Gate-free state preparation for fast variational quantum eigensolver simulations

Oinam Romesh Meitei* and Nicholas J. Mayhall†
Department of Chemistry, Virginia Tech, Blacksburg, VA 24061, USA

Bryan T. Gard,* George S. Barron, Sophia E. Economou, and Edwin Barnes
Department of Physics, Virginia Tech, Blacksburg, VA 24061, USA

David P. Pappas
Physics Department, University of Colorado, Boulder, CO 80309 and
National Institute of Standards and Technology, Boulder, CO 80305

I. SUPPLEMENTARY METHODS

A. Analytical Gradients

Instead of using finite difference approximations to obtain gradients for the pulse amplitudes, we use the analytical gradients derived below to speed up the classical simulations considerably. To derive the derivatives of the cost function with respect to the pulse amplitudes, we begin by rearranging the control Hamiltonian in the interaction frame, Equation 7 from the main text, and multiplying it by $\tau = \frac{T}{N}$. T is the total evolution time and N the number of Trotter steps. We also define the time of each Trotter step as $t_j = j\tau$, for $j = 0, ..., N$. We have

$$\hat{G}_t = \sum_k \Omega_k(t)\hat{\mathcal{O}}(t)\tau \tag{1}$$

$$\hat{\mathcal{O}}(t) = e^{i\hat{H}_D t}(e^{i\nu_k t}\hat{a}_k + e^{-i\nu_k t}\hat{a}_k^\dagger)e^{-i\hat{H}_D t} \tag{2}$$

The time evolution ansatz, Equation 6 in the main text, can be computed using a Trotter-Suzuki expansion as follows

$$|\psi\rangle = \prod_{j=1}^{\overleftarrow{N}} e^{-i\hat{G}_j} |\psi_0\rangle \tag{3}$$

such that the expectation value of the molecular Hamiltonian (estimate of the molecular energy) is given by,

$$E = \langle\psi| \hat{H}_M |\psi\rangle = \langle\psi_0| \prod_{j=1}^{\overrightarrow{N}} e^{i\hat{G}_j} \hat{H}_M \prod_{j=1}^{\overleftarrow{N}} e^{-i\hat{G}_j} |\psi_0\rangle \tag{4}$$

We have utilized $\overleftarrow{\prod}$ and $\overrightarrow{\prod}$ to denote products of elements that have increasing and decreasing indices from right to left respectively. Following this construction,

$$\prod_{j=1}^{\overleftarrow{3}} \hat{o}_j = \hat{o}_3\hat{o}_2\hat{o}_1, \qquad \prod_{j=1}^{\overrightarrow{3}} \hat{o}_j = \hat{o}_1\hat{o}_2\hat{o}_3 \tag{5}$$

The Energy derivative with respect to the pulse amplitude, $\Omega_k(t)$ at trotter step t can be expressed as:

*These two authors contributed equally
†Electronic address: nmayhall@vt.edu

$$\frac{\partial E}{\partial \Omega_k(t)} = \langle\psi_0| \prod_{j=1}^{\overrightarrow{N}} e^{i\hat{G}_j} \hat{H}_M \prod_{j=t}^{\overleftarrow{N}} e^{-i\hat{G}_j} \left(-i\mathcal{O}(t)\tau\right) \prod_{j=1}^{\overleftarrow{t-1}} e^{-i\hat{G}_j} |\psi_0\rangle + \tag{6}$$

$$\langle\psi_0| \prod_{j=1}^{\overrightarrow{t}} e^{i\hat{G}_j} (i\mathcal{O}(t)\tau) \prod_{j=t+1}^{\overrightarrow{N}} e^{i\hat{G}_j} \hat{H}_M \prod_{j=1}^{\overleftarrow{N}} e^{-i\hat{G}_j} |\psi_0\rangle \tag{7}$$

$$= -i\tau \left(\langle\psi| \hat{H}_M \prod_{j=t}^{\overleftarrow{N}} e^{-i\hat{G}_j} \mathcal{O}(t) |\psi_{t-1}\rangle - \langle\psi_{t-1}| \mathcal{O}(t) \prod_{j=t}^{\overrightarrow{N}} e^{i\hat{G}_j} \hat{H}_M |\psi\rangle \right) \tag{8}$$

$$= 2\tau Im \langle\psi| \hat{H}_M \prod_{j=t}^{\overleftarrow{N}} e^{-i\hat{G}_j} \mathcal{O}(t) |\psi_{t-1}\rangle \tag{9}$$

$$= 2\tau Im \langle\sigma_{t-1}| \mathcal{O}(t) |\psi_{t-1}\rangle \tag{10}$$

where,

$$|\psi_{t-1}\rangle = \prod_{j=1}^{\overleftarrow{t-1}} e^{-i\hat{G}_j} |\psi_0\rangle \tag{11}$$

$$|\sigma_{t-1}\rangle = \prod_{j=t}^{\overrightarrow{N}} e^{i\hat{G}_j} \hat{H}_M |\psi\rangle \tag{12}$$

Using the above Equation 9, the energy derivative with respect to the pulse amplitude can be computed at each Trotter step. Even with a large number of Trotter steps, the gradient evaluation only cost about 2.5 times the energy evaluation (Equation 4) per qubit. This is achieved by evaluating the fully evolved state, $|\psi\rangle$ at first and then applying a single matrix exponential at a time on $|\psi_0\rangle$ and $\hat{H}_M |\psi\rangle$, i.e. evaluating Equation 9 using Equations 10 and 11 for $t = N$ at first, then for $t = N - 1$, $t = N - 2$, upto $t = 1$.

Because the gradients are obtained for the amplitude at each Trotter step, pulses of arbitrary shapes can be easily optimized. For the square pulse restrictions used in this work, the gradient with respect to the amplitude for a given time segment (square pulse) can be easily obtained by summing all the gradients evaluated within the time window.

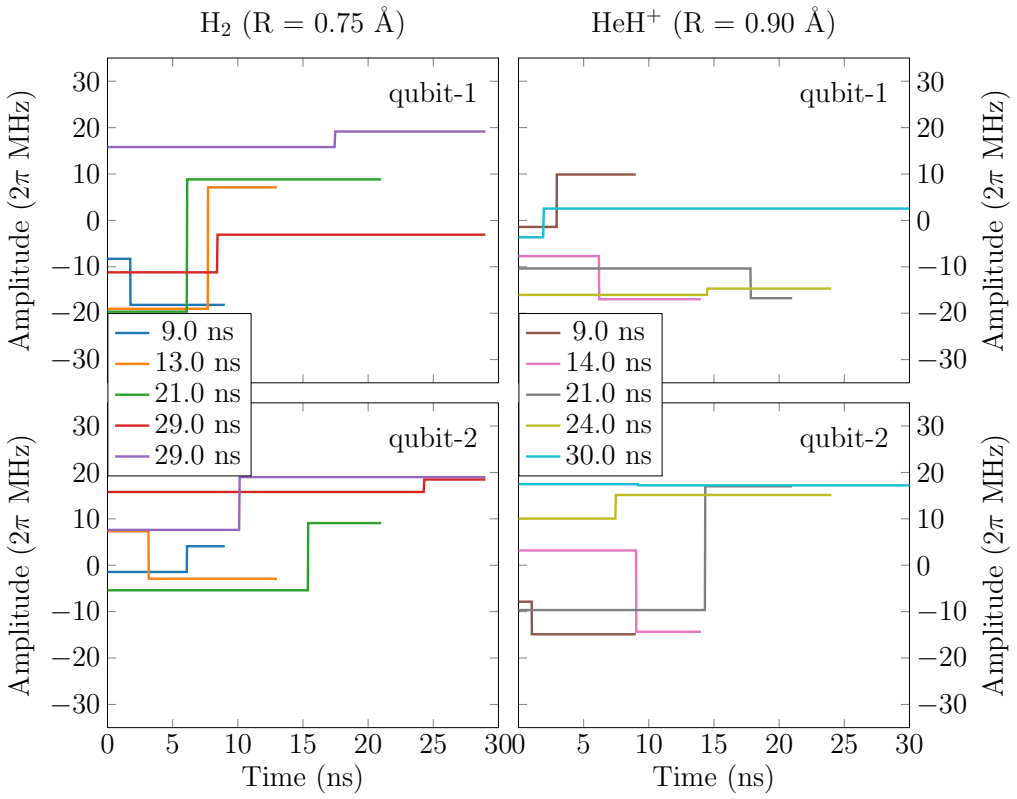
B. Noisy Parameters

There are two different scales for our optimization parameters. The amplitude of the drives are restricted to $20 \times 2\pi$ MHz which sets the scale for all frequency terms to tens of MHz. The scale for the timing of pulses is set by the smallest time scale that is typically discernable in practical wave generators, one nanosecond. This then means that our two scales differ by a relative factor of 100 (tens of MHz compared to ns). Therefore when choosing the noisy distributions for the random selection of parameters, we apply these scaling factors to the corresponding standard deviations in order to maintain a clear implementation of noise, across all parameters. We can then describe each noise distribution by its standard deviation and relate it back to the relevant time scale. In this case the standard deviations we use in the main text are related to the relevant parameters by $\sigma = \{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\} \propto \{0.1, 1, 10, 100\}$ MHz or $\{0.01, 0.1, 1, 10\}$ ns, depending on the relevant noisy parameter. As discussed in the text, for the distributions we investigate, this then means that the optimal pulses are highly robust and only begin to deviate significantly once the noise is of the same order as the magnitude of the individual parameters. For this data, Nelder-Mead was used as the optimizer.

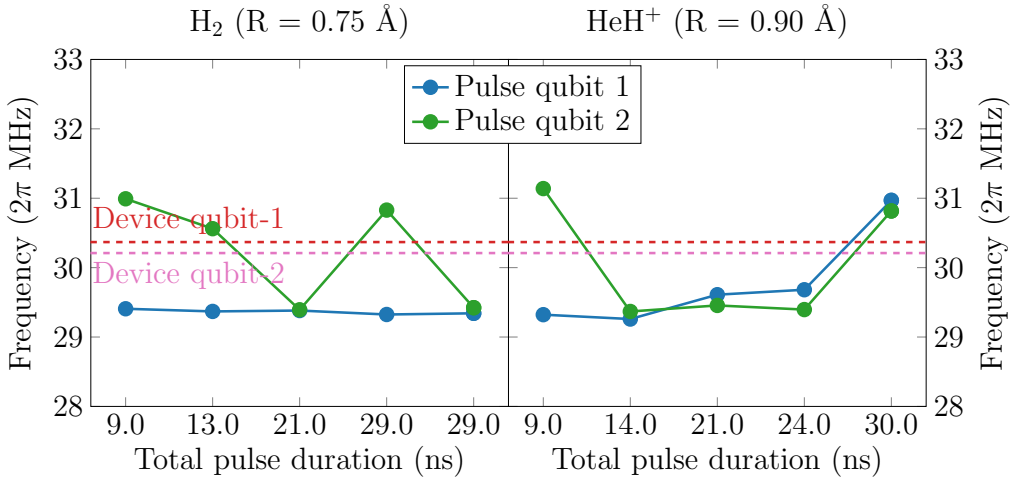
II. SUPPLEMENTARY DISCUSSION

A. Optimal square pulse shapes

In this section, the optimal pulse shapes discussed in Section IV.E. of the main text is illustrated.



Supplementary Figure 1: Amplitudes of the square pulses corresponding to Supplementary Figure 5 and 6 of the main text for H_2 and HeH^+ resp. on transmon qubit-1 (top) and qubit-2 (bottom). The optimal square pulses with different total pulse durations are used to drive the transmon qubits to the ground state of H_2 (left) and HeH^+ (right).



Supplementary Figure 2: Frequencies of the square pulses corresponding to Supplementary Figure **1**. The optimal square pulses with different total pulse durations are used to drive the transmon qubits to the ground state of H_2 (left) and HeH^+ (right).

B. Results using unnormalized energy

In this section, we present numerical results obtained from using unnormalized energy in the objective function of `ctrl-VQE`. Furthermore, the objective function here is the correlation energy defined by $E_{FCI} - E_{HF}$. This is simply

obtained by subtracting the Hartree Fock energy from the diagonal elements of the molecular Hamiltonian in the qubit basis. The objective function, Equation 8 in the main text, then becomes,

$$E(\Omega_n(t), \nu_n) = \left\langle \psi^{\text{trial}} \right| \hat{H}_{\text{corr}}^{\text{molecule}} \left| \psi^{\text{trial}} \right\rangle \tag{13}$$

where,

$$\hat{H}_{\text{corr}}^{\text{molecule}} = \hat{H}^{\text{molecule}} - IE_{HF} \tag{14}$$

The results obtained from using the unnormalized energy in the objective function of `ctrl-VQE` is tabulated in Table I. Here, the simple square pulses as well the device parameters described in the main text were used. As mentioned in the main text, the results are comparable to the results obtained therein.

TABLE I: Performance of `ctrl-VQE` using the unnormalized energy in the objective function: Error in energy from the full CI method, leakage to higher states, total pulse duration and overlap with the exact full CI states for selected bond distances of the H_2 molecule. Square pulses with two time segments were used.

Bond Distance (Angstrom)	Energy Error (Hartree)	Leakage (%)	Duration (ns)	Overlap (%)
0.5	1.45×10^{-5}	0.06	10	99.88
0.6	1.89×10^{-5}	0.06	10	99.88
0.7	1.18×10^{-5}	0.03	12	99.94
0.8	2.11×10^{-5}	0.04	18	99.91
0.9	1.24×10^{-5}	0.02	19	99.96

C. Comparison of square and gaussian control pulses

Here, we present a set of numerical results obtained with the simple square pulse used in the main text and a sum of two gaussian pulses. The parameterized gaussian pulse is given by,

$$\Omega_k(t) = \sum_{j=1}^{2} c_j \, \exp\left(-\frac{(t - t_j)^2}{2\sigma_j^2} \right), \tag{15}$$

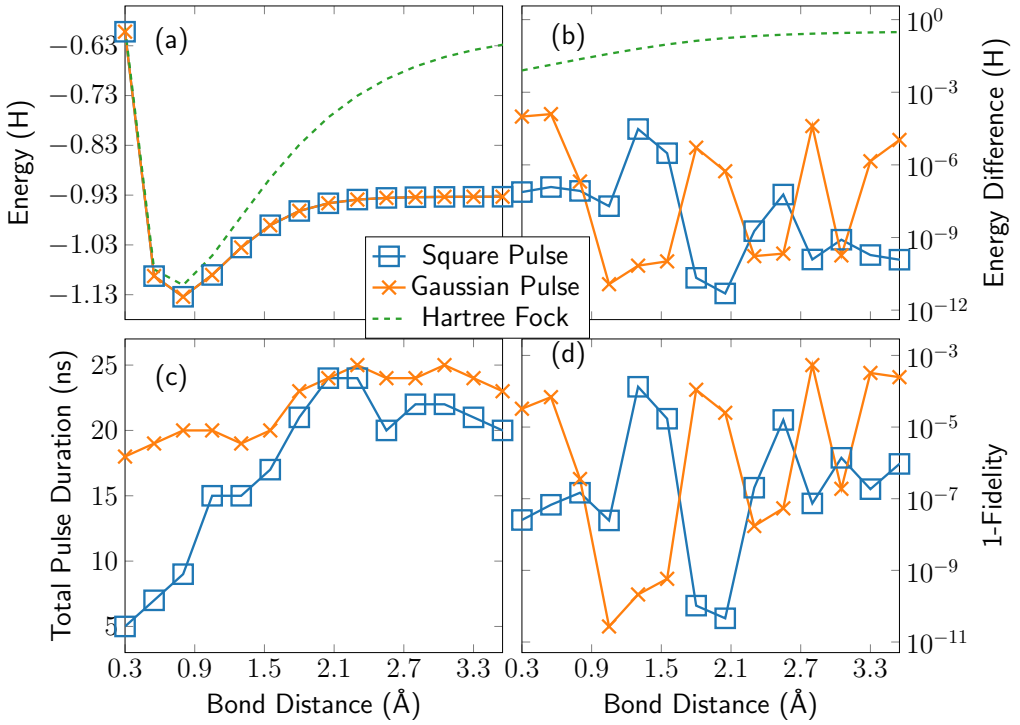
where t_j, σ_j^2 are the mean and variance of each Gaussian and c_j is the amplitude. With N transmons and j number pulses on each transmon, we have $3Nj + N$ parameters to optimize for the gaussian pulse. Note that there is no constraint on the individual amplitudes c_j, but there is the same 20 MHz constraint placed on the resulting sum. The driving frequency ν_k are constrained to $(\omega_k - 5) \leq \nu_k \leq (\omega_k + 5)$ where ω_k, ν_k are in units of 2π GHz.

D. Large detuning device parameters

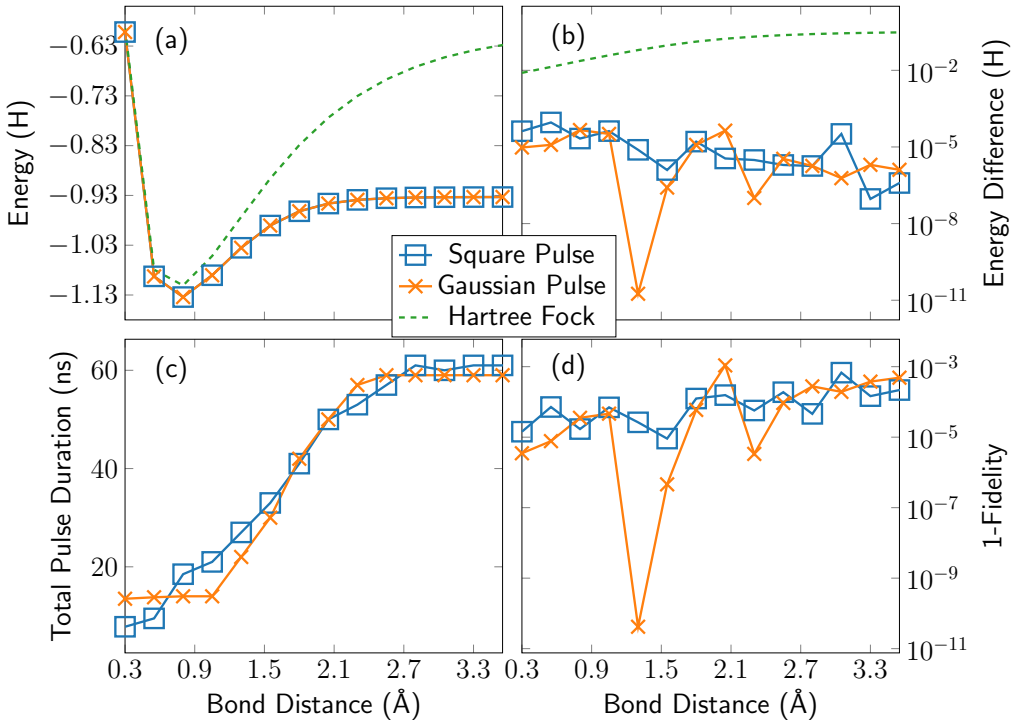
In this section, we present results obtained with a different set of device parameters appearing in Equation 3 in the main text. This set of device parameters which we refer to as the large detuning parameters is given in Table II.

ω_1	ω_2	δ_1	δ_2	g
5.3272	5.0761	0.3291	0.2616	0.0204

TABLE II: Large detuning device parameters. Note the difference in $|\omega_1 - \omega_2|$ as compared to the one used in the main text. All units are 2π GHz.



Supplementary Figure 3: Results of ctrl-VQE applied to the dissociation of H_2 when using the Δ_s parameters of Table 1 of the main text. a) Dissociation of H_2 as the bond distance between H and H is stretched. b) The energy difference between the state found with ctrl-VQE and the exact ground state. c) Total pulse durations for each Gaussian and square pulse across the various bond distances. d) One minus the overlap or 1-Fidelity between the resulting ctrl-VQE state and the exact ground state.



Supplementary Figure 4: Result of ctrl-VQE using the large detuning device parameters in Table II: (a) Bond dissociation of H_2 obtained using square pulse with two time segments and sum of two gaussians on each transmon for a two-transmon system. (b) Difference in energies between ctrl-VQE and FCI along the bond dissociation of H_2 (c) Total pulse durations along the bond dissociation of H_2. (d) One minus the overlap or 1-Fidelity between the resulting ctrl-VQE state and the exact FCI ground state.

In Supplementary Figure **4**, the difference in the pulse duration at the equilibrium and at far bond distance is significantly different from the one presented in the main text. Note the difference in $|\omega_1 - \omega_2|$ in the device parameters used in the main text. For a detailed characterisation, we examine pulse shapes at three different bond distances of H_2 molecule viz, near the equilibrium (0.75 angstrom), intermediate (2.0 angstrom) and far bond distance (5.0 angstrom). Here, the simple square pulse with two time segments have been considered. Several optimizations were carried out at these bond distances with the initial guess of the pulse parameterizations being random numbers. A range of 10 to 50 ns pulse duration were chosen at each bond distances and hundreds of optimizations were carried out for each pulse duration. The numerical optimizations have numerous local minima (suggested by the multiple solutions that exists) and so this strategy was carried out to obtain as many optimal pulses as possible.

Near the equilibrium bond distance, the shortest pulse duration found was 14.0 ns. Short pulses (14.0 ns) could be found for the large bond distances. However, the overlap of the resulting states with the exact states, however, were about 50% at the far bond distances. This likely indicates that the solution is actually a superposition state of the ground singlet and the lowest triplet state. The two states become degenerate at the large bond distances. This degeneracy can be distinguished since these states differ by their total spin eigenvalue. By including the penalty term dependent on the total spin discussed in Section II E, we can always obtain the target singlet ground state, even if there are degenerate states present. This is also demonstrated in Supplementary Figure **4** (d). As for the intermediate bond distance, the shortest pulse duration found with a low leakage was 50.0 ns. Pulses with a duration of 32.0 ns do exist, however, the leakage was 46.0%. Including higher levels (four) on each transmon resulted with a pulse of duration 33.0 ns but with a leakage of 33.0%. The significant difference in the pulse duration at the intermediate bond distance is suggestive of the control pulse optimization reflecting the complex electronic structure involved in the bond dissociation of H_2 at the intermediate bond distance.

E. Objective function penalties

Our generic objective function is the measurement of energy of the chemical Hamiltonians, $\langle H \rangle$. However, in practice this simple form may encounter issues due to inaccuracies brought about by leakage outside of the qubit subspace and degenercies with degenerate states. To mitigate both of these processes we add simple (convex) penalty

terms to the generic objective function. In order to mitigate leakage outside of the qubit subspace, we add a penalty term proportional to the population outside the qubit subspace. This then modifies the objective function to be

$$\langle H\rangle + \alpha(1 - \||\psi\rangle_s\|)^2, \tag{16}$$

where $|\psi\rangle_s$ is the state projected onto the qubit subspace and α is a constant penalty factor. Additionally we include a second penalty term proportional to the total spin of the time dependent state. In total then the objective function is given by,

$$\langle H\rangle + \alpha(1 - \||\psi\rangle_s\|)^2 + \beta(\langle s\rangle - s_t)^2, \tag{17}$$

where s_t is the target total spin eigenvalue and s is the total spin operator for our reduced space. For the case of a singlet state, $s_t = 0$. The total spin operator for this reduced space is given in the computational basis as,

$$s = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}. \tag{18}$$

We then have an objective function which mimizes the energy along with minimizing leakage and targeting the singlet state. Note that s is modified from the textbook total spin operator since we have reduced the original four spin-orbital problem to two qubits. In principle this addition to the objective function requires the additional measurement of the total spin of the state. However, this measurement contains the same order of terms ($\mathcal{O}(N^4)$) that are required to measure the chemical hamiltonian. In addition, the total spin operator typically contains Pauli strings that also appear in the chemical hamiltonian itself and therefore only require a small number (less than $\mathcal{O}(N^4)$) of new terms to measure.