^4). \end{aligned} \quad (\text{B56})$$

Thus, receiving a measurement result of ‘0’ implies that the state has been left mostly unchanged, except primarily for two effects: the desired components related to  $\rho_A$  have decreased by an amount  $\Lambda_A \rho_A$ , while the unwanted components related to  $\rho_B$  have increased by an amount  $\Lambda_B \rho_B$ . In other words, the measurement has caused a shift towards the subspace of undesired states, with the magnitude of the shift depending on  $\delta$  and  $p_{\text{suc}}$ . From this, we see a trade-off in play when picking the parameter  $\delta$ . As shown above, the largest probability of successful heralded projection requires  $\delta \approx \frac{\pi}{2}$ , but in this case the desired  $A$  part of the state is completely lost whenever the measurement outcome ‘0’ occurs (see eqn (B50)). On the other hand, a small value of  $\delta$  implies a small probability of successful heralded projection but also a small state perturbation in the case of the outcome ‘0’, with both scaling as  $\delta^2$  in the small- $\delta$  limit. In this sense,  $\delta$  represents the power of the measurement, with strong measurements yielding a higher chance of detecting a successful molecular merger but also a higher



disruptive impact of the measurement on the state. Picking good schedules for adjusting  $\delta$  has previously been studied in the context of Grover search,<sup>101–103</sup> and will likely also play a significant role in determining the performance of the simulation approaches presented here.

## Acknowledgements

The authors thank Juan Bernardo Perez Sanchez, Kouhei Nakaji, Mohammad Ghazi Vakili, FuTe Wong, and Sumner Alperin-Lea for helpful discussions and comments. We further thank Jakob S. Kottmann for confirming that LBK is always right. JACGA acknowledges support from the National Institute of Advanced Industrial Science and Technology of Japan (AIST). AA gratefully acknowledges King Abdullah University of Science and Technology (KAUST) for the KAUST Ibn Rushd Postdoctoral Fellowship. LBK acknowledges support from the Carlsberg Foundation. This research was developed with funding from the Defence Advanced Research Projects Agency (DARPA) and the Department of the Navy, Office of Naval Research. The views, opinions, and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. AAG acknowledges support from the NSERC Industrial Research Chairs Program and the Canada 150 Research Chairs. This work is supported by the Novo Nordisk Foundation, grant number NNF22SA0081175, NNF Quantum Computing Programme. Resources used in preparing this research were provided, in part, by the Province of Ontario, the Government of Canada through CIFAR, and companies sponsoring the Vector Institute (<https://vectorinstitute.ai/partnerships/current-partners/>).

## Notes and references

† Note that energy scales in general also depend on the fidelity of the problem representation, *e.g.*, the resolution of the discretization or the size of the plane-wave basis.<sup>20</sup> Thus, a polynomial scaling of these quantities is implicitly assumed in this argument.

‡ The  $S$ -matrix has transition elements concerning a process represented by the evolution step in Fig. 3, independent of the state preparation scheme in the molecule factory part.

§ To see this in the case of the tilded operators, compare their effects on states from the complete basis  $\{|\psi(\{r, R\})\rangle\}$ .

- 1 P. Benioff, *J. Stat. Phys.*, 1980, **22**, 563.
- 2 R. P. Feynman, *Int. J. Theor. Phys.*, 1982, **21**, 467.
- 3 Y. Manin, *Computable and Uncomputable*, Sovetskoye Radio, Moscow, 1980, vol. 28, p. 128.
- 4 C. Zalka, *Proc. R. Soc. London, Ser. A*, 1998, **454**, 313.
- 5 S. Wiesner, *arXiv*, 1996, preprint, arXiv:quant-ph/9603028, DOI: [10.48550/arXiv.quant-ph/9603028](https://doi.org/10.48550/arXiv.quant-ph/9603028).
- 6 D. S. Abrams and S. Lloyd, *Phys. Rev. Lett.*, 1997, **79**, 2586.
- 7 D. Aharonov and A. Ta-Shma, in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing, STOC '03*, Association for Computing Machinery, New York, NY, USA, 2003, pp. 20–29.

- 8 D. W. Berry, G. Ahokas, R. Cleve and B. C. Sanders, *Commun. Math. Phys.*, 2007, **270**, 359.
- 9 S. Lloyd, *Science*, 1996, **273**, 1073.
- 10 D. S. Abrams and S. Lloyd, *Phys. Rev. Lett.*, 1999, **83**, 5162.
- 11 J. Kempe, A. Kitaev, and O. Regev, in *FSTTCS 2004: Foundations of Software Technology and Theoretical Computer Science, Lecture Notes in Computer Science*, ed. K. Lodaya and M. Mahajan, Springer, Berlin, Heidelberg, 2005, pp. 372–383.
- 12 Y. Zheng, J. Leng, Y. Liu, and X. Wu, *arXiv*, 2024, preprint, arXiv:2411.05120, DOI: [10.48550/arXiv.2411.05120](https://doi.org/10.48550/arXiv.2411.05120).
- 13 S. Bravyi, A. J. Bessen, and B. M. Terhal, *arXiv*, 2006, preprint, arXiv:quant-ph/0611021, DOI: [10.48550/arXiv.quant-ph/0611021](https://doi.org/10.48550/arXiv.quant-ph/0611021).
- 14 A. Y. Kitaev, *arXiv*, 1995, preprint, arXiv:quant-ph/9511026, DOI: [10.48550/arXiv.quant-ph/9511026](https://doi.org/10.48550/arXiv.quant-ph/9511026).
- 15 A. Aspuru-Guzik, A. D. Dutoi, P. J. Love and M. Head-Gordon, *Science*, 2005, **309**, 1704.
- 16 J. D. Whitfield, J. Biamonte and A. Aspuru-Guzik, *Mol. Phys.*, 2011, **109**, 735.
- 17 M. Reiher, N. Wiebe, K. M. Svore, D. Wecker and M. Troyer, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 7555.
- 18 P.-W. Huang, G. Boyd, G.-L. R. Anselmetti, M. Degroote, N. Moll, R. Santagati, M. Streif, B. Ries, D. Marti-Dafcik, H. Jnane, *et al.*, *arXiv*, 2025, preprint, arXiv:2508.16719, DOI: [10.48550/arXiv.2508.16719](https://doi.org/10.48550/arXiv.2508.16719).
- 19 A. M. Childs, Y. Su, M. C. Tran, N. Wiebe and S. Zhu, *Phys. Rev. X*, 2021, **11**, 011020.
- 20 G. H. Low and I. L. Chuang, *Quantum*, 2019, **3**, 163.
- 21 D. W. Berry, C. Gidney, M. Motta, J. R. McClean and R. Babbush, *Quantum*, 2019, **3**, 208.
- 22 D. Motlagh and N. Wiebe, *PRX Quantum*, 2024, **5**, 020368.
- 23 E. Campbell, *Phys. Rev. Lett.*, 2019, **123**, 070503.
- 24 K. Nakaji, M. Bagherimehrab and A. Aspuru-Guzik, *PRX Quantum*, 2024, **5**, 020330.
- 25 M. Poernic, M. Hagan, J. Carrasquilla, D. Segal and N. Wiebe, *Phys. Rev. Res.*, 2024, **6**, 013224.
- 26 A. Rajput, A. Roggero and N. Wiebe, *Quantum*, 2022, **6**, 780.
- 27 M. Hagan and N. Wiebe, *Quantum*, 2023, **7**, 1181.
- 28 M. Bagherimehrab, L. M. Calderon, D. W. Berry, P. Schleich, M. G. Vakili, A. Aldossary, J. A. C. G. Angulo, C. Gorgulla, and A. Aspuru-Guzik, Faster algorithmic quantum and classical simulations by corrected product formulas, *arXiv*, 2024, preprint, arXiv:2409.08265 [quant-ph], DOI: [10.48550/arXiv.2409.08265](https://doi.org/10.48550/arXiv.2409.08265).
- 29 J. L. Bosse, A. M. Childs, C. Derby, F. M. Gambetta, A. Montanaro and R. A. Santos, *Nat. Commun.*, 2025, **16**(1), 2673.
- 30 Y. Su, D. W. Berry, N. Wiebe, N. Rubin and R. Babbush, *PRX Quantum*, 2021, **2**, 040332.
- 31 T. N. Georges, M. Bothe, C. Sünderhauf, B. K. Berntson, R. Izsák and A. V. Ivanov, *npj Quantum Inf.*, 2025, **11**, 55.
- 32 A. Caesura, C. L. Cortes, W. Pol, S. Sim, M. Steudtner, G.-L. R. Anselmetti, M. Degroote, N. Moll, R. Santagati, M. Streif, *et al.*, *PRX Quantum*, 2025, **6**, 030337.
- 33 V. C. Olaya-Agudelo, B. Stewart, C. H. Valahu, R. J. MacDonell, M. J. Millican, V. G. Matsos,



F. Scuccimarra, T. R. Tan and I. Kassal, *Phys. Rev. Res.*, 2025, **7**, 023215.

34 N. Maskara, S. Ostermann, J. Shee, M. Kalinowski, A. McClain Gomez, R. Araiza Bravo, D. S. Wang, A. I. Krylov, N. Y. Yao, M. Head-Gordon, M. D. Lukin and S. F. Yelin, *Nat. Phys.*, 2025, **21**, 289–297.

35 P.-O. Löwdin, in *Advances in Chemical Physics*, John Wiley & Sons, Ltd, 1958, pp. 207–322.

36 P.-O. Löwdin, *Int. J. Quantum Chem.*, 1995, **55**, 77.

37 S. Lee, J. Lee, H. Zhai, Y. Tong, A. M. Dalzell, A. Kumar, P. Helms, J. Gray, Z.-H. Cui, W. Liu, M. Kastoryano, R. Babbush, J. Preskill, D. R. Reichman, E. T. Campbell, E. F. Valeev, L. Lin and G. K.-L. Chan, *Nat. Commun.*, 2023, **14**, 1952.

38 S. Simon, G.-L. R. Anselmetti, R. Santagati, M. Degroote, N. Moll, M. Streif, and N. Wiebe, *arXiv*, 2025, preprint, arXiv:2504.03465, DOI: [10.48550/arXiv.2504.03465](https://doi.org/10.48550/arXiv.2504.03465).

39 A. Ambainis, in *2014 IEEE 29th Conference on Computational Complexity (CCC)*, 2014, pp. 32–43, iSSN: 1093-0159.

40 J. D. Watson and J. Bausch, *arXiv*, 2021, preprint, arXiv:2105.13350, DOI: [10.48550/arXiv.2105.13350](https://doi.org/10.48550/arXiv.2105.13350).

41 J. Bausch, T. S. Cubitt and J. D. Watson, *Nat. Commun.*, 2021, **12**, 452.

42 C.-F. Chen, H.-Y. Huang, J. Preskill, and L. Zhou, in *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, 2024, pp. 1323–1330.

43 N. Schuch and F. Verstraete, *Nat. Phys.*, 2009, **5**, 732.

44 J. D. Whitfield, P. J. Love and A. Aspuru-Guzik, *Phys. Chem. Chem. Phys.*, 2012, **15**, 397.

45 B. O’Gorman, S. Irani, J. Whitfield and B. Fefferman, *PRX Quantum*, 2022, **3**, 020322.

46 I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni and A. Aspuru-Guzik, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 18681.

47 H. H. S. Chan, R. Meister, T. Jones, D. P. Tew and S. C. Benjamin, *Sci. Adv.*, 2023, **9**, eab07484.

48 S. Simon, R. Santagati, M. Degroote, N. Moll, M. Streif and N. Wiebe, *PRX Quantum*, 2024, **5**, 010343.

49 R. Babbush, D. W. Berry, I. D. Kivlichan, A. Y. Wei, P. J. Love and A. Aspuru-Guzik, *New J. Phys.*, 2016, **18**, 033032.

50 I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan and R. Babbush, *Phys. Rev. Lett.*, 2018, **120**, 110501.

51 R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler and H. Neven, *Phys. Rev. X*, 2018, **8**, 041015.

52 R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven and G. K.-L. Chan, *Phys. Rev. X*, 2018, **8**, 011044.

53 J. R. McClean, F. M. Faulstich, Q. Zhu, B. O’Gorman, Y. Qiu, S. R. White, R. Babbush and L. Lin, *New J. Phys.*, 2020, **22**, 093015.

54 J. S. Pedernales, R. Di Candia, I. L. Egusquiza, J. Casanova and E. Solano, *Phys. Rev. Lett.*, 2014, **113**, 020505.

55 E. Kökcü, H. A. Labib, J. Freericks and A. F. Kemper, *Nat. Commun.*, 2024, **15**, 3881.

56 F. Tacchino, A. Chiesa, S. Carretta and D. Gerace, *Adv. Quantum Technol.*, 2020, **3**, 1900052.

57 L. Lin and Y. Tong, *PRX Quantum*, 2022, **3**, 010318.

58 N. L. Diaz, P. Braccia, M. Larocca, J. M. Matera, R. Rossignoli and M. Cerezo, *Phys. Rev. Res.*, 2025, **7**(3), 033294.

59 P. J. Ollitrault, G. Mazzola and I. Tavernelli, *Phys. Rev. Lett.*, 2020, **125**, 260511.

60 H. Liu, G. H. Low, D. S. Steiger, T. Häner, M. Reiher and M. Troyer, *Mater. Theory*, 2022, **6**, 11.

61 R. Bhuyan, J. Mony, O. Kotov, G. W. Castellanos, J. Gómez Rivas, T. O. Shegai and K. Börjesson, *Chem. Rev.*, 2023, **123**, 10877.

62 B. F. E. Curchod and T. J. Martínez, *Chem. Rev.*, 2018, **118**, 3305.

63 D. Motlagh, R. A. Lang, P. Jain, J. A. Campos-Gonzalez-Angulo, W. Maxwell, T. Zeng, A. Aspuru-Guzik and J. M. Arrazola, *Quantum Sci. Technol.*, 2025, **10**(4), 045048.

64 D. J. Tannor, *Introduction to Quantum Mechanics*, University Science Books, 2007.

65 V. Vitale, J. Dziedzic, S. M.-M. Dubois, H. Fangohr and C.-K. Skylaris, *J. Chem. Theory Comput.*, 2015, **11**, 3321.

66 N. T. Hunt, *Chem. Soc. Rev.*, 2009, **38**, 1837.

67 I. Noda, *Appl. Spectrosc.*, 1990, **44**, 550.

68 L. Bassman Oftelie, K. Klymko, D. Liu, N. M. Tubman and W. A. de Jong, *Phys. Rev. Lett.*, 2022, **129**, 130603.

69 J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe and S. Lloyd, *Nature*, 2017, **549**, 195.

70 M. Cerezo, G. Verdon, H.-Y. Huang, L. Cincio and P. J. Coles, *Nat. Comput. Sci.*, 2022, **2**, 567.

71 H.-Y. Huang, M. Broughton, J. Cotler, S. Chen, J. Li, M. Mohseni, H. Neven, R. Babbush, R. Kueng, J. Preskill and J. R. McClean, *Science*, 2022, **376**, 1182.

72 A. Angulo, L. Yang, E. S. Aydil and M. A. Modestino, *Digital Discovery*, 2022, **1**, 35.

73 J. F. Joung, M. Han, J. Hwang, M. Jeong, D. H. Choi and S. Park, *JACS Au*, 2021, **1**, 427.

74 L. R. Liu, J. D. Hood, Y. Yu, J. T. Zhang, N. R. Hutzler, T. Rosenband and K.-K. Ni, *Science*, 2018, **360**, 900.

75 D. K. Ruttle, A. Guttridge, S. Spence, R. C. Bird, C. R. Le Sueur, J. M. Hutson and S. L. Cornish, *Phys. Rev. Lett.*, 2023, **130**, 223401.

76 R. C. Bird, C. R. Le Sueur and J. M. Hutson, *Phys. Rev. Res.*, 2023, **5**, 043086.

77 R. C. Bird and J. M. Hutson, *Phys. Rev. Res.*, 2025, **7**(3), 033022.

78 A. Aspuru-Guzik, A. D. Dutoi, P. J. Love and M. Head-Gordon, *Science*, 2005, **309**, 1704.

79 Z. Ding, C.-F. Chen and L. Lin, *Phys. Rev. Res.*, 2024, **6**, 033147.

80 D. W. Berry, Y. Tong, T. Khattar, A. White, T. I. Kim, G. H. Low, S. Boixo, Z. Ding, L. Lin, S. Lee and G. K. L. Chan, *PRX Quantum*, 2025, **6**(2), 020327.

81 W. J. Huggins, O. Leimkuhler, T. F. Stetina and K. B. Whaley, *PRX Quantum*, 2025, **6**(2), 020319.

82 C.-F. Chen, M. J. Kastoryano, F. G. Brandão, and A. Gilyén, *arXiv*, 2023, preprint, arXiv:2303.18224, DOI: [10.48550/arXiv.2303.18224](https://doi.org/10.48550/arXiv.2303.18224).

83 M. Hagan and N. Wiebe, *arXiv*, 2025, preprint, arXiv:2502.03410, DOI: [10.48550/arXiv.2502.03410](https://doi.org/10.48550/arXiv.2502.03410).



84 D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, in *Proceedings of the 46th Annual ACM Symposium on Theory of Computing – STOC '14*, ACM Press, New York, USA, 2014, pp. 283–292.

85 L.-P. Wang, A. Titov, R. McGibbon, F. Liu, V. S. Pande and T. J. Martínez, *Nat. Chem.*, 2014, **6**, 1044.

86 P. A. M. Neto and H. M. Nussenzveig, *Europhys. Lett.*, 2000, **50**, 702.

87 J. E. Molloy and M. J. Padgett, *Contemp. Phys.*, 2002, **43**, 241.

88 P. Casavecchia, *Rep. Prog. Phys.*, 2000, **63**, 355.

89 S. Y. T. van de Meerakker, H. L. Bethlehem and G. Meijer, *Nat. Phys.*, 2008, **4**, 595.

90 C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Photons and Atoms: Introduction to Quantum Electrodynamics*, Wiley, 1st edn, 1997.

91 G. Lindblad, *Commun. Math. Phys.*, 1976, **48**, 119.

92 J. Mehra and E. Sudarshan, *Il Nuovo Cimento*, 1972, **11**, 215.

93 R. Cleve and C. Wang, in *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017) (Schloss-Dagstuhl-Leibniz Zentrum für Informatik)*, 2017.

94 Z. Ding, M. Junge, P. Schleich and P. Wu, *Commun. Math. Phys.*, 2025, **406**, 60.

95 M. Pocanic, D. Segal and N. Wiebe, *J. Phys. A: Math. Theor.*, 2025, **58**(30), 305302.

96 X. Dan, E. Geva and V. S. Batista, *J. Chem. Theory Comput.*, 2025, 1530–1546.

97 C. Wittig, *J. Phys. Chem. B*, 2005, **109**, 8428.

98 T. Takekoshi, M. Debatin, R. Rameshan, F. Ferlaino, R. Grimm, H.-C. Nägerl, C. R. Le Sueur, J. M. Hutson, P. S. Julienne, S. Kotochigova, *et al.*, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2012, **85**, 032506.

99 M. Bagherimehrab and A. Aspuru-Guzik, *Quantum Sci. Technol.*, 2024, **9**, 035010.

100 A. Lund, *New J. Phys.*, 2011, **13**, 053024.

101 P. Andrés-Martínez and C. Heunen, *Quantum Sci. Technol.*, 2022, **7**, 025007.

102 B. Yan, S. Wei, H. Jiang, H. Wang, Q. Duan, Z. Ma and G.-L. Long, *Sci. Rep.*, 2022, **12**, 14339.

103 A. Mizel, *Phys. Rev. Lett.*, 2009, **102**, 150501.

104 D. M. Jackson, U. Haeusler, L. Zaporski, J. H. Bodey, N. Shofer, E. Clarke, M. Hugues, M. Atatüre, C. Le Gall and D. A. Gangloff, *Phys. Rev. X*, 2022, **12**, 031014.

105 W. G. Hoover, *Computational Statistical Mechanics*, Elsevier, 1st edn, 1991.

106 Z. Liu, X. Li, C. Wang, and J.-P. Liu, *arXiv*, 2024, preprint, arXiv:2411.03972, DOI: [10.48550/arXiv.2411.03972](https://doi.org/10.48550/arXiv.2411.03972).

107 R. Babbush, D. W. Berry, R. Kothari, R. D. Somma and N. Wiebe, *Phys. Rev. X*, 2023, **13**, 041041.

108 L. Lang, H. M. Cezar, L. Adamowicz and T. B. Pedersen, *J. Am. Chem. Soc.*, 2024, **146**, 1760.

109 H. E. Li, Y. Zhan and L. Lin, *npj Quantum Inf.*, 2025, **11**(1), 183.

110 G. H. Low and N. Wiebe, *arXiv*, 2018, preprint, arXiv:1805.00675, DOI: [10.48550/arXiv.1805.00675](https://doi.org/10.48550/arXiv.1805.00675).

111 Y. R. Sanders, G. H. Low, A. Scherer and D. W. Berry, *Phys. Rev. Lett.*, 2019, **122**, 020502.

112 X.-M. Zhang and X. Yuan, *npj Quantum Inf.*, 2024, **10**, 42.

113 Z. Sun, G. Boyd, Z. Cai, H. Jnane, B. Koczor, R. Meister, R. Minko, B. Pring, S. C. Benjamin and N. Stamatopoulos, Low-depth phase oracle using a parallel piecewise circuit, *Physical Review A*, 2025, **111**(6), 062420, DOI: [10.1103/m32k-7nq2](https://doi.org/10.1103/m32k-7nq2).

114 D. W. Berry, M. Kieferová, A. Scherer, Y. R. Sanders, G. H. Low, N. Wiebe, C. Gidney and R. Babbush, *npj Quantum Inf.*, 2018, **4**, 1.

115 Z. Liu, A. M. Childs, and D. Gottesman, *arXiv*, 2024, preprint, arXiv:2411.04019, DOI: [10.48550/arXiv.2411.04019](https://doi.org/10.48550/arXiv.2411.04019).

