Supplemental Materials: Digital quantum simulation of open quantum systems using quantum imaginary time evolution

Hirsh Kamakari,[1] Shi-Ning Sun,[1] Mario Motta,[2] and Austin J. Minnich[1,*]

[1]*Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA*

[2]*IBM Quantum, IBM Research Almaden, San Jose, CA 95120, USA*

(Dated: November 16, 2021)

I. DERIVATION OF THE QITE LINEAR SYSTEM FOR ALGORITHM II

Here we derive the QITE linear systems which need to be solved to obtain the time evolution of the density operator ansatz. Consider $|\rho\rangle = \sum_x p_x T|x\rangle \otimes \overline{T}|x\rangle$, with T unitary. The complex time propagator is the same as in the vectorization method,

$$X(t) = \exp\left(\left[-i\mathbb{I}\otimes H + iH^\top \otimes \mathbb{I} + \sum_k (\overline{L_k}\otimes L_k - \frac{1}{2}\mathbb{I}\otimes(L_k^\dagger L_k) - \frac{1}{2}(L_k^\top \overline{L_k})\otimes\mathbb{I})\right] t\right). \tag{S1}$$

Trotterizing results in

$$X(t) = \left\{\left[\exp\left(iH^\top\tau\right)\otimes\exp\left(-iH\tau\right)\right]\prod_k\left[\exp\left(-\frac{L_k^\top\overline{L_k}\tau}{2}\right)\otimes\exp\left(-\frac{L_k^\dagger L_k\tau}{2}\right)\right]\exp\left(\overline{L_k}\otimes L_k\tau\right)\right\}^n + \mathcal{O}\left(\tau^2 n\right) \tag{S2}$$

with $\tau = t/n$. Using the identity $\overline{\exp(-iA)} = \exp(iA^\top)$ for A Hermitian and $\overline{\exp(B)} = \exp(\overline{B})$ for arbitrary B, the propagator can be rewritten as

$$X(t) = \left\{[U\otimes\overline{U}]\prod_k [V_k\otimes\overline{V_k}]W_k\right\}^n + \mathcal{O}\left(\tau^2 n\right) \tag{S3}$$

with $U := \exp(iH^\top\tau)$, $V_k := \exp(-L_k^\top\overline{L_k}\tau/2)$, and $W_k := \exp(\overline{L_k}\otimes L_k\tau)$. It is immediate that evolution with $U\otimes\overline{U}$ preserves the ansatz, as $(U\otimes\overline{U})\sum_x p_x T|x\rangle\otimes\overline{T}|x\rangle =$

* aminnich@caltech.edu

$\sum_x p_x(UT)|x\rangle \otimes (\overline{UT})|x\rangle$. The term $V \otimes \overline{V}$ also preserves the ansatz, but is an imaginary time evolution with Hamiltonian $L_k^\top \overline{L_k}/2$, and so requires a modified QITE algorithm, described below, for implementing $\exp(-L_k^\top \overline{L_k}\tau/2)T|x\rangle$. Due to the non-unitarity of V_k, we expect that in addition to a unitary evolution of the state, the weights p_x will also evolve in time. The final term, $W_k = \exp(\overline{L_k} \otimes L_k\tau)$, does not preserve the ansatz, and we use a modified version of QITE, described below, to effectively apply W_k while preserving the form of ansatz.

A. Implementing $V_k \otimes \overline{V_k}$ via a QITE adaptation

Under real time evolution by the non-unitary operator $V_k \otimes \overline{V_k}$, the evolution of $|\rho\rangle$ can be expressed as

$$V_k \otimes \overline{V_k} \sum_x p_x T|x\rangle \otimes \overline{T}|x\rangle = \sum_x (p_x + q_x) \exp(iA)T|x\rangle \otimes \exp(-i\overline{A})\overline{T}|x\rangle + O(\tau^2), \tag{S4}$$

with $q_x \in \mathbb{R}$ and A a Hermitian operator with $\|A\|_2 = \mathcal{O}(\tau)$. Defining $B := (1/2)L_k^\top \overline{L_k}$, we then have, to first order in τ,

$$\exp(-\tau B) \otimes \exp(-\tau\overline{B}) \sum_x p_x T|x\rangle \otimes \overline{T}|x\rangle = \sum_x (p_x + q_x) \exp(iA)T|x\rangle \otimes \exp(-i\overline{A})\overline{T}|x\rangle. \tag{S5}$$

Expanding both sides to first order in τ and discarding higher order terms results in

$$\begin{aligned} -\tau \sum_x p_x (BT|x\rangle \otimes \overline{T}|x\rangle + T|x\rangle \otimes \overline{BT}|x\rangle) = \\ \sum_x q_x T|x\rangle \otimes \overline{T}|x\rangle + i \sum_x p_x (AT|x\rangle \otimes \overline{T}|x\rangle - T|x\rangle \otimes \overline{AT}|x\rangle). \end{aligned} \tag{S6}$$

Taking the inner product of Eq. (S6) with $\langle y|T^\dagger \otimes \langle y|T^\top$ results in

$$\begin{aligned} -\tau \sum_x p_x (\langle y|T^\dagger BT|x\rangle \langle y|T^\top \overline{T}|x\rangle + \langle y|T^\dagger T|x\rangle \langle y|T^\top \overline{BT}|x\rangle) = \\ \sum_x q_x \langle y|T^\dagger T|x\rangle \langle y|T^\top \overline{T}|x\rangle + i \sum_x p_x (\langle y|T^\dagger AT|x\rangle \langle y|T^\top \overline{T}|x\rangle - \langle y|T^\dagger T|x\rangle \langle y|T^\top \overline{AT}|x\rangle). \end{aligned} \tag{S7}$$

Using the identities $\langle y|T^\dagger T|x\rangle = \langle y|T^\top \overline{T}|x\rangle = \delta_{xy}$ for T unitary results in

$$-\tau p_y (\langle y|T^\dagger BT|y\rangle + \langle y|T^\top \overline{BT}|y\rangle) = q_y + i p_y (\langle y|T^\dagger AT|y\rangle - \langle y|T^\top \overline{AT}|y\rangle). \tag{S8}$$

Because A is Hermitian, we additionally have $\langle y|T^\dagger AT|y\rangle = \langle y|T^\top \overline{AT}|y\rangle$ so that the last two terms on the right-hand side cancel, resulting in

$$q_y = -2\tau p_y \mathrm{Re}\left[\langle y|T^\dagger BT|y\rangle\right] \tag{S9}$$

With the q_x's determined, we can now determine the operator A. Rearranging Eq. (S6), we first isolate the terms containing A:

$$\begin{aligned} i\sum_x p_x(AT|x\rangle \otimes \overline{T}|x\rangle - T|x\rangle \otimes \overline{AU}|x\rangle) = \\ -\tau \sum_x p_x(BT|x\rangle \otimes \overline{T}|x\rangle + T|x\rangle \otimes \overline{BT}|x\rangle) - \sum_x q_x T|x\rangle \otimes \overline{T}|x\rangle. \end{aligned} \tag{S10}$$

We define the right hand side as

$$|\Phi\rangle = -\tau \sum_x p_x(BT|x\rangle \otimes \overline{T}|x\rangle + T|x\rangle \otimes \overline{BT}|x\rangle) - \sum_x q_x T|x\rangle \otimes \overline{T}|x\rangle. \tag{S11}$$

We then decompose A into a sum over Pauli strings with domain size D, $A = \sum_j a_j \sigma_j$, where the σ_j are Pauli strings acting on at most D qubits, $a_j \in \mathbb{R}$ and $a_j = \mathcal{O}(\tau)$ for all j. Substituting into the left hand side of Eq. (S10) yields

$$i\sum_{x,j} p_x a_j(\sigma_j T|x\rangle \otimes \overline{T}|x\rangle - T|x\rangle \otimes \overline{\sigma_j}\,\overline{T}|x\rangle) = \sum_j a_j|v_j\rangle = |\Phi\rangle, \tag{S12}$$

where we have defined the vectors $|v_j\rangle := \sum_x p_x(\sigma_j T|x\rangle \otimes T^*|x\rangle - T|x\rangle \otimes \sigma_j^* T^*|x\rangle)$. Denoting by f the function

$$f(a) = \left\| |\Phi\rangle - i\sum_j a_j|v_j\rangle \right\|^2 = \langle\Phi|\Phi\rangle + i\sum_j (a_j^*\langle v_j|\Phi\rangle - a_j\langle\Phi|v_j\rangle) + \sum_{jk} a_j^* a_k \langle v_j|v_k\rangle), \tag{S13}$$

the optimal coefficients a_j are determined by minimizing f. This results in the set of equations

$$0 = \frac{\partial f}{\partial a_k} = -\mathrm{Im}\left[\langle v_k|\Phi\rangle\right] + \sum_j a_j \mathrm{Re}\left[\langle v_k|v_j\rangle\right]. \tag{S14}$$

Defining the matrix $S_{jk} := \mathrm{Re}\left[\langle v_j|v_k\rangle\right]$ and the vector $b_j := \mathrm{Im}\left[\langle v_j|\Phi\rangle\right]$, the optimal coefficients a are the solution to the linear system $Sa = b$.

Using the definition's of $|v_j\rangle$ and $|\Phi\rangle$, we calculate calculate the matrix elements of S as

$$S_{jk} = \mathrm{Re}\left[\langle v_j|v_k\rangle\right] = \sum_x p_x^2 \mathrm{Re}\left[\langle x|T^\dagger(\sigma_j\sigma_k + \sigma_k\sigma_j)T|x\rangle\right] - 2\sum_{xy} p_x p_y \mathrm{Re}\left[\langle x|T^\dagger\sigma_j T|y\rangle\langle x|T^\dagger\sigma_k T|y\rangle\right], \tag{S15}$$

and the elements of b as

$$b_j = -2\tau\left(\sum_x p_x^2 \mathrm{Im}\left[\langle x|T^\dagger\sigma_j BT|x\rangle\right] + \sum_{xy} p_x p_y \mathrm{Im}\left[\langle x|T^\dagger\sigma_j T|y\rangle\langle y|T^\dagger B^\dagger T|x\rangle\right]\right). \tag{S16}$$

B. Implementing W_k via a QITE adaptation

The real time evolution corresponding to $W_k = \exp\left(\tau \overline{L_k} \otimes L_k\right)$ can be determined completely analogously to that of $V_k \otimes \overline{V_k}$ above. The resulting equations are

$$\begin{cases} q_y = \tau \sum_x p_x \left| \langle y|T^\dagger L_k T|x\rangle \right|^2 \\ S_{jk} = \mathrm{Re}\left[\langle v_j|v_k\rangle\right] \\ b_j = \mathrm{Im}\left[\langle v_j|\Phi\rangle\right] \end{cases} \tag{S17}$$

where $Sa = b$ gives the optimal Pauli strings. The matrix elements for S are the same, as the vectors $|v_j\rangle$ are identical in both cases. Since $|\Phi\rangle$ has a different form, the elements of b are modified and given by

$$b_j = \mathrm{Im}\left[\langle v_j|\Phi\rangle\right] = 2\tau \sum_{xy} p_x p_y \mathrm{Im}\left[\langle x|T^\dagger \sigma_j L_k T|y\rangle \langle y|T^\dagger L_k^\dagger T|x\rangle\right]. \tag{S18}$$

C. Conservation of Probability

The trace of the density operator, given by $\mathrm{Tr}\left(\rho\right) = \sum_x p_x = 1$, is preserved by time evolution generated by the Lindblad equation. Here we show that time evolution via Algorithm II also maintains the trace. The trace is preserved if the sum of all q_x's is zero at each time step. This requires summing the contributions to the q_x's from both the V_k and W_k terms as follows:

$$\begin{aligned} \sum_y q_y &= \sum_y \left(-\tau p_y \mathrm{Re}\left[\langle y|T^\dagger L_k^\dagger L_k T|y\rangle\right] + \tau \sum_x p_x \left|\langle y|T^\dagger L_k T|x\rangle\right|^2 \right) \\ &= -\tau \sum_y p_y \langle y|T^\dagger L_k^\dagger L_k T|y\rangle + \tau \sum_x p_x \sum_y \langle x|T^\dagger L_k^\dagger T|y\rangle \langle y|T^\dagger L_k T|x\rangle \\ &= -\tau \mathrm{Tr}\left(\rho L_k^\dagger L_k\right) + \tau \sum_x p_x \langle x|T^\dagger L_k^\dagger L_k T|x\rangle \\ &= -\tau \mathrm{Tr}\left(\rho L_k^\dagger L_k\right) + \tau \mathrm{Tr}\left(\rho L_k^\dagger L_k\right) \\ &= 0 \end{aligned} \tag{S19}$$

D. Measuring Observables

The result of the above real and imaginary time evolution, the weights $p_x(t)$ and the Hermitian operator $A(t)$ can be used to calculate the expectation value of any observ-

able. Inverting the Choi-Jamiołkowski isomorphism gives us the density operator $\rho(t) = \sum_x p_x(t)T(t)|x\rangle\langle x|T^\dagger(t)$. Observables O are then calculated as

$$\begin{aligned}\langle O(t)\rangle = \mathrm{Tr}\left(O\rho(t)\right) &= \sum_{xy} p_x(t)\langle y|OT(t)|x\rangle\langle x|T^\dagger(t)|y\rangle \\ &= \sum_{xy} p_x\langle x|T^\dagger|y\rangle\langle y|OT|x\rangle \\ &= \sum_{x} p_x\langle x|T^\dagger\left(\sum_y |y\rangle\langle y|\right)OT|x\rangle \\ &= \sum_{x} p_x(t)\langle x|T^\dagger(t)OT(t)|x\rangle.\end{aligned} \tag{S20}$$

Beyond a certain number of qubits, storing all the $p_x(t)$'s is not possible, and a stochastic sampling approach is needed. Locality conditions suggest one possible approach to efficient sampling, described at the end of section (I E), which converges faster than uniform random sampling.

E. Measuring Matrix Elements

To obtain the coefficients q_x and a_i, we need to measure various matrix elements. In general, we can decompose any operator into a sum over Pauli strings, $X = \sum_j x_i\sigma_i$. Since $\langle x|X|y\rangle = \sum_j x_i\langle x|\sigma_i|y\rangle$, we then need to measure $\langle x|\sigma_i|y\rangle$ for all Pauli strings σ_i. This can be done using the following identities:

$$2\mathrm{Re}\left[\langle x|X|y\rangle\right] = \frac{\langle x| + \langle y|}{\sqrt{2}}X\frac{|x\rangle + |y\rangle}{\sqrt{2}} - \frac{\langle x| - \langle y|}{\sqrt{2}}X\frac{|x\rangle - |y\rangle}{\sqrt{2}}, \tag{S21}$$

$$2\mathrm{Im}\left[\langle x|X|y\rangle\right] = \frac{\langle x| + i\langle y|}{\sqrt{2}}X\frac{|x\rangle - i|y\rangle}{\sqrt{2}} - \frac{\langle x| - i\langle y|}{\sqrt{2}}X\frac{|x\rangle + i|y\rangle}{\sqrt{2}}. \tag{S22}$$

In general, the state $(|x\rangle + i^p|y\rangle)/\sqrt{2}$, with $p \in \{0,1,2,3\}$, requires a quantum circuit comprising m CNOT gates and having depth $m+1$, where m is the Hamming distance between the binary strings x, y [**?**]. Indeed, one can find an index k such that $x_k \neq y_k$. Without loss of generality, one can assume that $x_k = 1$ (otherwise, just invert the roles of x, y and replace p with $-p$ mod 4). One can then define the sets $S = \{l : x_l = 1, l \neq k\}$, $T = \{l : x_l \neq y_l, l \neq k\}$. Finally, starting from a register of n qubits prepared in $|0\rangle^{\otimes n}$, the desired state is obtained by: (i) applying a product of X gates on qubits in the set S, $\prod_{l\in S} X_l$, (ii) applying to qubit k the gate $g_p = H, SH, ZH, ZSH$, for $p = 0,1,2,3$,

respectively, and (iii) applying a product of CNOT gates to qubits in T controlled by qubit k, $\prod_{l\in S} c_k X_l$.

For local observables the state preparation is simpler, as described in the following. Consider a k-qubit observable $X^{(k)} \otimes \mathbb{I}_{n-k}$, with $X^{(k)}$ acting non-trivially on k qubits out of a total of n qubits. Then

$$\langle x|X^{(k)} \otimes \mathbb{I}_{n-k}|y\rangle = \langle x_1, \ldots, x_k, x_{k+1}, \ldots, x_n|X^{(k)} \otimes \mathbb{I}_{n-k}|y_1, \ldots, y_k, y_{k+1}, \ldots, y_n\rangle \tag{S23}$$

$$= \delta_{x_{k+1},y_{k+1}} \cdots \delta_{x_n,y_n} \langle x_1, \ldots, x_k|X^{(k)}|y_1, \ldots, y_k\rangle. \tag{S24}$$

Thus we need only to prepare the states

$$\frac{|x_1, \ldots, x_k, x_{k+1}, \ldots, x_n\rangle + |y_1, \ldots, y_k, x_{k+1}, \ldots, x_n\rangle}{\sqrt{2}} = \frac{|x_1, \ldots, x_k\rangle + |y_1, \ldots, y_k\rangle}{\sqrt{2}} \otimes |x_{k+1}, \ldots, x_n\rangle. \tag{S25}$$

Since k is typically small, only 1 or 2 qubits in most cases and independent of the system size, this state can be efficiently prepared. The form of Eq. (S25) suggests a stochastic sampling method to determine which p_x's to store classically. For simplicity we describe the case of qubits in a line, and the indices $1, \ldots, n$ labelling the sites with the observable acting on the first k qubits. The general case is similar. Since the matrix elements will depend more heavily on qubits $1, \ldots, k+m$ for some cutoff m, we can sample with higher frequency on the first $k+m$ qubits and with lower frequency on the rest. In addition, in many cases we expect the dissipation channels L_k to reduce long range correlations, further increasing the convergence rate of local sampling.

II. HARDWARE ERRORS IN THE TWO LEVEL SYSTEM SIMULATION

We compare the hardware results to noiseless numerical emulations in Fig. S1 so as to further separate hardware and algorithmic errors. The noiseless numerical emulations were run with the same circuits as in the hardware trials but using IBM's *qasm_simulator*. From Fig. S1, we see that Algorithm I has a larger deviation between the emulation and hardware data. This difference can be accounted for by the fact that the hardware experiment for Algorithm II requires only a single qubit, so the density matrix for all time steps can be obtained from only single qubit rotations. Single qubit simulations can always be compiled to a constant depth regardless of the number of time steps, resulting in lower depth circuits and correspondingly lower total gate error. In addition, since 2-qubit gates are generally lower fidelity than single-qubit gates, there is no infidelity contribution due to 2-qubit gates in Algorithm II.

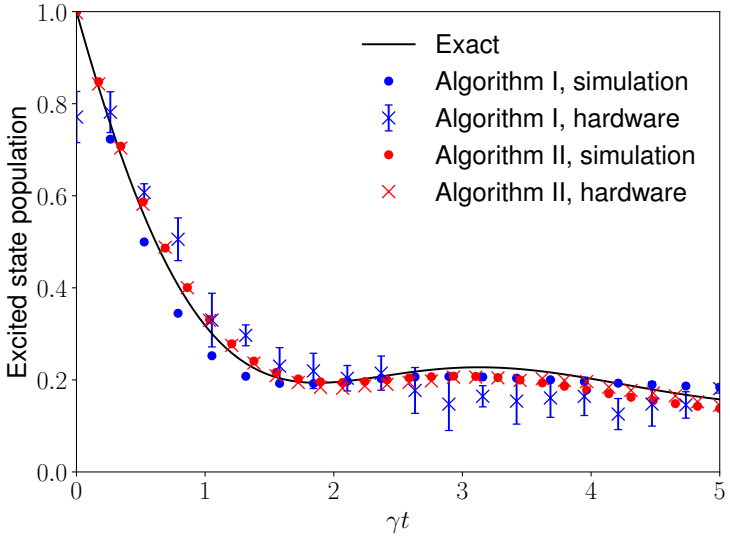

FIG. S1: Excited state population for the two level system (TLS). The solid curve is the exact solution and the blue and red dots are noiseless numerical emulations of Algorithm I and II, respectively. The blue and red crosses are the hardware results presented in the main text for Algorithm I and II, respectively. The deviation between hardware and simulation results for Algorithm I are larger than for Algorithm II, which we attribute to hardware error resulting from the larger circuit depth and number of qubits needed for the Algorithm I.

III. NUMBER OF PAULI STRINGS IN THE TRANSVERSE FIELD ISING MODEL

Exactly simulating the 2-site TFIM using Algorithm I requires measuring expectation values of the 256 Pauli strings on 4 qubits. To reduce the runtime of the Algorithm, we use a subset of all Pauli strings. We show in Fig. S2 that increasing the number of included Pauli strings beyond 16 has only a minor effect on the observables.

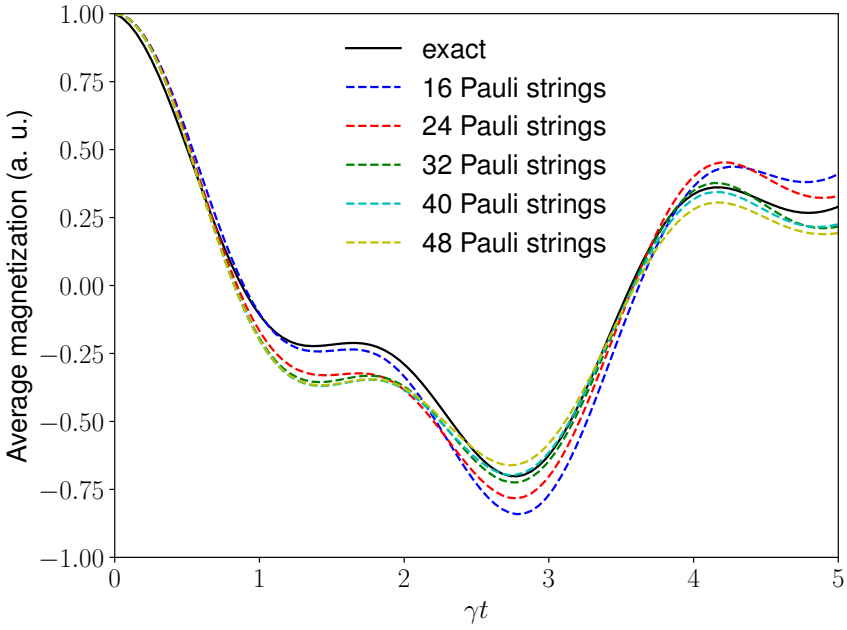

FIG. S2: Noiseless numerical simulations for the transverse field Ising model (TFIM) using Algorithm I with increasing number of Pauli strings included. Here the dissipation rate is $\gamma = 0.1$. The black solid curve is the exact result, and the blue dashed curve is a simulation of Algorithm I using the same 16 Pauli's as in the main text. The red, green, light blue, and yellow dashed curves are noiseless numerical simulations of Algorithm I obtained from including an increasing number of Pauli strings in the simulation. From these simulations we see that only marginal increase in accuracy is obtained from including a larger number of Pauli strings.

IV. EFFECT OF DISSIPATION RATE ON ALGORITHM PRECISION

Here we study the effects of increasing dissipation rates on the accuracy of both algorithms. In general, we should expect larger algorithmic errors when larger dissipation rates are simulated since both the Trotter error and QITE error increase with the operator norm of the Lindblad operators. Larger dissipation rates correspond to Lindblad operators with larger norms and hence larger algorithmic errors. To understand how increasing dissipation rates affect both algorithm's errors, we performed simulations of the 2-site TFIM with dissipation rates ranging from $\gamma = 0$ to $\gamma = 1$. Figure S3 shows the results of the simulations. We see that in this specific case, which includes 16 Pauli strings for Algorithm I and all possible bit-strings for Algorithm II, Algorithm II performs qualitatively better than Algorithm I for all dissipation rates. Although Algorithm II performs better for the simulations shown in Fig. S3, we have not considered the error due to bit-string selection. For larger systems where all bit-strings cannot be included, there will be additional errors introduced by including only a strict-subset of all possible bit-strings.

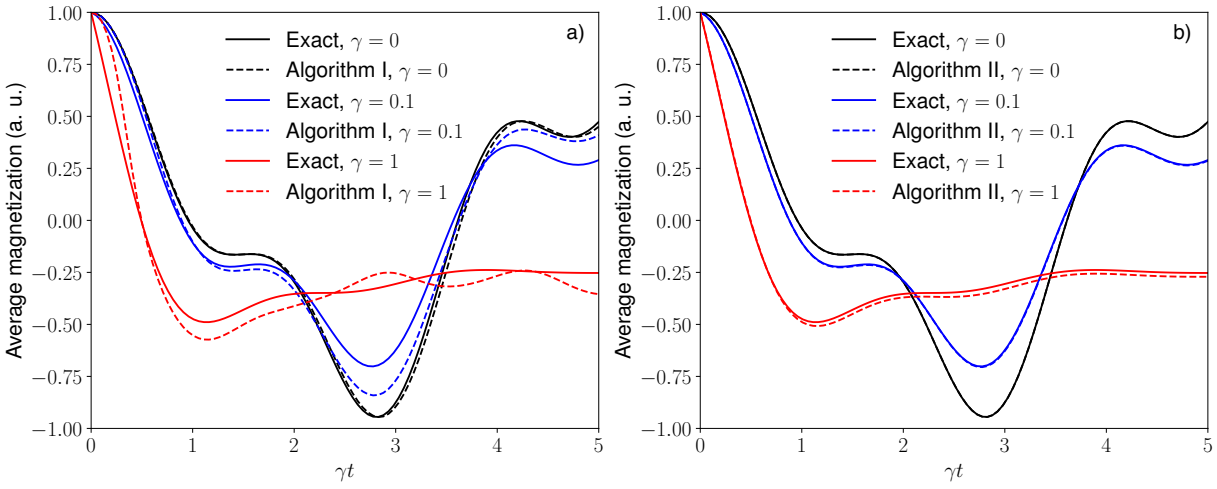

FIG. S3: The effect of increasing the dissipation rate from $\gamma = 0$ to $\gamma = 1$. a) Noiseless simulation of Algorithm I using 16 Pauli strings. The same qualitative error is obtained for all dissipation rates simulated. b) Noiseless simulation of Algorithm II.