



**HAL**  
open science

# Leveraging Analog Quantum Computing with Neutral Atoms for Solvent Configuration Prediction in Drug Discovery

M. D'arcangelo, L.-P. Henry, L. Henriët, D. Loco, Nicolaï Gouraud, S. Angebault, J. Sueiro, J. Forêt, P. Monmarché, Jean-Philip J.-P.

► **To cite this version:**

M. D'arcangelo, L.-P. Henry, L. Henriët, D. Loco, Nicolaï Gouraud, et al.. Leveraging Analog Quantum Computing with Neutral Atoms for Solvent Configuration Prediction in Drug Discovery. *Physical Review Research*, 2024, 6 (4), pp.043020. 10.1103/PhysRevResearch.6.043020 . hal-04214318

**HAL Id: hal-04214318**

**<https://hal.science/hal-04214318v1>**

Submitted on 5 Oct 2024


**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.


L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.




Distributed under a Creative Commons Attribution 4.0 International License

## Leveraging analog quantum computing with neutral atoms for solvent configuration prediction in drug discovery

Mauro D’Arcangelo <sup>\*</sup>, Louis-Paul Henry, and Loïc Henriët<sup>†</sup>  
*PASQAL, 7 Rue Léonard de Vinci, 91300 Massy, France*

Daniele Loco,<sup>\*</sup> Nicolai Gouraud , Stanislas Angebault, Jules Sueiro, and Jérôme Forêt  
*Qubit Pharmaceuticals, Advanced Research Department, 24 rue du Faubourg Saint-Jacques, 75014 Paris, France*

Pierre Monmarché  
*Sorbonne Université, Laboratoire Jacques-Louis Lions, UMR 7589 CNRS, 75005 Paris, France*

Jean-Philip Piquemal <sup>‡</sup>  
*Qubit Pharmaceuticals, Advanced Research Department, 24 rue du Faubourg Saint-Jacques, 75014 Paris, France  
 and Sorbonne Université, Laboratoire de Chimie Théorique, UMR 7616 CNRS, 75005 Paris, France*



(Received 4 October 2023; accepted 25 August 2024; published 4 October 2024)

We introduce an approach to sampling equilibrium solvent water molecule configurations within proteins that leverages analog quantum computing. We present a complete end-to-end study from the molecular biology application to the development of the quantum algorithm to the implementation on a neutral atom quantum processing unit (QPU). To do so, we combine a quantum placement strategy to the 3D Reference Interaction Site Model, an approach capable of predicting continuous solvent distributions. The intrinsic quantum nature of such coupling guarantees molecules not to be placed too close to each other, a constraint usually imposed by hand in classical approaches. We present first a full quantum adiabatic evolution model that uses a local Rydberg Hamiltonian to cast the general problem into an antiferromagnetic Ising model. Its solution is embodied into a Rydberg atom array QPU. Following a classical emulator implementation, a QPU portage allows to experimentally validate the algorithm performances on an actual quantum computer. As a perspective of use on next generation devices, we emulate a second hybrid quantum-classical version of the algorithm. Such a variational quantum approach uses a classical Bayesian minimization routine to find the optimal laser parameters. Overall, these Quantum-3D-RISM algorithms open a route towards the application of analog quantum computing in molecular modeling and drug design.

DOI: [10.1103/PhysRevResearch.6.043020](https://doi.org/10.1103/PhysRevResearch.6.043020)

### I. INTRODUCTION

The direct manipulation of quantum systems to perform quantum computations has become an intense field of interdisciplinary research aiming to apply quantum computing to various fields ranging, from theoretical chemistry and many-body physics to material sciences and drug design. For example, such technologies are believed to be ultimately able to improve the accuracy of quantum chemistry methods [1–4] through algorithms exponentially faster than classical ones [1,2,5]. If no widely accepted proof of exponential advantage for quantum computing calculations has been yet provided [6], then in practice a polynomial advantage could still

contribute to bring more accurate simulations at hand in near-term quantum devices [1,5–7]. More pragmatically, a lower energy consumption of quantum devices is expected to be more rapidly envisioned.

Reasons for this present situation are linked to near-term quantum technologies limitations in term of qubit count and to the absence of fully error-corrected gate-based machines, whose simulations are thus still suffering from a high level of noise [2,7,8]. In that context, analog quantum computing is an alternative promising strategy [9,10].

On the software side, many groups have contributed to developing algorithms that are expected to scale efficiently with the size of chemical systems, once that more advanced hardware will become available [1,2]. Such techniques include the variational quantum algorithms (VQAs) [5,7], a family of algorithms allowing to exploit present hardware through coupling with classical optimization. In a nutshell, if one is able to formulate a problem in terms of a mathematical cost function to be minimized or maximized with respect to the parameters of the quantum machine, then the search for the best parameters can be outsourced to a classical hardware, the quantum computer being used to compute the cost

<sup>\*</sup>These authors contributed equally to this work.

<sup>†</sup>Contact author: loic.henriet@pasqal.com

<sup>‡</sup>Contact author: jean-philip.piquemal@sorbonne-universite.fr

function, a task usually too expensive to compute classically. The solution to the problem is then given in terms of the best parameters for the quantum evolution. VQAs have nowadays well-known limitations, e.g., barren plateaus encountered in the classical optimization loop [11], so that their interest in practical usage in the near and long term is still an open question. Further research is compulsory to address this issue and to further progress in the understanding of this class of algorithms, that lead the way in different fields of applications and research in quantum computing.

Concerning the application of quantum computing to drug design, a necessary tool for accurate drug-target affinity predictions in structure-based drug discovery (DD) is related to methods capable of addressing the problem of predicting the hydration sites in protein structures or any biomolecular system of interest. Predicting the solvent structure inside important protein pockets thanks to user-friendly and computationally efficient methods is a determining task to improve the predicting power of numerical simulations. An often complex balance between enthalpic and entropic contributions of the water hydration effect is known to potentially affect the molecular recognition mechanism at the protein-ligand interface or on a larger scale the biomolecule structure stability and their conformational changes. The role of water has been especially investigated in protein hydration, both numerically and experimentally. Simulations based on molecular dynamics approaches (MD) have shown the dynamic nature of protein hydration: Water molecules networks can be temporarily present or absent in nonpolar regions of the protein, or exchange between the bulk of the solvent and more polar regions [12–15]. Such dynamic equilibrium and the usual size of the molecular systems targeted in DD (and not only), characterized by a daunting number of degrees of freedom, requires to deal with the computational effort of sampling a highly dimensional phase space. The relevance of water effects can be even more relevant in cases of poorly accessible regions in the interior of the protein as deeply buried pockets [14]. There are methods such as the three-dimensional (3D) Reference Interaction Site Model (3D-RISM) that can avoid this effort through the use of continuous distributions [16–21] based on a statistical mechanics integral equation formalism. However they lack information about the exact locations of molecules. In this work, we present the first algorithms, exploiting the analog quantum computing paradigm, to analyze the hydration structures within proteins' binding sites thanks to the 3D-RISM continuous distribution. Indeed, a poor treatment of the protein-hydration problem is often a limiting factor preventing the access to accurate protein-ligand affinity predictions, a crucial step for practitioners in drug discovery. Our strategy is applied to small molecules of interest in drug design for which the effect of solvent water molecules is substantial. The main advantage of this proposed approach is its potential to efficiently sample the distribution of water molecules inside protein cavities. Therefore, this work presents the design of quantum algorithmic strategies in that direction. To do so, we implemented the approach on both PASQAL's Pulser emulator and real neutral atom quantum computer, which, thanks to its flexible connectivity and ease of implementation of analog operations, has demonstrated the potential to solve combinatorial problems. The sampling of

water molecules in biological systems distribution can also be seen as a combinatorial problem on a graph, since it corresponds to finding the most probable disposition of solvent molecules given the constraints imposed by both the protein structure and the rest of the solvent atomistic structure.

We remark here that a number of similar applications, in combinatorial optimization with quantum annealers [22,23], can be found in the literature, e.g., using photonic hardware [24–28] or superconductive qubits [29–33]. A portion of these works present extremely promising theoretical applications [24,26,27] of Gaussian boson sampling in photonic technologies. Recent advances have also been done in developing the photonic approach towards real-life applications, as in drug discovery to address the molecular docking problem [25,28] or macromolecules folding [25]. Two alternative formulations of the docking problem in QUBO form are found in Refs. [30,31], which have been implemented on D-Wave annealers. The former uses the Rosetta energy function to encode atomic interactions, while the latter is purely geometric and leverages a weighted subgraph isomorphism representation of the affinity between ligand and protein cavity. From a classical method perspective, molecular docking and solvent configuration predictions are addressed with radically different approaches. From a theoretical perspective their aim is not equivalent: Molecular docking is a screening technique used to predict optimal configurations of a molecule in interaction, often, with a protein, under defined structural constraints [34]. It is used as a cheap computational approach to screen large libraries of compounds, to efficiently find the most promising ones for a given target biomolecule. The problem of solvation is instead related to statistical mechanics, to find the distribution of solvent molecules in interactions with a desired chemical species (the solute). The amount of solvent molecules is usually overwhelming compared to that of solutes, and its configuration not only depends on the solute, but on the mutual effects between solute and solvents [35]. An accurate account of such effects have implications in drug discovery, and numerically expensive methods can be needed due to the large dimensionality of the problem and the intrinsically slow solvent dynamics around solutes [35].

## II. THE IMPORTANCE OF WATER PLACEMENT

The distribution of water molecules impacts the protein's structure and determines its overall shape. Water molecules also mediate the interaction between proteins and other small molecules, the so-called ligands. The presence of water generally influences the binding capacity of a ligand to a specific protein site [35,36], a problem of particular interest for the discovery of new drugs. In practice, water molecules can occupy ligand-protein interaction sites, and clusters of water molecules, inducing complex solvation or desolvation thermodynamics processes, can deeply affect the thermodynamics of ligand binding modes. Consequently, accounting for protein solvation effects is crucial in structure-based drug design to reliably model the ligand activity through a rational design of the ligand properties [35,36]. To a certain extent, the presence of water molecules can be experimentally established using x-ray crystallography, or other techniques. However, experiments can present accuracy limitations due to poor

resolution [37] and crystal structures with resolution  $>2.7 \text{ \AA}$  [38] represent a limit for a reliable modeling of water molecules and it is then critical to refine the structure using accurate molecular modeling tools. Therefore a consistent, reliable, and fast water placement procedure is relevant in different application scenarios.

The use of grid-based techniques such as 3D-RISM [20] is useful to complement experimental information, avoiding the use of long and more expensive, even if accurate, sampling methods. 3D-RISM has theoretical basis in the statistical mechanical integral equation theories of liquids [20]. A grid of points, on which the equations are solved, has to be defined in real space around the solute system, e.g., a protein. One of the main outputs is the solvent structure distribution function, namely  $g(\mathbf{r})$  around that solute. More details about 3D-RISM can be found by the interested reader in the Supplemental Material (SM) [39]. A realistic grid size for applications in DD can easily be represented by a cube with 80 points for each side, spaced by  $0.5 \text{ \AA}$ , for a total of around 512 000 points. These are quite standard parameters [40–43], even if the number of grid points rapidly increases for higher accuracy in the representation of the solvent distribution. The  $g(\mathbf{r})$  density represents in a continuous fashion (in practice on a given set of grid points) the water networks around the solute, and maxima of this function can be associated to stable water molecules in the region of space where the value is localized (i.e., the grid point location). Maxima, in the distribution or in the associated thermodynamics functions, can be misleading since the structure of the  $g(\mathbf{r})$  is potentially not trivial, representing clusters of water molecules organized around the solute. Increasing the number of grid points to achieve a better accuracy has also the effect of increasing the complexity of decomposing the full density in localized atomic positions.

To address this problem, we translate the search for discrete water molecule positions from a 3D-RISM density to a Gaussian mixture problem. In this formulation each Gaussian component represents a water molecule position. There is a vast literature on learning mixtures of Gaussians in arbitrary dimensions, especially where a separation condition in the centers of the Gaussians is imposed [44–47]. Such algorithms would be in principle suitable for our purposes, since in our setting a clear separation is naturally required by the minimal physical distance between water molecules. However, it should be noted that these methods are very sensitive to outliers, where data are corrupted by noise [48]. Although in this sense robust and efficient algorithms with theoretical guarantees have been developed [48,49], we have developed an alternative approach based on an Ising model that presents certain favorable characteristics: It scales independently of the number of centers, and it can be applied equally well to non-Gaussian mixtures. Such algorithm can be implemented on near-term quantum computers, as shown in the next few sections in the special case of a neutral atom quantum processing unit (QPU).

### III. ANALOG QUANTUM COMPUTING WITH NEUTRAL ATOMS

To address the solvent sampling problem, we propose an approach based on a neutral atom QPU coupled to a 3D-RISM

computation. The machine employs arrays of rubidium atoms arranged in arbitrary 2D configurations defined by a layout of optical traps of the type shown in Fig. 1(a). Each atom is described by its position  $\vec{q}_i$  and its internal state, which can be in general an arbitrary superposition of the ground state  $|0\rangle$  and a highly excited Rydberg state  $|1\rangle$ .

A laser system is tuned close to the resonant frequency of the  $|0\rangle$  to  $|1\rangle$  transition, so that each atom is effectively a qubit living in a two-dimensional Hilbert space. The programmable time-dependent control fields of the driving laser are the Rabi frequency and detuning, denoted  $\Omega(t)$  and  $\Delta(t)$ , respectively. After evolving for a period of time  $T$ , the system can be measured in the computational basis by a fluorescence imaging process that will find certain atoms in  $|0\rangle$ , while the rest are inferred to be in  $|1\rangle$  [9].

The term *ground state* will be used in two different contexts. The ground state of a single isolated atom is denoted  $|0\rangle$ , and the *all-ground state* of a noninteracting  $M$ -qubit system refers to the state  $|\psi_0\rangle := |0\rangle \otimes \dots \otimes |0\rangle = |0\rangle^{\otimes M}$  which is often the initial state of any quantum algorithm. The ground state of the system coupled with interactions and a driving laser, however, corresponds to the state of minimal energy of the Hamiltonian, which is not necessarily the all-ground state. For instance, the Hamiltonian at  $\Omega = 0$  and  $\Delta > 0$  can be related to a known graph-theoretical problem. By interpreting each neutral atom as a vertex in a graph, and connecting with an edge vertices situated at a distance closer than a certain threshold dictated by the physics of interactions, the ground state is a solution to the maximum independent set (MIS) problem [50], i.e., the problem of selecting the largest number of vertices in a graph that are not connected to each other directly by an edge. The connection between neutral atoms and the MIS problem has been extensively explored in recent experimental implementations with up to 256 qubits [51,52].

In the following we will show how to translate the water placement problem to a combinatorial optimization problem, and how the neutral atom Hamiltonian can be used to sample approximate solutions via an adiabatic quantum algorithm.

### IV. A QUANTUM ISING MODEL FORMULATION OF THE PLACEMENT PROBLEM

To study the water placement problem from a quantum perspective, we first place ourselves in the formalism of 3D-RISM. Indeed, this algorithm produces a continuous map of the oxygen atom density inside the protein cavity. We want to translate this continuous information into a discrete set of water molecule positions. In order to do so, we assume that each water molecule randomly oscillates around a definite stable position in the protein cavity. If one tracks over time the oscillating motion of a water molecule and assigns to each unit of space the probability of finding the molecule there, then the resulting probability map should resemble a Gaussian of a certain width centered around the stable position. In this ansatz, the 3D-RISM density is interpreted as the sum of all the Gaussians associated to each water molecule in the cavity. Therefore, the joint density is decomposed into several single-particle densities expressing the uncertainty in the position of each water molecule, and the problem of locating the position of water molecules is transposed to the problem of finding

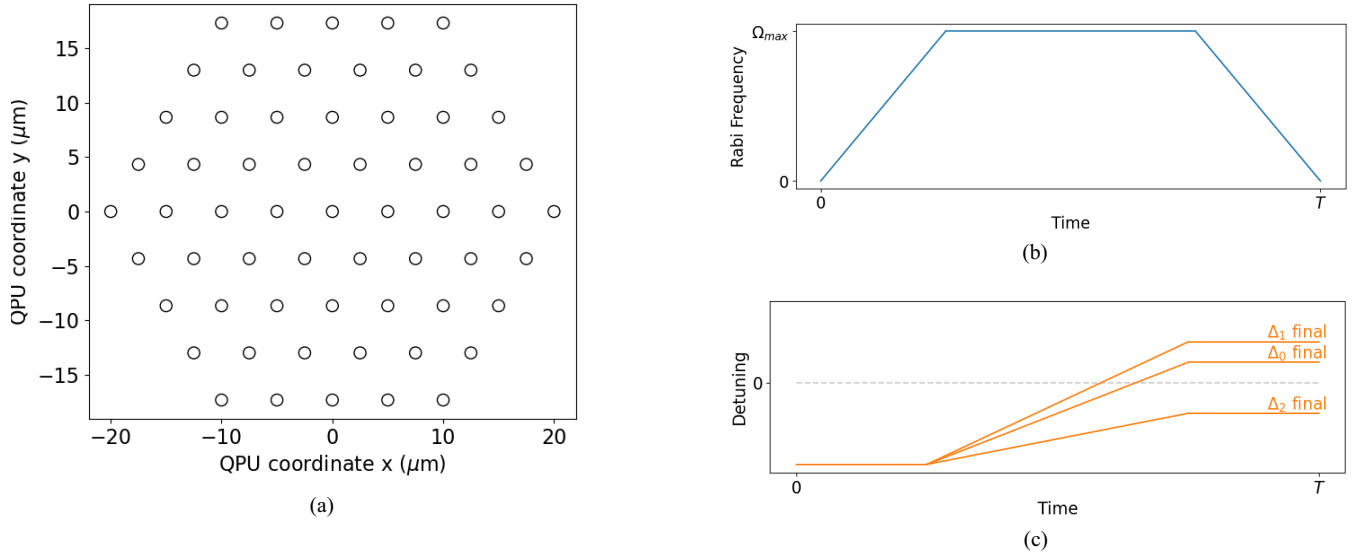


FIG. 1. (a) An example of layout of optical traps in a neutral atom QPU. Each circle represents a site where a Rubidium atom can be trapped and used for quantum computations. The units correspond to the physical size of the trap layout in  $\mu\text{m}$ . (b) and (c) Schematic example of an adiabatic path parametrized by  $\Omega_i(t)$  and  $\Delta_i(t)$  in the Rydberg Hamiltonian (13) for three qubits. The Rabi frequency is kept the same for all qubits, so  $\Omega_i(t) = \Omega(t)$  for all  $i$ , and it vanishes at the extremal points of the path. The Detunings  $\Delta_i(t)$  are ramped up from a negative value to some final value related to the one-body terms of the problem Hamiltonian (3), and therefore they will be different for each qubit.

the best Gaussian mixture that approximates the 3D-RISM density.

Formally, a protein cavity can be modeled as a connected subset  $\mathcal{C} \subset \mathbb{R}^3$  and a 3D-RISM density is a scalar function  $g: \mathcal{C} \rightarrow \mathbb{R}$  normalized such that  $\int_{\mathcal{C}} g(\mathbf{r}) d\mathbf{r} = 1$ . Denote by  $\mathcal{G}(\mu, \sigma^2)$  a normalized isotropic Gaussian with mean  $\mu \in \mathbb{R}^3$  and variance  $\sigma^2$ . For practical purposes, we restrict the possible values of the mean of the Gaussians to a finite discrete lattice  $\mathcal{Q} := \{\mathbf{q}_1, \dots, \mathbf{q}_M\} \subset \mathcal{C}$  rather than the whole continuous space. Assigning a binary variable  $n_i \in \{0, 1\}$  to each point of  $\mathcal{Q}$ , an arbitrary sum of Gaussians can be written:

$$\sum_{i=1}^M \mathcal{G}(\mathbf{q}_i, \sigma^2) n_i. \quad (1)$$

The value of each  $n_i$  will then act as a switch that indicates whether or not a Gaussian is placed in position  $\mathbf{q}_i$ . We want to find the optimal assignment of  $\{n_i\}_{i=1, \dots, M}$  such that the following  $L^2$  norm is minimized:

$$I^2 := \int_{\mathcal{C}} \left( g(\mathbf{r}) - \sum_{i=1}^M \mathcal{G}(\mathbf{q}_i, \sigma^2)(\mathbf{r}) n_i \right)^2 d\mathbf{r}. \quad (2)$$

Expanding the square, one sees that  $I^2$  defines the energy of an Ising model:

$$I^2 = K - \sum_{i=1}^M \Gamma_i n_i + \sum_{i \neq j=1}^M V_{ij} n_i n_j, \quad (3)$$

where  $K$  is an unimportant constant, and the coefficients of the linear and quadratic terms are given by:

$$\Gamma_i := 2 \int_{\mathcal{C}} g(\mathbf{r}) \mathcal{G}(\mathbf{q}_i, \sigma^2)(\mathbf{r}) d\mathbf{r} - \int_{\mathcal{C}} (\mathcal{G}(\mathbf{q}_i, \sigma^2)(\mathbf{r}))^2 d\mathbf{r} \quad (4)$$

and

$$V_{ij} := \int_{\mathcal{C}} \mathcal{G}(\mathbf{q}_i, \sigma^2)(\mathbf{r}) \mathcal{G}(\mathbf{q}_j, \sigma^2)(\mathbf{r}) d\mathbf{r}. \quad (5)$$

For fixed  $\sigma^2$ , the coefficients can be computed numerically, but it is helpful to gain further insight into the interaction term  $V_{ij}$ . When the Gaussians are concentrated far enough from the boundary of  $\mathcal{C}$ , it is possible to approximate the integral by extending the integration to the whole of  $\mathbb{R}^3$ . By performing the change of variables  $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{q}_i$ , the term becomes the convolution of two Gaussians centered in zero:

$$\begin{aligned} V_{ij} &\simeq \int_{\mathbb{R}^3} \mathcal{G}(\mathbf{0}, \sigma^2)(\mathbf{r}) \mathcal{G}(\mathbf{r}_{ij}, \sigma^2)(\mathbf{r}) d\mathbf{r} \\ &= [\mathcal{G}(\mathbf{0}, \sigma^2) * \mathcal{G}(\mathbf{0}, \sigma^2)](\mathbf{r}_{ij}), \end{aligned} \quad (6)$$

where  $\mathbf{r}_{ij} = \mathbf{q}_j - \mathbf{q}_i$ . It is a known result that the convolution of two Gaussians is itself a Gaussian, therefore one has:

$$V_{ij} \sim \exp(-\alpha |\mathbf{r}_{ij}|^2). \quad (7)$$

Equation (3) is therefore a classical Ising model with exponentially decaying interactions. In addition, the water placement problem requires two water molecules to be placed at a minimal physical distance from each other. This translates to an extra constraint on the Ising model (3) in the form of  $n_i = n_j = 1 \Rightarrow |\mathbf{r}_i - \mathbf{r}_j| > R$  for some  $R > 0$ . If  $\{n_1^*, \dots, n_M^*\}$  is the ground state of the constrained Ising model, then the solution to the water placement problem is then defined as:

$$\mathcal{W} := \{\mathbf{q}_i \in \mathcal{Q} \mid n_i^* = 1\}, \quad (8)$$

$$\mathcal{N} := \sum_{i=1}^M n_i^*, \quad (9)$$

with  $\mathcal{N}$  the number of placed water molecules and  $\mathcal{W}$  their positions. This formulation has far-reaching applications to



general finite Gaussian mixture models. See the SM for more details.

The quantum version of an Ising model is obtained by replacing the binary variable  $n_i$  with the number operator  $\hat{n}_i$ , whose spectrum is  $\{0, 1\}$ . In terms of Pauli matrices, the number operator is  $\hat{n}_i = (\hat{\sigma}_i^z + \mathbb{1})/2$ . Performing the substitution  $n_i \rightarrow \hat{n}_i$ , we can establish a direct identification between the classical Ising model (3) and the (diagonal) Hamiltonian of a system of interacting spins located in  $\mathbf{q}_1, \dots, \mathbf{q}_M$ :

$$I^2 \rightarrow \hat{I}^2 := - \sum_{i=1}^N \Gamma_i \hat{n}_i + \sum_{i \neq j=1}^N V_{ij} \hat{n}_i \hat{n}_j \quad (10)$$

which we will be referring to as the *problem Hamiltonian*. If  $\mathcal{H}$  denotes the Hilbert space of the quantum system and  $\mathcal{B}(\mathcal{H})$  denotes its computational basis (the basis in which the operator in Eq. (10) is diagonal), then the ground state of (10) corresponds to the bitstring that minimizes (3), but it is now seen as the computational basis vector  $|e^*\rangle \in \mathcal{B}(\mathcal{H})$  that minimizes the expectation value of  $\hat{I}^2$ :

$$|e^*\rangle := \arg \min_{|e\rangle \in \mathcal{B}(\mathcal{H})} \langle e | \hat{I}^2 | e \rangle. \quad (11)$$

The position and number of the water molecules are then given by:

$$\begin{aligned} \mathcal{W} &:= \{\mathbf{q}_i \in \mathcal{Q} \mid \langle e^* | \hat{n}_i | e^* \rangle = 1\} \\ \mathcal{N} &:= \sum_{i=1}^M \langle e^* | \hat{n}_i | e^* \rangle. \end{aligned} \quad (12)$$

## V. SOLVING THE QUANTUM ISING PROBLEM WITH A LOCAL RYDBERG HAMILTONIAN

A system of neutral atoms coupled to an optical laser can be crafted in such a way as to evolve according to the following time-dependent Hamiltonian:

$$\hat{H}(t) = \sum_{i=1}^M \frac{\Omega_i(t)}{2} \hat{\sigma}_i^x - \sum_{i=1}^M \Delta_i(t) \hat{n}_i + \sum_{i < j=1}^M U_{ij} \hat{n}_i \hat{n}_j, \quad (13)$$

where  $\Omega_i(t)$  is the Rabi frequency of the driving laser on qubit  $i$ ,  $\Delta_i(t)$  is the detuning of the laser,  $U_{ij}$  is the interaction coefficient between Rydberg excitations:

$$U_{ij} = \frac{C_6}{|\mathbf{r}_{ij}|^6}, \quad (14)$$

and  $C_6$  is a physical constant. Before describing each term more in detail, one can immediately see the similarity between (10) and (13), and consequently why systems of neutral atoms might be particularly suited for solving such a problem. The biggest difference lies in the coefficient of the two-body term: In the Rydberg Hamiltonian  $U_{ij}$  decays as a power law, while in the problem Hamiltonian  $V_{ij}$  decays exponentially. However, one should keep in mind that (10) does not directly encode the proximity constraint between neighboring excitations, which is taken care of precisely by the  $r^{-6}$  part of the Rydberg Hamiltonian.

To see this point more explicitly, We compare the two functions  $e^{-r^2}$  and  $r^{-6}$ . At  $r = 0$  the former is capped at

a finite value, while the latter goes rapidly to infinity. Under the problem Hamiltonian  $\hat{I}^2$ , two nearby qubits  $i$  and  $j$  subject to strong one-body terms will be able to be excited simultaneously, since there always exist finite values of  $\Gamma_i, \Gamma_j$  that can overcome the interaction penalty  $V_{ij}$ . In the Rydberg Hamiltonian instead the physical repulsion between Rydberg excitations becomes so strong at close distance that whole sectors of the Hilbert space characterized by nearby excitations are effectively inaccessible. This means that using the Rydberg Hamiltonian (13) as a replacement for the problem Hamiltonian (10) would only be valid for those problem instances where low-energy states of the problem Hamiltonian have no closeby excitations. We expect, based on physical grounds that is indeed the case, since excitations represent the positions of physical objects, water molecules, that would naturally be subject to strong repulsion at close distance. This translates to the impossibility of having in  $I^2$  linear coefficients that prevail over the interactions at close distance. To understand better this point, we show in Fig. 2 the close-, medium-, and long-range interaction sectors. In the  $r \sim 0$  region (sector 1 in Fig. 2) the Rydberg interaction diverges, as opposed to the exponential interaction which is capped at a finite value. This is the largest numerical discrepancy between the two potentials; however, the numerical discrepancy is not necessarily relevant: At close range, both the Rydberg and the problem Hamiltonian are strong enough to prevent nearby excitations; they only differ in their absolute strength. The relevant sectors for the dynamics are the medium- and long-range ones (sector 2 in Fig. 2), where interactions are comparable to the linear coefficients of the Hamiltonian. For this reason the similarity in the behavior of the two Hamiltonians is better understood by neglecting the close-range sector 1.

Moving away from sector 1, we give now an indication of how well the medium- and long-range interactions match in both models. To this end, we selected a test problem with 14 qubits, and calculated the total interaction strength felt by each qubit using both interaction potential forms, i.e., for qubit  $i$ :  $\sum_j V_{ij}$  and  $\sum_j U_{ij}$ . However, we include in the sum only pairs at a distance lying in sector 2 of Fig. 2, in order to see past the close-distance discrepancy between the two interaction potentials discussed above. The plot is shown in Fig. 2. Although not matching perfectly, the qualitative behavior is nearly identical, reflecting the importance of the interaction strength being in general a decaying function of the distance between qubits, giving indication of the fact that the Rydberg Hamiltonian could indeed work as a replacement for the problem Hamiltonian.

Having control over the Rydberg Hamiltonian (13), the solution to the water placement problem can be found by using the quantum adiabatic algorithm [53]. In a nutshell, the strategy is to identify a path  $\hat{H}(t)$  in the space of admissible Hamiltonians for some time parameter  $t \in [0, T]$  such that the system is initialized in the ground state of  $\hat{H}(0)$  (which therefore has to be known and easy to prepare), while the only known property of the ground state of  $\hat{H}(T)$  is that it encodes the solution to a hard combinatorial optimization problem. If the evolution is performed slowly enough, then the system remains at all times in the instantaneous ground state of  $\hat{H}(t)$  for all  $t$ , and therefore measurements of the system at  $t = T$  give the solution to the optimization problem.

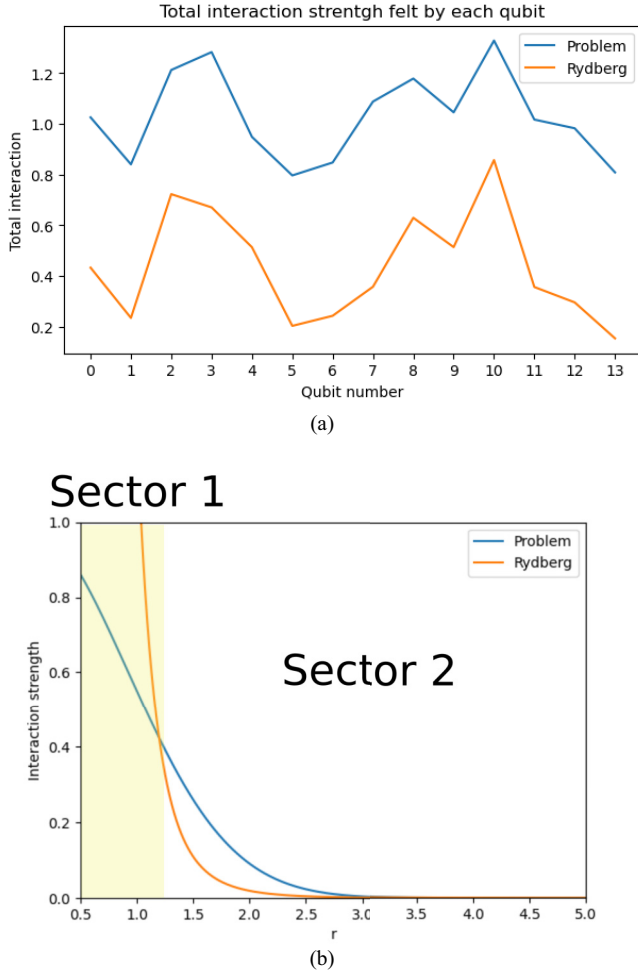


FIG. 2. (a) In a test system with 14 qubits, the total medium- and long-range interaction strength felt by each qubit is reported for the problem Hamiltonian (blue) and the Rydberg Hamiltonian (orange). The curves present a comparable behavior confirming the viability of using the Rydberg Hamiltonian as an approximation to the problem Hamiltonian. The discrepancy in the raw numerical value of the interaction strength in the two models is just a consequence of a scale difference between the two, which can be adjusted by adjusting the scale of the linear coefficients. (b) Plot of the corresponding problem and Rydberg interaction strength as a function of particle-particle distance; the threshold described in the text corresponds to the  $r$  value at the crossing between sector 1 (yellow region) and sector 2 areas.

We can exploit the quantum adiabatic evolution (QAE) algorithm by choosing  $\hat{H}(0)$  to be (13) with  $\Omega_i(0) = 0$ ,  $\Delta_i(0) = -c$  for some large positive constant  $c$ , so that the ground state of  $\hat{H}(0)$  is, to a very good approximation, the all-ground state  $|0\rangle^{\otimes M}$ . The final Hamiltonian  $\hat{H}(T)$  instead is chosen in such a way as to maximize the overlap between its low-energy spectrum and the one of the problem Hamiltonian  $\hat{I}^2$ . This can be achieved by mapping the one-body terms  $\Gamma_i$  in (3) to the final detunings  $\Delta_i(T)$  in (13). A schematic example of  $\Omega_i(t)$  and  $\Delta_i(t)$  is shown in Figs. 1(b) and 1(c).

The QAE algorithm, described in Algorithm 1, can be tested using the Pulser classical emulator [54].

We first tested this approach using synthetic densities. Results can be found in the Supplemental Material [39].

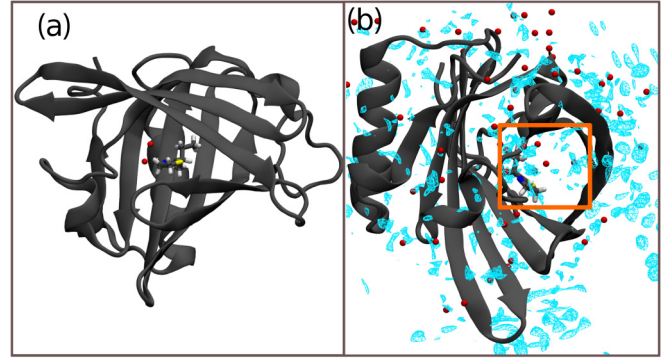


FIG. 3. (a) MUP-I protein complexed with the small molecule 2-sec-butyl-4,5-dihydrothiazole (PDB entry: 1i06.pdb), inside the protein binding pocket. The small molecule is represented in licorice style; the red spheres represent the two oxygen atoms of the crystal water molecules inside the protein binding pocket, while the protein scaffold is represented in gray. (b) The 3D-RISM density isosurface (isovalue 5.5) superposed to the protein and ligand cocrystal structure; the density is represented in wireframe style, in cyan. The VMD software is used for the visualization.

## VI. LOCAL ALGORITHM EMULATION USING 3D-RISM DENSITIES

To test the algorithm just presented, we compute the 3D-RISM solvent density within a real protein. We chose the major urinary protein (MUP-I) pocket, where a small ligand, the 2-sec-butyl-4,5-dihydrothiazole, is binding to the protein (see Fig. 3). This choice is motivated by the fact that the protein-ligand complex structure has been cocrystallized, its complete atomic structure being available from the Protein Data Bank (PDB). Interestingly, the crystal structure presents two structural (i.e., stable) water molecules in the vicinity of the ligand. Such water molecules are clearly visible in the 3D-RISM density, which exhibits high density spots in the same region as the positions of the oxygen atoms belonging to the two water molecules present in the protein crystal structure, near the ligand (see Fig. 3).

Despite the algorithm being well defined in any dimension, only a few current generation neutral atom quantum computers can operate in 3D [55], while the majority

ALGORITHM 1. QAE algorithm that solves the corresponding Ising problem in Eq. (10) using local lasers.

---

**inputs:** reference 2D density  $g(\vec{r})$  and set of qubit positions  $\mathcal{Q} := \{\vec{q}_i\}$ .

**outputs:** water molecule positions  $\mathcal{W}$  and their number  $\mathcal{N}$ .

1: **procedure** QAEAlg( $\vec{r}$ ),  $\{\vec{q}_i\}$

2: Build the adiabatic pulses  $\Omega(t)$ ,  $\Delta_i(t)$  as shown in Figs. 1(b) and 1(c), where the final detunings  $\Delta_i(T)$  are obtained from the mapping of Eq. (15) and  $\max_i \Omega(t)$  is defined to enforce the correct Rydberg blockade

3: Select  $|e^*\rangle$ , the basis state sampled from the quantum state after adiabatic evolution, that best represents  $g(\vec{r})$

4: **return**  $(\mathcal{W}, \mathcal{N})$ , the positions and number of water molecules in the protein cavity extracted from  $|e^*\rangle$  as defined in Eq. (12)

---

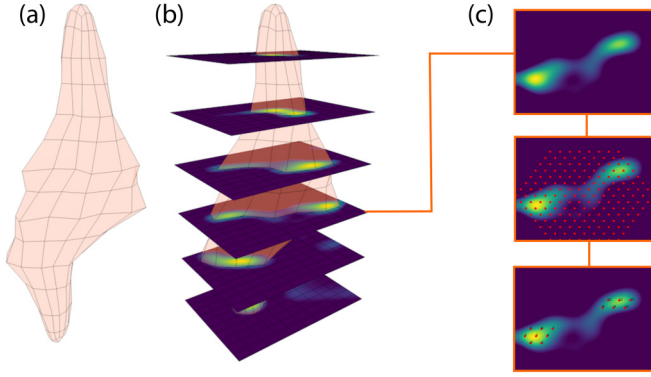


FIG. 4. (a) Example representation of a 3D-RISM density as a closed isosurface, with a define, constant, density value, as obtained starting from the AmberTools output, discretized on a given chosen grid of point. (b) Projection of the 3D-RISM density on 2D planes, namely slices, passing through the cross section of the density; each slice is a 2D density map, where the 3D-RISM density value discretized on the grid is projected on the plane according to a proximity criterion. (c) Process of selection of the qubits array on each 2D density slice: Starting from the slice, a regular mesh of traps is disposed uniformly on the density, to then select a limited number of locations where to place qubits according to the local density value; a threshold on the density is set, so that only traps close to a defined density or higher will be occupied.

operates in 2D. For this reason we cut the 3D-RISM density (which, by construction, is defined in three dimensions) into two-dimensional slices, and we apply the algorithm separately to each slice. For simplicity, we restrict ourselves to a small region of the protein around the two crystal water molecules. The plane corresponding to the first slice is defined by the axis connecting the two crystallographic water molecules and a third random direction. A series of slices is then produced from this first one in the direction normal to this plane, separated by a spacing of  $0.5 \text{ \AA}$ , for a total of six different slices. To smooth the 3D-RISM density in each slice, we apply a Laplacian of Gaussian filter as provided in the SciPy Python library [56], with a  $\sigma = 8$  (a.u.). The high-density regions of each smoothed-out slice are then covered with a qubit register. The process is depicted schematically in Fig. 4. Overall, this test implementation requires 4 to 14 qubits, which remains in the range of qubit count that can be classically emulated with high precision by a state vector solver with a realistic noise model. In order to mimic a physical implementation, we fit the registers in the same trap layout of Fig. 1(a), as shown in Fig. 5. The density slices with associated qubit registers are shown in Fig. 6.

For each register, we design an adiabatic pulse to find the low energy states of the Rydberg Hamiltonian (13). The final detunings  $\Delta_i(T)$  are computed from the  $\Gamma_i$  coefficients of each problem Hamiltonian by using the following heuristic mapping:

$$\tilde{\Delta}_i := \Gamma_i - \frac{1}{|N(i)|} \sum_{j \in N(i)} \Gamma_j, \quad \Delta_i(T) := \tilde{\Delta}_i \frac{\Delta_{\max}}{\max_j |\tilde{\Delta}_j|}, \quad (15)$$

where  $N(i)$  denotes nodes situated within a Rydberg blockade distance from node  $i$ , and  $\Delta_{\max}$  is a fixed value. The mapping

(15) can be understood as a two-step process. First, each  $\Gamma_i$  is shifted so that it is centered around the mean value of neighboring nodes, obtaining the intermediate detunings  $\tilde{\Delta}_i$ . Then, the  $\tilde{\Delta}_i$  are rescaled so that the maximum one corresponds to a certain fixed value  $\Delta_{\max}$ , obtaining the final value  $\Delta_i(T)$ .

We simulate the resulting adiabatic dynamics with Pulser, and sample the final state 1000 times. For each computational basis state sampled, we calculate its cost according to the problem Hamiltonian, and select as a solution the one with the minimal cost. The resulting sampling, where the basis states are represented as bitstrings, is shown in Fig. 7 for all six slices. The orange bar in each plot identifies the bitstring corresponding to the best solution for that slice. The fact the bitstring with the lowest cost is not always the most sampled one is a consequence of the differences between the low energy spectrum of problem and Rydberg Hamiltonians, which is due to the different nature of the interaction terms. The main limitation of the algorithm is indeed getting the two Hamiltonians to match as closely as possible, but on the other hand no pulse shaping or variational procedure is needed, making it in principle scalable to larger sizes without much of an overhead.

Once the best bitstring for each slice is identified, the coordinates of the excited qubits are mapped back to coordinates in the protein cavity, obtaining a list of water molecule positions. Since each 2D slice is treated independently, this procedure produces several water molecules, some of which might actually be the same water molecule replicated throughout several slices. To avoid this replica effect, we chose to employ a  $k$ -means algorithm as provided in the scikit-learn Python package [57], to bring down the final number of water molecules to just 2. The final prediction for the water molecule positions in the protein cavity is shown in Fig. 8(b). The agreement with the known crystallographic positions is good, with a maximum discrepancy of  $0.83 \text{ \AA}$ .

## VII. EXPERIMENTAL MEASUREMENT ON QPU EXPLOITING THE LOCAL ALGORITHM RESULTS

The emulations presented in the previous section correspond to the capabilities of a neutral atom QPU capable of local addressing. At the time the experiment was performed, no commercial neutral atom device was capable of local addressing. Only recently a neutral atom device with local addressing has been made available [52]. For this reason we tested here instead a device with global addressing.

In this simplified setting we exploit the results obtained in the previous section with the local algorithm to perform a grid-based search of the best solution for each register under a constant global pulse. Schematically, the experiment is comprised of the following steps:

(1) For each register presented in Fig. 6 perform a numerical emulation of the local algorithm to select the bitstring  $b_i$  that encodes the solution to the water placement problem for the  $i$ th register.

(2) Prepare the registers in the QPU in the all-ground state  $|0\rangle^{\otimes N}$ .

(3) Evolve the state under a constant pulse  $\Omega(t) = \Omega = \text{const.}$ ,  $\Delta(t) = \Delta = \text{const.}$  for a time  $T$ .



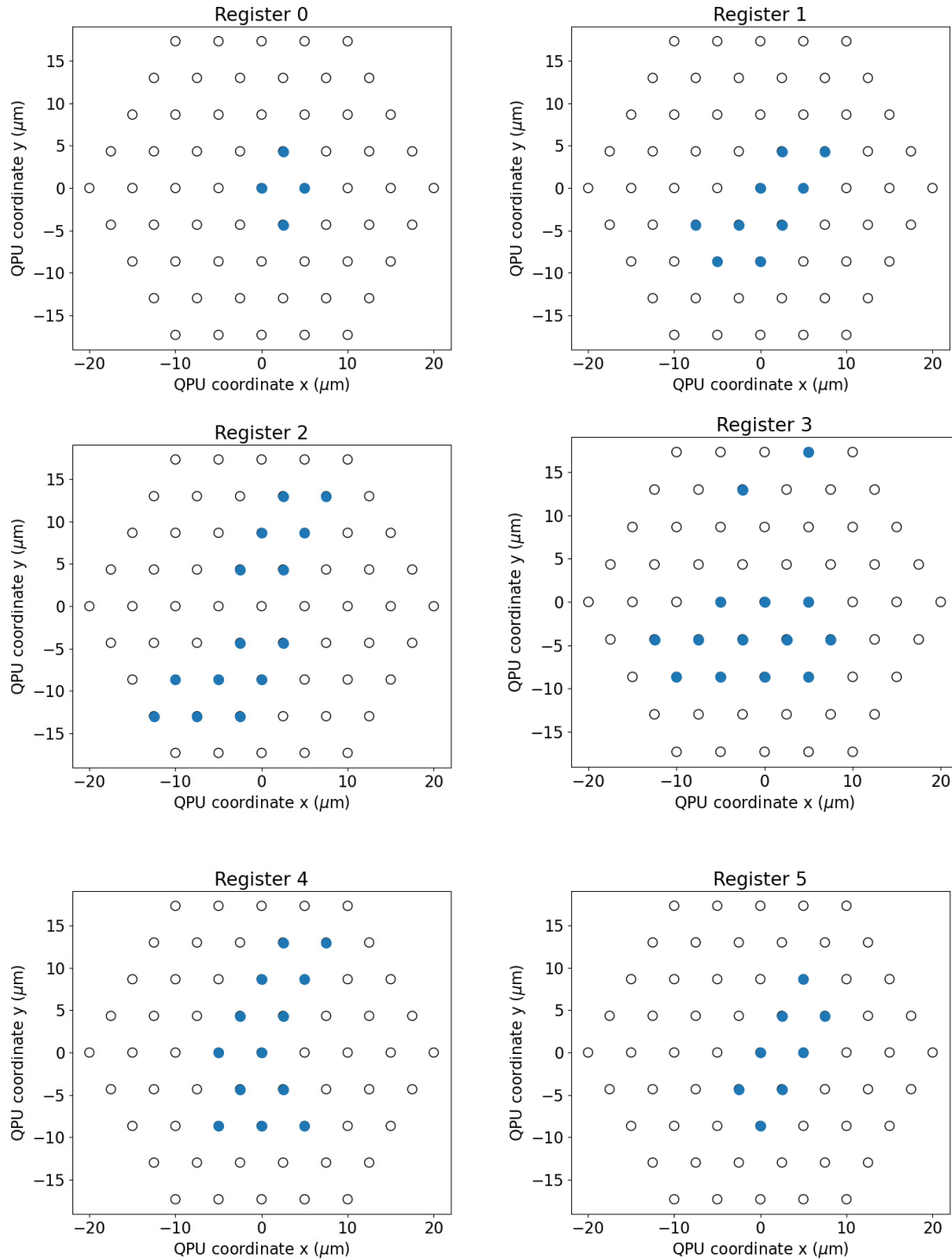


FIG. 5. Registers of Fig. 6, reshaped to fit in the trap layout available on the QPU [see Fig. 1(a)]. The units correspond to the physical dimension of the registers in micrometers.

(4) Estimate the expectation value of the projector  $|b_i\rangle \langle b_i|$  by averaging over 500 measurements in the computational basis.

(5) Repeat the procedure for several values of  $\Delta$  and  $T$  spanning a two-dimensional parameter space, keeping  $\Omega$  fixed to a value that would enforce nearest-neighbour blockade.

The parameter space window was pre-selected to explore a region where the probability of measuring the correct bitstring

is large and presents a good contrast. The parameters space being composed of  $8 \times 9 = 72$  points, each point obtained by attempting 500 measurements of the final state, gives a total of 36 000 measurements per register. The average execution time per register was measured to be 2 h and 47 min, resulting in a raw repetition rate (number of measurements per second) of 3.4 Hz. On 500 measurements attempted, on average 377 were successful, the unsuccessful ones being those for which

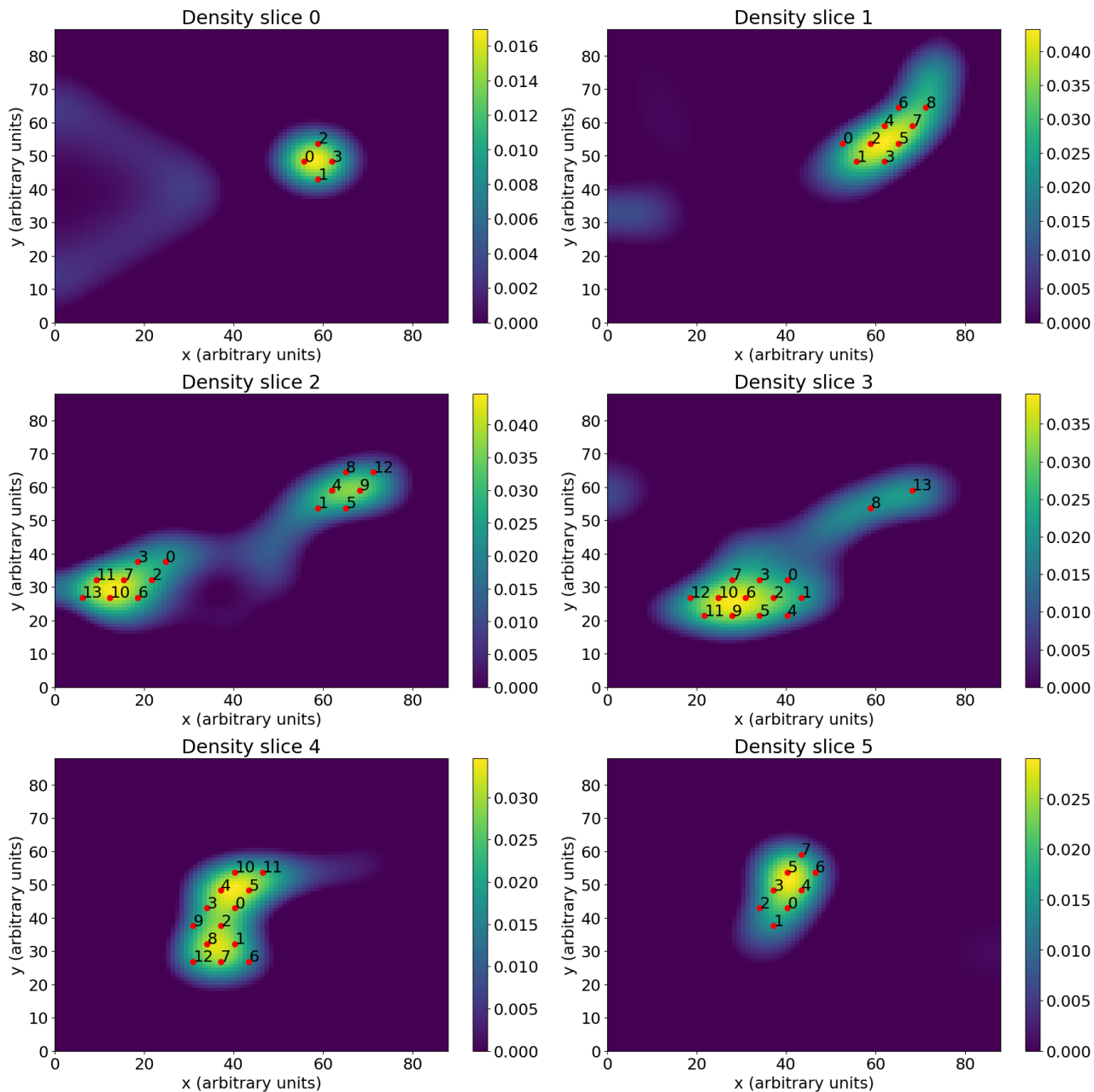


FIG. 6. Two-dimensional 3D-RISM density slices and qubit registers used in the local algorithm emulation. The registers vary in size from 4 to 14 qubits. The units displayed in the  $x$  and  $y$  directions correspond to the 3D-RISM discretization grid, and they are not representative of neither the size of the protein cavity in  $\text{\AA}$ , nor the size of the qubit registers in  $\mu\text{m}$ . To get an idea of the physical size of the registers, consider the spacing between neighboring qubits to be fixed at  $5 \mu\text{m}$ .

initial state preparation failed. This gives a state preparation success of 75.4%, and an effective repetition rate of 2.5 Hz.

The experimental measurements of the projector for all four registers are shown in Fig. 9, together with theoretical expectations given a reasonable estimate for false-positive and false-negative error rates on the machine. On a qualitative level, the experiments can be seen to be compatible with theoretical predictions obtained from emulation. These results indicate that meaningful solutions to the water placement problem belong to the space of physically allowed configura-

tions of the QPU. To our knowledge, such use of Rydberg physics had not been explored before in a real quantum physics experiment. Furthermore, the experiment shows that probability landscapes in this simplified setting can be accurately resolved by the machine, even in the absence of error correction or error mitigation techniques. This is a key result knowing the small magnitude of the involved probabilities that are capped at only a few percentages in certain systems.

In order to better understand and quantitatively validate the experimental data produced by the setup, we considered

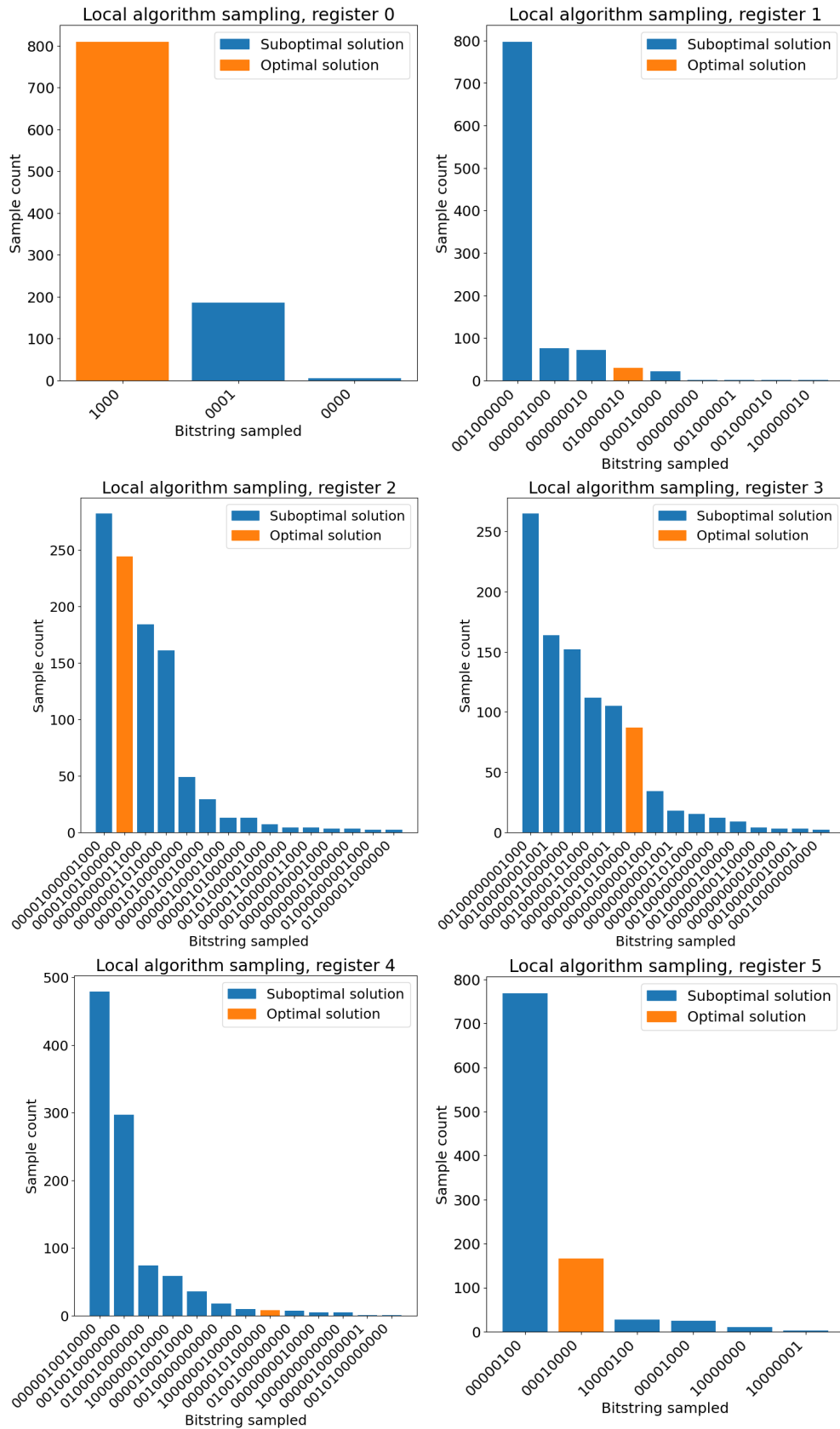


FIG. 7. Histogram of 1000 samples of the quantum state obtained with the local algorithm based on adiabatic evolution. Each plot corresponds to one of the registers shown in Fig. 6. The orange bar corresponds to the bitstring with the minimal cost among the ones that were sampled, and it is taken as the solution of the water placement problem for that register.

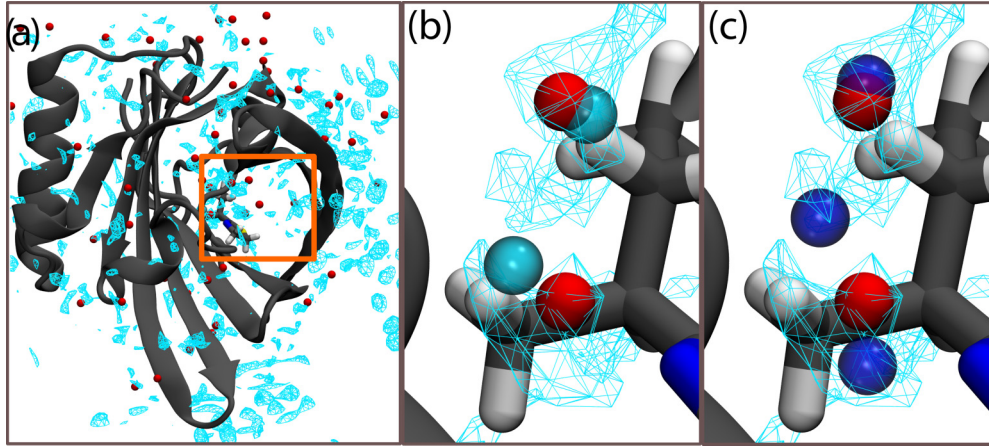


FIG. 8. (a) MUP-I protein complexed with the 2-sec-butyl-4,5-dihydrothiazole, represented as in Fig. 3. (b) Zoom inside the protein cavity: oxygens of the crystal water molecules position (red) compared with the placement performed by the QAE Algorithm 1, combining all the slices (cyan semitransparent spheres); the cyan wireframe represents a 3D-RISM isosurface. (c) Zoom inside the protein cavity: oxygens of the crystal water molecules position (red) compared with the placement performed by the VQA Algorithm 2, using a single slice (blue semitransparent spheres); the cyan wireframe represents a 3D-RISM isosurface.

a more sophisticated error model. Calibration of the control devices can only be achieved with finite precision, resulting in static uncertainties in global spacing of the atomic array ( $\approx 1\%$ ) or in the spatial homogeneity of  $\Omega$  on this array ( $\approx 4\%$ ). In addition, fluctuations of laser intensity induces a shot-to-shot variation of  $\Omega$  ( $\approx 5\%$ ). While the laser frequency can be set with high precision, variation in  $\Omega$  indirectly alters  $\Delta$ , resulting in small detuning shifts of the order of  $2\pi \times 0.06$  MHz. Decay processes are also taken into account by solving the Master equation with an effective decay rate  $\Gamma_{\text{eff}}/2\pi$ . Finally, the measurement phase is inherently flawed by several physical processes like atomic losses due to background-gas collisions or Rydberg state finite lifetime, whose effects can all be encompassed as first approximation into two detection error terms,  $\varepsilon$  and  $\varepsilon'$ . The various values of experimental and noise parameters are usually fitted by comparing the expectation value of easy-to-access observables between simulated and experimentally acquired data. For instance, emulating the dynamics described by the Hamiltonian (16) in presence or absence of noise processes and measuring the occupation  $\hat{n}_i$  at each site enables to estimate for the magnitude of the various error sources. An example of fitting curves at a given detuning  $\Delta$  for the first register is shown in Fig. 10. While the errors on calibration parameters are directly measured on the experiment, the remaining effective noise parameter are found to be  $\Gamma_{\text{eff}}/2\pi = 0.05$  MHz,  $\varepsilon = 2\%$ , and  $\varepsilon' = 18\%$ .

Overall, given the experimental errors sources that were discussed, the experimental data appear fully compatible with the emulated ones (see Figs. 9 and 10) confirming the viability of the proposed approach on an actual neutral atom device. We want to stress here that thanks to the quantum nature of the algorithm, two water molecules are guaranteed to never be placed too close to each other, a constraint that is imposed by hand in classical approaches such as Placevent [58] or GASol [59] with the risk of incurring in suboptimal local solutions or nonergodicity. This constitutes one of the main advantage of the algorithm.

### VIII. BEYOND A LOCAL ALGORITHM: VARIATIONAL ALGORITHM USING A GLOBAL RYDBERG HAMILTONIAN

As we discussed, technical limitations of the present hardware prevent us from using a local Hamiltonian as in Eq. (13). However, through emulation, we can still plan for an algorithm that will be able to run on more short-term devices using a global Rydberg Hamiltonian:

$$\hat{H}(t) = \frac{\Omega(t)}{2} \sum_{i=1}^M \hat{\sigma}_i^x - \Delta(t) \sum_{i=1}^M \hat{n}_i + \sum_{i<j=1}^M \frac{C_6}{r_{ij}^6} \hat{n}_i \hat{n}_j. \quad (16)$$

Such formulation allows to further approximate the problem Hamiltonian in Eq. (10), since the linear term in the qubits excitation is no longer local. A variational procedure is then established to minimize the cost function  $\langle \Psi | \hat{I}^2 | \Psi \rangle$ , where  $\hat{I}^2$  is the problem Hamiltonian (10) and  $|\Psi\rangle$  is a quantum state obtained from the quantum evolution under the physical Hamiltonian (16).

The variational procedure models the cost function landscape using a set of laser parameters  $\{\Omega_k(t), \Delta_k(t)\}$  indexed by an integer  $k$  as the result of a Bayesian search algorithm, for which the solution is computed as:

$$(\Omega(t), \Delta(t))^* = \arg \min_{\Omega_k(t), \Delta_k(t)} \{ \langle \Psi_k(T) | \hat{I}^2 | \Psi_k(T) \rangle \}, \quad (17)$$

where  $|\Psi_k(T)\rangle$  is the quantum state obtained from an evolution of duration  $T$  with parameters  $\Omega_k(t), \Delta_k(t)$ . This algorithm is therefore part of the VQA family. The full VQA developed for this scope is described in Algorithm 2.

The objective of the numerical procedure is to maximize the probability of sampling the basis state (11), from the optimized  $N$ -qubits wave function. This corresponds to finding one or more configurations of excited qubits best representing the 3D-RISM density distribution as a sum of gaussian distributions. The final output is still described by the quantities in Eq. (12).



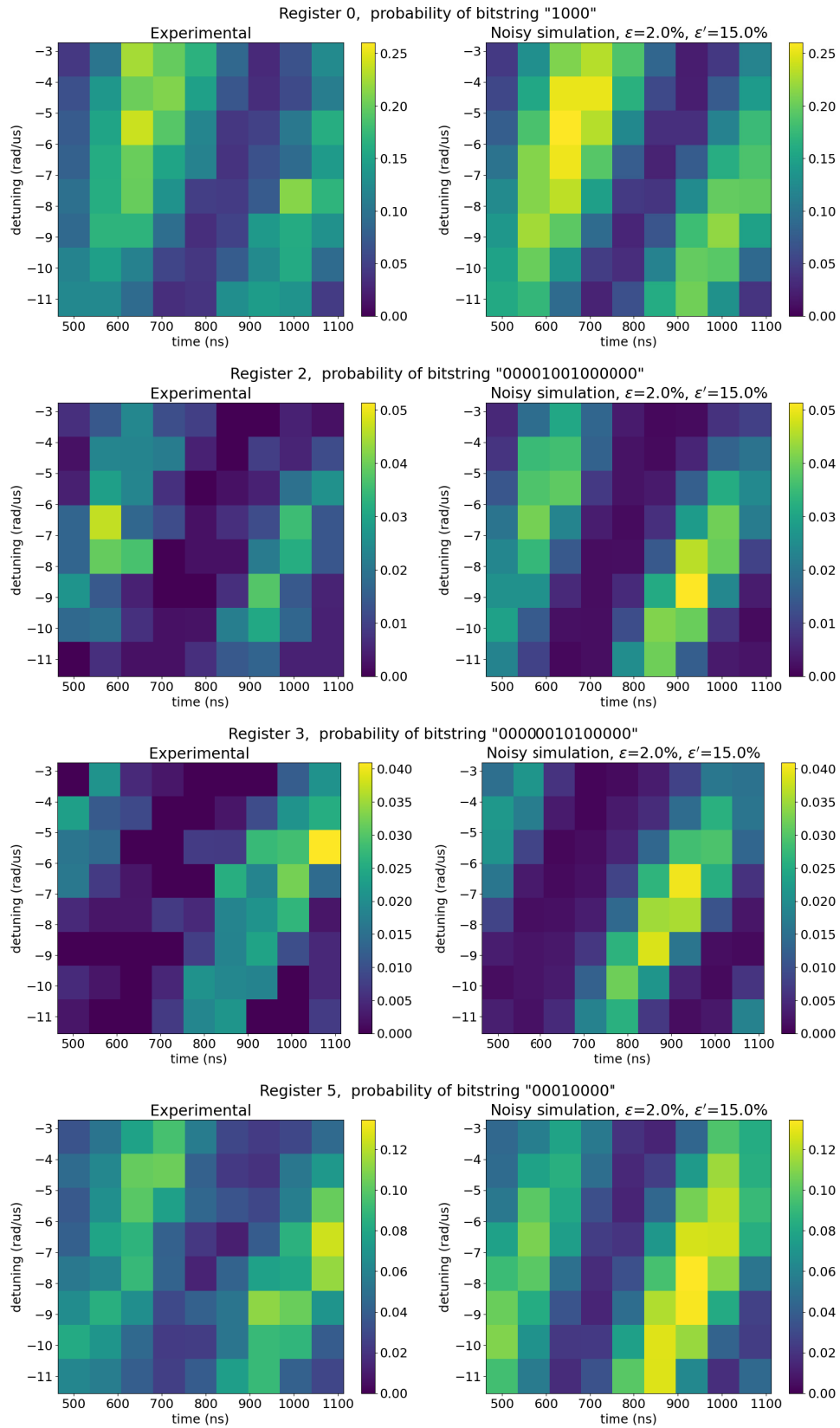


FIG. 9. Expectation value of the projector operator  $|b_i\rangle \langle b_i|$  associated to the solutions  $b_i$ ,  $i = 1, 2, 3, 4$  of the four of the 3D-RSM 2D slices. The area scanned is a two-dimensional region parametrized by pulse duration (horizontal axis) and detuning (vertical axis) of the driving laser. The plot on the left represents the experimental values, while the plot on the right represents the values obtained from a classical simulation of the Rydberg Hamiltonian with measurement errors  $\epsilon = 0.02$  and  $\epsilon' = 0.15$  related to the probability of false-positive and false-negative detection.

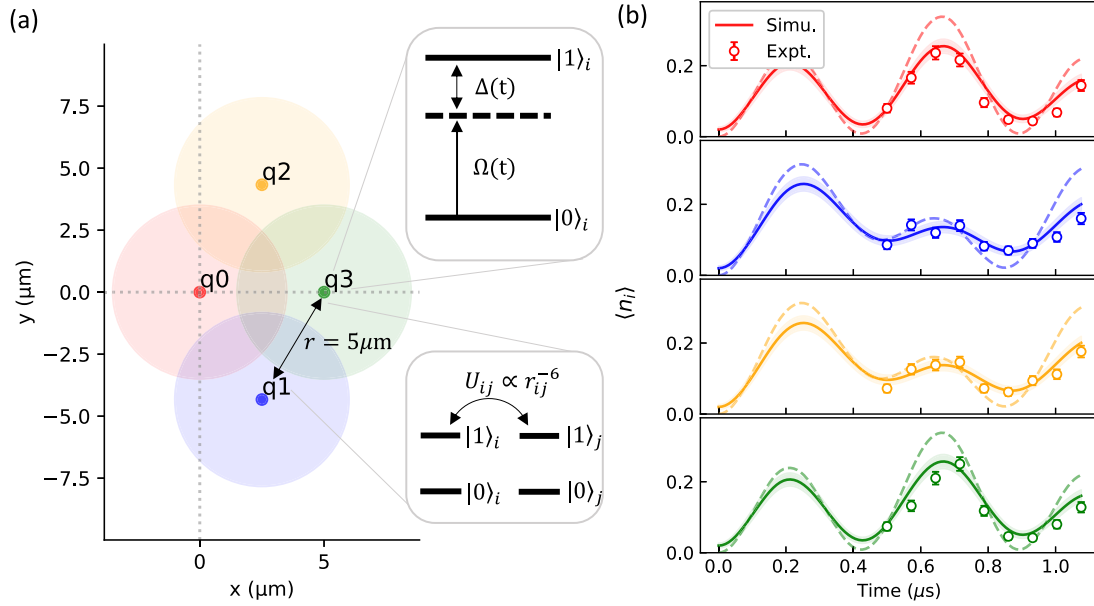


FIG. 10. (a) A laser pulse of constant  $\Omega$  and  $\Delta$  is applied for various times on the first atomic register comprising of 4 qubits (colored). (b) For a fixed detuning  $\Delta/2\pi = -0.87$  MHz, the expectation value of the occupation of each qubit site,  $\langle n_i \rangle$ , is plotted for noiseless simulation (dashed), noisy simulation (line) and experimental data (dot). The error bars and uncertainty regions are the standard deviation computed over  $N_{\text{shots}} = 500$  samples,  $\sigma_{n_i} = \sqrt{\langle n_i \rangle (1 - \langle n_i \rangle) / N_{\text{shots}}}$ .

To test the performances of the algorithm we use again simple synthetic densities to limit the number of qubits to employ, so that a cycle of optimization, emulating multiple times the quantum evolution of the system on the CPU, is performed for each test case. The results are reported in the Supplemental Material [39]. They show that the algorithm is able to give the correct positions of the Gaussian distributions centroids. Due to the high computational cost involved in the emulation of a numerical solution of the time-dependent

ALGORITHM 2. Hybrid quantum-classical VQA using global lasers, with cost function issued from the Ising model, namely  $\langle \Psi | \hat{I}^2 | \Psi \rangle$ .

- 
- inputs:** reference 2D density  $g(\vec{r})$  and set of qubit positions  $Q := \{\vec{q}_i\}$ .  $n_c$  and  $n_r < n_c$  are parameters for the Bayesian optimization procedure, being, respectively, the total number of cycles and the number of cycles initiated with randomized values of laser parameters  $\Omega$  and  $\Delta$ .
- outputs:** water molecule positions  $\mathcal{W}$  and their number  $\mathcal{N}$ .
- 1: **procedure** VQAlgog( $\vec{r}$ ),  $\{\vec{q}_i\}$ ,  $n_r$ ,  $n_c$
  - 2: **for**  $k = 1, \dots, n_c$  **do**
  - 3:   **if**  $k \leq n_r$  **then**
  - 4:      $\Omega_k(t)$ ,  $\Delta_k(t)$  are generated uniformly at random
  - 5:   **else**
  - 6:      $\Omega_k(t)$ ,  $\Delta_k(t)$  are obtained as trial parameters for the Bayesian minimization of the cost function in Eq. (17)
  - 7:    Perform quantum evolution with parameters  $(\Omega(t), \Delta(t))^*$ , the laser parameters that minimize the cost function according to Eq. (17), to obtain  $|\Psi^*(T)\rangle$
  - 8:    Select  $|e^*\rangle$ , the basis state sampled from  $|\Psi^*(T)\rangle$  that best represents  $g(\vec{r})$
  - 9: **return**  $(\mathcal{W}, \mathcal{N})$ , the positions and number of water molecules in the protein cavity extracted from  $|e^*\rangle$  as defined in Eq. (12)
- 

Shrödinger equation, we could not extend our tests towards larger systems.

Furthermore, We used as an additional test case one of the slices obtained from the 3D-RISM density in the MUP-I system already presented for the QAE approach. The sliced density, together with the qubit register, are shown in Fig. 11(a). We perform 50 cycles of Bayesian optimization, using 200 samples to represent each wave function produced by the corresponding set of laser parameters. The Gaussians used to dress the excited qubits, so to assign scores to the corresponding bitstring, have amplitudes proportional to the local value of the 3D-RISM density and an uniform variance  $\sigma^2 = 5$  (a.u.). The final state obtained using the best laser parameters is represented in the histogram in Fig. 11(b), in terms of the computational basis. From the resulting final wave function, we obtain the 3D coordinates of the oxygen atoms of the placed water molecules. In Fig. 8 we report the best water molecules configuration found by the algorithm (cyan atoms), corresponding to the best bitstring  $|101001\rangle$  found in the optimization. We want to stress here that also two excited qubits states, as  $|010100\rangle$  and  $|010001\rangle$  show good sampling probability [see corresponding histogram in Fig. 11(b)].

Additional emulated experiments are performed on all the other slices and the results are reported in the SM. The algorithm exhibits equivalent performances in all these cases. Again, we stress that this variational quantum algorithm, like the version using the local lasers, prevents two water molecules from being placed too close to each other, a constraint that is imposed by hand in classical approaches.

## IX. ALTERNATIVE FORMULATIONS AND QUANTUM ADVANTAGE

There exists a general connection between physical qubit Hamiltonians on one side and combinatorial optimization

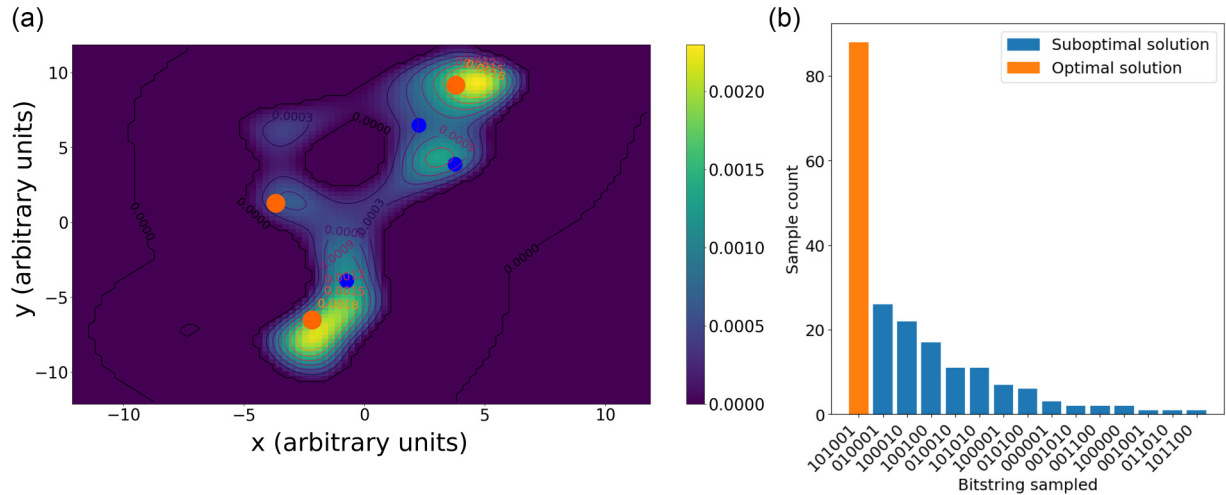


FIG. 11. (a) Representation of the smoothed 3D-RISM density slice used for testing VQA, with the six qubits, orange and blue circles, array used for the emulation on CPU through Pulser. The orange circles are used to highlight the solution obtained from the emulation, corresponding to the bitstring  $|101001\rangle$ . (b) Histograms reporting the six-qubit quantum state composition, in terms of basis states, as obtained from the VQA emulated on CPU, without noise, with most sampled bitstring highlighted.

problems formulated in a quadratic minimization framework on the other [60,61]. Solving combinatorial optimization problems through quantum annealing boils down to finding the ground state of the qubit Hamiltonian, which maps to the ground state of the combinatorial optimization problem, using algorithms such as the quantum adiabatic algorithm [53] or QAOA [62]. See Ref. [63] for a recent review on this subject for neutral atoms. Results regarding the performance of the quantum algorithms compared to classical approaches such as simulated annealing [64] are mostly numerical in nature, and not decisive. Depending on the specific problem instance, a quantum speedup is only possible at times [65–68].

The aim of the present work is to showcase a complete end-to-end solution from problem formulation and algorithm development to implementation on a real neutral atom QPU. We are not interested at this stage to show unequivocal quantum computational advantage over classical methods. A quantification and comparison of speed and accuracy of both classes of methods has not been investigated extensively enough for such a claim. A more comprehensive study in this sense might include an implementation of the present algorithm on other quantum architectures, as well as alternative quantum formulations.

As an example, the algorithm presented could be implemented on quantum annealers that allow configurability of all the coupling strengths, such as the D-Wave Advantage System [69]. However, the flexibility in atomic position, high scalability, and long coherence times of neutral atoms would arguably constitute a major asset in potential future large-scale implementations. Moreover, an important improvement of the present quantum algorithm is treating the full 3D problem instead of working with 2D slices. This would happen natively in neutral atoms once the technology will be mature enough, while it would introduce a large overhead in architectures with a rigid and incomplete 2D connectivity.

Among possible quantum alternatives to the present approach, we want to highlight [25,28], based on a mapping between docking and the maximum weighted clique problem

[70]. In light of the trivial mapping between the maximum weighted clique problem and the maximum weighted independent set problem on the complement graph [71], such approach is of potential interest in neutral atom quantum computing. In terms of alternative QUBO formulations of the docking problem, we highlight [30,31], which have been implemented on D-wave annealers.

On the other hand, a large variety of classical methods and software exists to address the water location problem, trying to tackle the intrinsic complexity of the task when dealing with complex systems [35,38,72]. To define a quantum advantage, also the performances of such methods need to be assessed and compared to a quantum alternative. Many of those rely on indirect approaches, to avoid explicit, direct atomistic, physics based dynamics simulations, to reduce the computational cost of sampling large phase spaces. At the best of our knowledge, available reviews are mostly comparative literature analysis, but no coherent studies on the performances of all or a subset of water location methods exist, so that a cross-comparison between each method could be possible, to fairly evaluate their accuracy and computational cost. More details can be found, by the interested reader, in the Appendix. We note here that a detailed benchmark of these methods could be the object of a further development of this project. Such a benchmark could be preliminary to a further investigation on the scaling of our quantum approach with respect to problem size.

## X. CONCLUSION AND PERSPECTIVES

In this work we presented two quantum algorithms able to sample equilibrium solvent configurations within proteins. We proposed a first fully local quantum adiabatic evolution version to be used on next generation devices, whereas a second version, belonging to the variational quantum algorithm family of algorithms, has also been introduced as a viable short-term alternative. This class of algorithms correspond to quantum versions of the 3D Reference Interaction Site

Model (Quantum-3D-RISM or Q-3D-RISM) since they use continuous solvent distributions as initial input. As a proof of concept, both algorithms have been shown to successfully be able to locate density maxima in nontrivial densities. In the case of the second VQA algorithm, a classical optimization is performed to find the best set of laser pulses, capable of producing the expected distribution of maxima. A test to confirm that solutions of the QAE emulations can be found in a real experiment with a suitable parameterization of the control fields was implemented on a neutral atom QPU, which represents to the best of our knowledge the first application of an analog quantum algorithm to the protein solvation problem. Presently, we limited ourselves to a qubit count of 14. This number was constrained by the present machine layout [see Figs. 1(a) and 5], by the state vector emulation capabilities (roughly 16 qubits can be emulated easily with Pulser) and the need of coupling Pulser and the actual QPU for the quantum adiabatic evolution model implementation. In that connection, concerning the local algorithm, future availability of time-dependent pulses and local addressing will totally remove the need of the Pulser emulation to prepare QPU data. In the same line, only time-dependent pulses are required for the VQA version of Quantum-3D-RISM: It should enable us to perform larger simulations at short term. Moreover, since the quantum versions of 3D-RISM comes with the native advantage of preventing water molecules from being placed on top of each other, it will be interesting to compare the performances of the classical and quantum versions of 3D-RISM since the next QPU implementation will provide us with the possibility to test at large scale the accuracy of such techniques. To do so, a careful study of the machine noise will be necessary but one key advantage of analog computing is to exhibit relatively constant noise levels with increasing system size making us optimistic about the prospect of QPU simulations encompassing a high number of qubits.

In conclusion, it will be possible, with further technological advancing, to prepare any molecular system with such algorithms in order to couple them to state-of-the-art molecular dynamics engines [73,74] for further properties evaluations. Overall, this Quantum-3D-RISM (Q-3D-RISM) family of algorithms demonstrates promises in predicting the solvation structure within biomolecular systems of interest for drug discovery applications, providing concrete use cases for the application of analog quantum computing in life sciences.

#### ACKNOWLEDGMENTS

This work was made possible thanks to the Pack Quantique grant from région Ile de France and GENCI, project ACQMED (Convention No. 20012758). Funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, project EMC2 (Grant No. 810367), is also acknowledged (J.-P.P) as funding from PEPR Epiq (ANR-22-PETQ-0007) and HQI programs. The authors thank the Fresnel team (Pasqal) for their support: J. Armougom, D. Benvenuti, L. Beguin, L. Bourachot, J. Briand, C. Briosne Frejaville, N. Carrez, T. Cartry, A. Charpentier, D. Claveau, L. Colin, G. Cournez, L. Couturier, J. De Hond, S. Desire, A. Dumas, S. Dutartre, P. Favier, G. Firenze, D. Kaczor, C. Hamot, G. Herce, J.

Heurtebize, V. Hully, B. Labarre, L. Lassabliere, H. Le Bars, L. Leclerc, A. Lindberg, G. Meriaux, F. Nambi, T. Pansiot, G. Pariente, J. Pellegrino, L. Ponsot, S. Roche, H. Silverio, G. Villaret, and J.-M. Wipff.

M.D. and D.L. performed simulations and contributed code; M.D., D.L., N.G., P.M., S.A., J.S., J.F., L.-P.H. and J.-P.P contributed methodology (theory); the Fresnel team and L. H. designed and performed the experiments; M.D., D.L., the Fresnel team, N.G., L.-P.H., L.H., and J.-P.P. analyzed data; M.D., D.L., part of the Fresnel team, L.-P.H., L.H., and J.-P.P. wrote the paper with the input of all authors; and L.H. and J.-P.P. designed the research.

#### APPENDIX A: QUANTUM COMPUTATIONAL RESOURCES

For the quantum computing sequences, we make use of Fresnel, an industrial neutral atom QPU made of single  $^{87}\text{Rb}$  atoms trapped in arrays of optical tweezers, conceived and manufactured by PASQAL. We operate the QPU in the ground-Rydberg qubit basis with global analog control [9]. The qubits are encoded into the ground state  $|0\rangle = |5S_{1/2}, F = 2, m_F = 2\rangle$  and a Rydberg state  $|1\rangle = |60S_{1/2}, m_J = 1/2\rangle$ . This effective two-level system is addressed with a two-photon laser excitation through an intermediate state  $|6P_{3/2}, F = 3, m_F = 3\rangle$ . The first (respectively second) photon excitation is generated by a 420-nm (1013-nm) laser beam. Details about the Pulser control software used to program the experiment can be found in Ref. [54].

#### APPENDIX B: CLASSICAL METHODS FOR HYDRATION SITES PREDICTION

An extended literature exists on the development of methods to predict stable water locations around molecular solutes. Here we report only few reviews on many available methods and the related numerical implementations [35,38,72]. Such methods have been described based on the different levels of approximations introduced to treat the problem, which is fundamentally related to predicting or evaluating molecular interactions and statistical sampling of distributions of interacting molecules. As an example, statistics or knowledge based methods [35,72] deeply relies on crystallography or biological data to apply a variety of machine learning, or machine learning like, techniques to avoid direct physics based simulations, which require potentially long simulations to reach a reliable representation of the statistical distribution of a large number of water molecules around the target solute. Some methods are able to assign a score to crystal waters to evaluate their stability in the same crystal position, or identify conserved crystal waters across related biological structures, i.e., proteins of the same family or experiments on closely related structures. Therefore such methods (see Refs. [38] and [35] for some specific examples) often do not find new stable hydration sites on a target structure. Even when they do predict stable positions of water (e.g., WAT-GEN [75]) they heavily rely on the database used to make the prediction, and the quality of the water crystal location, which is potentially questionable. Even if such approaches



are in principle less computationally expensive than standard sampling techniques, as Monte Carlo (MC) algorithms, and represent an extremely valuable resource, they just partially address the problem of water modeling and they just avoid the fundamental sampling problem. In between empirical and direct simulation methods as MD or MC, there exist grid based methods [35,75]. This class of methods are coupling molecular interactions simulation, like what is used in MD or MC, with a statistical mechanics theory to account for the statistical distribution of water molecules and their thermodynamics effects. These methods usually only partially retain the atomistic representation of the solvent. Grid-based approaches are numerically more tractable, since the heavy task of sampling the important molecular configurations at the protein-solvent interface is addressed through a statistical

mechanics approach, or optimization technique on randomly generated water guess distributions. Depending on the computational complexity of the method, different test dataset sizes have been used and reported in the literature to evaluate the method's performance: from several thousands of biological molecular structures (for empirical methods) to around 10 for MC sampling techniques. As a matter of fact, one can expect knowledge based and grid based methods may require from minutes to few hours on very limited computational resources, while MD or MC simulations may require several hours to few days on powerful machines. The accuracy is often calculated based on the number of crystal waters that each method is able to reproduce, eventually attributing to such waters a quality, based on some score or thermodynamics analysis.

- 
- [1] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, Quantum chemistry in the age of quantum computing, *Chem. Rev.* **119**, 10856 (2019).
- [2] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [3] C. Feniou, M. Hassan, D. Traoré, E. Giner, Y. Maday, and J.-P. Piquemal, Overlap-ADAPT-VQE: Practical quantum chemistry on quantum computers via overlap-guided compact Ansätze, *Commun. Phys.* **6**, 192 (2023).
- [4] C. Feniou, B. Claudon, M. Hassan, A. Courtat, O. Adjoua, Y. Maday, and J.-P. Piquemal, Greedy gradient-free adaptive variational quantum algorithms on a noisy intermediate scale quantum computer, [arXiv:2306.17159](https://arxiv.org/abs/2306.17159).
- [5] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum algorithms, *Rev. Mod. Phys.* **94**, 015004 (2022).
- [6] 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, Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry, *Nat. Commun.* **14**, 1952 (2023).
- [7] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
- [8] W. Sennane, J.-P. Piquemal, and M. J. Rančić, Calculating the ground-state energy of benzene under spatial deformations with noisy quantum computing, *Phys. Rev. A* **107**, 012416 (2023).
- [9] L. Henriët, L. Beguin, A. Signoles, T. Lahaye, A. Browaeys, G.-O. Reymond, and C. Jurczak, Quantum computing with neutral atoms, *Quantum* **4**, 327 (2020).
- [10] P. Scholl, M. Schuler, H. J. Williams, A. A. Eberharter, D. Barredo, K.-N. Schymik, V. Lienhard, L.-P. Henry, T. C. Lang, T. Lahaye, and A. B. Läuchli, Quantum simulation of 2D antiferromagnets with hundreds of Rydberg atoms, *Nature (London)* **595**, 233 (2021).
- [11] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, *Nat. Commun.* **9**, 4812 (2018).
- [12] R. Baron, P. Setny, and J. A. McCammon, Water in cavity-ligand recognition, *J. Am. Chem. Soc.* **132**, 12091 (2010).
- [13] J. Michel, J. Tirado-Rives, and W. L. Jorgensen, Prediction of the water content in protein binding sites, *J. Phys. Chem. B* **113**, 13337 (2009).
- [14] H.-J. Woo, A. R. Dinner, and B. Roux, Grand canonical Monte Carlo simulations of water in protein environments, *J. Chem. Phys.* **121**, 6392 (2004).
- [15] Y. Levy and J. N. Onuchic, Water mediation in protein folding and molecular recognition, *Annu. Rev. Biophys. Biomol. Struct.* **35**, 389 (2006).
- [16] D. Chandler and H. C. Andersen, Optimized cluster expansions for classical fluids. II. Theory of molecular liquids, *J. Chem. Phys.* **57**, 1930 (1972).
- [17] D. Chandler, Derivation of an integral equation for pair correlation functions in molecular fluids, *J. Chem. Phys.* **59**, 2742 (1973).
- [18] D. Chandler, Equilibrium structure and molecular motion in liquids, *Acc. Chem. Res.* **7**, 246 (1974).
- [19] D. Beglov and B. Roux, An integral equation to describe the solvation of polar molecules in liquid water, *J. Phys. Chem. B* **101**, 7821 (1997).
- [20] A. Kovalenko, Three-dimensional rism theory for molecular liquids and solid-liquid interfaces, in *Molecular Theory of Solvation, Understanding Chemical Reactivity*, edited by F. Hirata (Springer Netherlands, Dordrecht, 2003), pp. 169–275.
- [21] D. Roy and A. Kovalenko, Biomolecular simulations with the three-dimensional reference interaction site model with the Kovalenko-Hirata closure molecular solvation theory, *Int. J. Mol. Sci.* **22**, 5061 (2021).
- [22] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, *Phys. Rev. E* **58**, 5355 (1998).
- [23] J. Brooke, D. Bitko, T. F. Rosenbaum, and G. Aeppli, Quantum annealing of a disordered magnet, *Science* **284**, 779 (1999).
- [24] Y.-H. Deng, Y.-C. Gu, H.-L. Liu, S.-Q. Gong, H. Su, Z.-J. Zhang, H.-Y. Tang, M.-H. Jia, J.-M. Xu, M.-C. Chen, J. Qin,

- L.-C. Peng, J. Yan, Y. Hu, J. Huang, H. Li, Y. Li, Y. Chen, X. Jiang, L. Gan *et al.*, Gaussian boson sampling with pseudo-photon-number-resolving detectors and quantum computational advantage, *Phys. Rev. Lett.* **131**, 150601 (2023).
- [25] S. Yu, Z.-P. Zhong, Y. Fang, R. B. Patel, Q.-P. Li, W. Liu, Z. Li, L. Xu, S. Sagona-Stophel, E. Mer, S. E. Thomas, Y. Meng, Z.-P. Li, Y.-Z. Yang, Z.-A. Wang, N.-J. Guo, W.-H. Zhang, G. K. Tranmer, Y. Dong, Y.-T. Wang *et al.*, A universal programmable Gaussian boson sampler for drug discovery, *Nat. Comput. Sci.* **3**, 839 (2023).
- [26] J. Bao, Z. Fu, T. Pramanik, J. Mao, Y. Chi, Y. Cao, C. Zhai, Y. Mao, T. Dai, X. Chen, X. Jia, L. Zhao, Y. Zheng, B. Tang, Z. Li, J. Luo, W. Wang, Y. Yang, Y. Peng, D. Liu *et al.*, Very-large-scale integrated quantum graph photonics, *Nat. Photon.* **17**, 573 (2023).
- [27] H. Zhu, H. Chen, S. Li, T. Chen, Y. Li, X. Luo, F. Gao, Q. Li, L. Zhou, M. F. Karim, X. Shang, F. Duan, H. Cai, L. K. Chin, L. C. Kwek, X. Zhang, and A.-Q. Liu, A dynamically programmable quantum photonic microprocessor for graph computation, *Laser Photon. Rev.* **18**, 2300304 (2024).
- [28] X. Tan, H. Song, Y. Ji, H. Tang, Y.-Y. Fang, X.-Y. Xu, Y.-Y. Li, X.-K. Li, K.-D. Zhu, and X.-M. Jin, Scalable and programmable three-dimensional photonic processor, *Phys. Rev. Appl.* **20**, 044041 (2023).
- [29] Z. Bian, F. Chudak, R. Israel, B. Lackey, W. G. Macready, and A. Roy, Discrete optimization using quantum annealing on sparse Ising models, *Front. Phys.* **2**, 56 (2014).
- [30] M. Pandey, T. Zaborniak, H. Melo, A. Galda, and V. K. Mulligan, Multibody molecular docking on a quantum annealer, [arXiv:2210.11401](https://arxiv.org/abs/2210.11401).
- [31] E. Triuzzi, R. Mengoni, D. Bonanni, D. Ottaviani, A. Beccari, and G. Palermo, Molecular docking via weighted subgraph isomorphism on quantum annealers, [arXiv:2405.06657](https://arxiv.org/abs/2405.06657).
- [32] R. Harris, Y. Sato, A. J. Berkley, M. Reis, F. Altomare, M. H. Amin, K. Boothby, P. Bunyk, C. Deng, C. Enderud, S. Huang, E. Hoskinson, M. W. Johnson, E. Ladizinsky, N. Ladizinsky, T. Lanting, R. Li, T. Medina, R. Molavi, R. Neufeld *et al.*, Phase transitions in a programmable quantum spin glass simulator, *Science* **361**, 162 (2018).
- [33] A. D. King, S. Suzuki, J. Raymond, A. Zucca, T. Lanting, F. Altomare, A. J. Berkley, S. Ejtemaee, E. Hoskinson, S. Huang, E. Ladizinsky, A. J. R. MacDonald, G. Marsden, T. Oh, G. Poulin-Lamarre, M. Reis, C. Rich, Y. Sato, J. D. Whittaker, J. Yao *et al.*, Coherent quantum annealing in a programmable 2,000 qubit Ising chain, *Nat. Phys.* **18**, 1324 (2022).
- [34] D. B. Kitchen, H. Decornez, J. R. Furr, and J. Bajorath, Docking and scoring in virtual screening for drug discovery: Methods and applications, *Nat. Rev. Drug Discov.* **3**, 935 (2004).
- [35] M. L. Samways, R. D. Taylor, H. E. Bruce Macdonald, and J. W. Essex, Water molecules at protein–drug interfaces: Computational prediction and analysis methods, *Chem. Soc. Rev.* **50**, 9104 (2021).
- [36] D. Bucher, P. Stouten, and N. Triballeau, Shedding light on important waters for drug design: Simulations versus grid-based methods, *J. Chem. Inf. Model.* **58**, 692 (2018).
- [37] A. Wlodawer, W. Minor, Z. Dauter, and M. Jaskolski, Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures, *FEBS J.* **275**, 1 (2008).
- [38] E. Nittinger, F. Flachsenberg, S. Bietz, G. Lange, R. Klein, and M. Rarey, Placement of water molecules in protein structures: From large-scale evaluations to single-case examples, *J. Chem. Inf. Model.* **58**, 1625 (2018).
- [39] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevResearch.6.043020> for additional details on 3D-RISM, test of both versions of the algorithm on synthetic density generated as Gaussian mixtures with known Gaussian components, with details on the numerical tests performed, a full derivation of the water placement problem in the form of a QUBO.
- [40] T. Imai, A. Kovalenko, and F. Hirata, Solvation thermodynamics of protein studied by the 3d-rism theory, *Chem. Phys. Lett.* **395**, 1 (2004).
- [41] T. Imai, R. Hiraoka, A. Kovalenko, and F. Hirata, Locating missing water molecules in protein cavities by the three-dimensional reference interaction site model theory of molecular solvation, *Proteins: Struct. Funct. Bioinf.* **66**, 804 (2007).
- [42] D. J. Sindhikara and F. Hirata, Analysis of biomolecular solvation sites by 3d-rism theory, *J. Phys. Chem. B* **117**, 6718 (2013).
- [43] M. C. Stumpe, N. Blinov, D. Wishart, A. Kovalenko, and V. S. Pande, Calculation of local water densities in biological systems: A comparison of molecular dynamics simulations and the 3d-rism-kh molecular theory of solvation, *J. Phys. Chem. B* **115**, 319 (2011).
- [44] S. Dasgupta, Learning mixtures of Gaussians, in *Proceedings of the 40th Annual Symposium on Foundations of Computer Science* (IEEE, Los Alamitos, CA, 1999), pp. 634–644.
- [45] A. T. Kalai, A. Moitra, and G. Valiant, Disentangling Gaussians, *Commun. ACM* **55**, 113 (2012).
- [46] A. Moitra and G. Valiant, Settling the polynomial learnability of mixtures of Gaussians, in *Proceedings of the IEEE 51st Annual Symposium on Foundations of Computer Science* (IEEE, Los Alamitos, CA, 2010), pp. 93–102.
- [47] J. Anderson, M. Belkin, N. Goyal, L. Rademacher, and J. R. Voss, The more, the merrier: The blessing of dimensionality for learning large Gaussian mixtures, in *Proceedings of the 27th Conference on Learning Theory (COLT'14)*, edited by M. Balcan, V. Feldman, and C. Szepesvári, JMLR Workshop and Conference Proceedings Vol. 35 (JMLR, 2014), pp. 1135–1164.
- [48] A. Bakshi, I. Diakonikolas, H. Jia, D. M. Kane, P. K. Kothari, and S. S. Vempala, *Robustly Learning Mixtures of  $k$  Arbitrary Gaussians* (ACM Press, New York, 2022) p. 1234–1247.
- [49] I. Diakonikolas, G. Kamath, D. Kane, J. Li, A. Moitra, and A. Stewart, Robust estimators in high-dimensions without the computational intractability, *SIAM J. Comput.* **48**, 742 (2019).
- [50] N. Bourgeois, B. Escoffier, V. Paschos, and J. M. M. van Rooij, Fast Algorithms for max independent set, *Algorithmica* **62**, 382 (2013).
- [51] S. Ebadi, A. Keesling, M. Cain, T. T. Wang, H. Levine, D. Bluvstein, G. Semeghini, A. Omran, J.-G. Liu, R. Samajdar, X.-Z. Luo, B. Nash, X. Gao, B. Barak, E. Farhi, S. Sachdev, N. Gemelke, L. Zhou, S. Choi, H. Pichler, S.-T. Wang, M. Greiner, V. Vuletić, and M. D. Lukin, Quantum optimization of maximum independent set using Rydberg atom arrays, *Science* **376**, 1209 (2022).

- [52] J. Wurtz, A. Bylinskii, B. Braverman, J. Amato-Grill, S. H. Cantu, F. Huber, A. Lukin, F. Liu, P. Weinberg, J. Long, S.-T. Wang, N. Gemelke, and A. Keesling, Aquila: QuEra's 256-qubit neutral-atom quantum computer, [arXiv:2306.11727](https://arxiv.org/abs/2306.11727).
- [53] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, A quantum adiabatic evolution algorithm applied to random instances of an np-complete problem, *Science* **292**, 472 (2001).
- [54] H. Silvério, S. Grijalva, C. Dalyac, L. Leclerc, P. J. Karalekas, N. Shammah, M. Beji, L.-P. Henry, and L. Henriët, Pulser: An open-source package for the design of pulse sequences in programmable neutral-atom arrays, *Quantum* **6**, 629 (2022).
- [55] M. Kim, K. Kim, J. Hwang, E.-G. Moon, and J. Ahn, Rydberg quantum wires for maximum independent set problems, *Nat. Phys.* **18**, 755 (2022).
- [56] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in Python, *Nat. Methods* **17**, 261 (2020).
- [57] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, Scikit-learn: Machine learning in Python, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [58] D. J. Sindhikara, N. Yoshida, and F. Hirata, Placevent: An algorithm for prediction of explicit solvent atom distribution-application to HIV-1 protease and F-ATP synthase, *J. Comput. Chem.* **33**, 1536 (2012).
- [59] L. Fusani, I. Wall, D. Palmer, and A. Cortes, Optimal water networks in protein cavities with GASol and 3D-RISM, *Bioinformatics* **34**, 1947 (2018).
- [60] A. Lucas, Ising formulations of many NP problems, *Front. Phys.* **2**, 5 (2014).
- [61] A. Das and B. K. Chakrabarti, Colloquium: Quantum annealing and analog quantum computation, *Rev. Mod. Phys.* **80**, 1061 (2008).
- [62] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, The quantum approximate optimization algorithm and the Sherrington-Kirkpatrick model at infinite size, *Quantum* **6**, 759 (2022).
- [63] C. Dalyac, L. Leclerc, L. Vignoli, M. Djellabi, W. da Silva Coelho, B. Ximenez, A. Dareau, D. Dreon, V. E. Elfving, A. Signoles, L.-P. Henry, and L. Henriët, Graph algorithms with neutral atom quantum processors, *Eur. Phys. J. A* **60**, 177 (2024).
- [64] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing, *Science* **220**, 671 (1983).
- [65] M. Cain, S. Chattopadhyay, J.-G. Liu, R. Samajdar, H. Pichler, and M. D. Lukin, Quantum speedup for combinatorial optimization with flat energy landscapes, [arXiv:2306.13123](https://arxiv.org/abs/2306.13123).
- [66] A. P. Young, S. Knysch, and V. N. Smelyanskiy, First-order phase transition in the quantum adiabatic algorithm, *Phys. Rev. Lett.* **104**, 020502 (2010).
- [67] M. Guidetti and A. P. Young, Complexity of several constraint-satisfaction problems using the heuristic classical algorithm walksat, *Phys. Rev. E* **84**, 011102 (2011).
- [68] I. Hen and A. P. Young, Exponential complexity of the quantum adiabatic algorithm for certain satisfiability problems, *Phys. Rev. E* **84**, 061152 (2011).
- [69] C. C. McGeoch and P. Farré, The D-wave advantage system: An overview TECHNICAL REPORT, <https://api.semanticscholar.org/CorpusID:222355754>.
- [70] L. Banchi, M. Fingerhuth, T. Babej, C. Ing, and J. M. Arrazola, Molecular docking with Gaussian boson sampling, *Sci. Adv.* **6**, eaax1950 (2020).
- [71] M. Pelillo, Heuristics for maximum clique and independent set, in *Encyclopedia of Optimization*, edited by C. A. Floudas and P. M. Pardalos (Springer US, Boston, MA, 2009), pp. 1508–1520.
- [72] K. Mitusińska, A. Raczynska, M. Bzówka, W. Bagrowska, and A. Góra, Applications of water molecules for analysis of macromolecule properties, *Comput. Struct. Biotechnol. J.* **18**, 355 (2020).
- [73] L. Lagardère, L.-H. Jolly, F. Lipparini, F. Aviat, B. Stamm, Z. F. Jing, M. Harger, H. Torabifard, G. A. Cisneros, M. J. Schnieders, N. Gresh, Y. Maday, P. Y. Ren, J. W. Ponder, and J.-P. Piquemal, Tinker-hp: A massively parallel molecular dynamics package for multiscale simulations of large complex systems with advanced point dipole polarizable force fields, *Chem. Sci.* **9**, 956 (2018).
- [74] O. Adjoua, L. Lagardère, L.-H. Jolly, A. Durocher, T. Very, I. Dupays, Z. Wang, T. J. Inizan, F. Célerse, P. Ren, J. W. Ponder, and J.-P. Piquemal, Tinker-hp: Accelerating molecular dynamics simulations of large complex systems with advanced point dipole polarizable force fields using gpus and multi-gpu systems, *J. Chem. Theory Comput.* **17**, 2034 (2021).
- [75] H.-H. Bui, A. J. Schiewe, and I. S. Haworth, Watgen: An algorithm for modeling water networks at protein-protein interfaces, *J. Comput. Chem.* **28**, 2241 (2007).