Lecture 4: Pauli propagation and matrix product states

Computational aspects of quantum simulation
ENS de Lyon

2026

Abstract

We first propagate a selected Pauli observable backwards through a Clifford+T circuit. Clifford gates preserve one Pauli term and a T-gate creates at most two, leading to an exact expectation-value algorithm exponential only in the T-count. We then introduce Schmidt decomposition, Schmidt rank, and optimal truncation across a bipartite cut. Finally, we define matrix product states and state the characterization of their minimal bond dimensions by Schmidt ranks. This characterization will be proved in Lecture 5, where we also develop MPS computations and circuit updates.

Learning objectives. After this lecture, students should be able to:

- propagate a Pauli observable backwards through a Clifford+T circuit;
- prove the $2^t \operatorname{poly}(N, m, t)$ runtime for an exact Pauli expectation value with t T-gates;
- explain why this observable-specific algorithm is not automatically a sampler for the complete output distribution;
- compute Schmidt decompositions and ranks across consecutive cuts;
- prove the optimal one-cut truncation bound and compute its error.
- define an open-boundary MPS and state the relation between its minimal bond dimensions and Schmidt ranks.

Two routes to compression. Pauli propagation stores only the observable being queried; its cost is controlled by the number of Pauli terms produced. Schmidt decomposition instead reveals when a state is a sum of few product vectors across a cut. We introduce an MPS description of the complete state and state its Schmidt-rank characterization here. The proof and the algorithms for computing with and updating an MPS come in Lecture 5.

Contents

1 Exercise from Lecture 3

Problem

Start from $|00\rangle$, apply H_1, $\mathrm{CNOT}_{1\to 2}$, and S_2, and then measure X_1, with outcome $\mu \in \{+1, -1\}$. If $\mu = -1$, apply Z_1Z_2; if $\mu = +1$, do nothing. Finally measure Y_2, with outcome $\nu \in \{+1, -1\}$. Simulate the complete adaptive computation using stabilizer generators.

Solution

The input generators are Z_1, Z_2. Conjugating them through the circuit gives

$$\begin{aligned} \langle Z_1, Z_2\rangle &\xmapsto{H_1} \langle X_1, Z_2\rangle \\ &\xmapsto{\mathrm{CNOT}_{1\to 2}} \langle X_1X_2, Z_1Z_2\rangle \\ &\xmapsto{S_2} \langle X_1Y_2, Z_1Z_2\rangle. \end{aligned}$$

The measured Pauli X_1 commutes with X_1Y_2 and anticommutes with Z_1Z_2. Hence μ is uniform, and after the measurement a valid generating set is

$$\langle X_1Y_2, \mu X_1\rangle = \langle \mu X_1, \mu Y_2\rangle.$$

The equality follows by multiplying the two signed generators.

If $\mu = +1$, no correction is made and the generators are already X_1, Y_2. If $\mu = -1$, conjugation by Z_1Z_2 changes both $-X_1$ and $-Y_2$ into their positive versions. Thus the two classical branches merge:

$$S_{\text{after correction}} = \langle X_1, Y_2\rangle.$$

The final Y_2-measurement is therefore deterministic with $\nu = +1$. The transcript distribution is

$$\Pr[(\mu,\nu) = (+1,+1)] = \frac{1}{2}, \qquad \Pr[(\mu,\nu) = (-1,+1)] = \frac{1}{2},$$

and all other transcripts have probability zero. The final state is

$$|+\rangle \otimes |+i\rangle, \qquad |+i\rangle = \frac{|0\rangle + i|1\rangle}{\sqrt{2}}.$$

This calculation uses exactly the gate, measurement, and adaptive-control updates in the Gottesman–Knill algorithm.

2 Pauli propagation beyond Clifford circuits

The stabilizer method evolves a concise description of the state. For a non-Clifford circuit that description need not remain closed. If only one expectation value is requested, a different strategy is possible: leave the input state fixed and propagate the observable backwards through the circuit.

2.1 Backward propagation of an observable

Let

$$U = U_m \cdots U_1,$$

where U_1 acts first, and let the input be $|\phi\rangle$. For an output observable O,

$$\langle O\rangle_{\text{out}} = \langle\phi|U^\dagger O U|\phi\rangle.$$

Set $O_m = O$, and for $j = m, m-1, \dots, 1$, define

$$O_{j-1} = U_j^\dagger O_j U_j.$$

Then $O_0 = U^\dagger O U$, so the desired answer is $\langle\phi|O_0|\phi\rangle$.

The Pauli operators in $\{I, X, Y, Z\}^{\otimes N}$ form an orthonormal basis of the complex vector space of operators on N qubits, for the normalized Hilbert–Schmidt inner product

$$\langle A, B\rangle_{\text{op}} = 2^{-N}\,\mathrm{Tr}(A^\dagger B).$$

Indeed, one-qubit trace orthogonality and the tensor-product trace formula give $2^{-N}\,\mathrm{Tr}(PQ) = \delta_{P,Q}$, and there are 4^N such Paulis, equal to the dimension of the operator space. Consequently every operator has a unique Pauli expansion

$$O = \sum_{P\in\{I,X,Y,Z\}^{\otimes N}} c_P P, \qquad c_P = 2^{-N}\,\mathrm{Tr}(PO).$$

Storing all 4^N coefficients is not useful. The algorithm below stores only the nonzero Pauli terms generated from the selected output observable.

2.2 The only branching rule: conjugation by a T gate

Let

$$T = \begin{pmatrix} 1 & 0 \\ 0 & \omega \end{pmatrix}, \qquad \omega = e^{i\pi/4} = \frac{1+i}{\sqrt{2}}.$$

Clifford+T is universal in the approximate sense: every N-qubit unitary can be approximated arbitrarily closely, up to global phase, by a circuit over $\{H, S, \mathrm{CNOT}, T\}$ [1]. In fact, H, T, CNOT already suffice, since $S = T^2$.

Since T is diagonal,

$$T^\dagger I T = I, \qquad T^\dagger Z T = Z.$$

Direct multiplication gives

$$T^\dagger X T = \begin{pmatrix} 0 & \omega \\ \overline{\omega} & 0 \end{pmatrix} = \frac{X - Y}{\sqrt{2}},$$

$$T^\dagger Y T = \begin{pmatrix} 0 & \overline{\omega} \\ \omega & 0 \end{pmatrix} = \frac{X + Y}{\sqrt{2}}.$$

Thus a local I or Z creates one Pauli term, while a local X or Y creates two. By contrast, if C is Clifford, then $C^\dagger P C$ is one signed Pauli whenever P is a Pauli.

2.3 Evaluating the final Paulis

At the end of backward propagation we must evaluate Pauli operators on the input. Recall from Lecture 3 that, given a Hermitian Pauli and a stabilizer state specified by its generators, we know how to compute the expectation value in polynomial time. We restate the rule here.

Proposition 2.1 (Pauli expectation on a stabilizer state)**.** *Let $S = \langle g_1, \dots, g_N \rangle$ stabilize $|\phi\rangle$, and let P be a Hermitian Pauli. Then*

$$\langle \phi | P | \phi \rangle \in \{0, +1, -1\}.$$

More precisely:

(i) if P anticommutes with some g_i, then the expectation is 0;

(ii) if P commutes with every g_i, then exactly one of P and $-P$ belongs to S, giving expectation $+1$ or -1, respectively.

The correct case and sign can be found in polynomial time from the tableau.

Proof. If $Pg_i = -g_iP$, then

$$\langle \phi | P | \phi \rangle = \langle \phi | g_i P g_i | \phi \rangle = -\langle \phi | P | \phi \rangle,$$

so the expectation vanishes. If P commutes with every generator, then $P|\phi\rangle$ lies in the one-dimensional common $+1$ eigenspace of the stabilizer group. Hence $P|\phi\rangle = \mu|\phi\rangle$ for some $\mu \in \{+1, -1\}$, and $\mu P \in S$. Solving the binary membership system and multiplying the selected signed generators determines μ. □

2.4 The exact propagation algorithm

Represent the current observable as a list of pairs (c, P), where P is a canonical tensor product of I, X, Y, Z and any sign has been absorbed into c. Write $\mathcal{L}_j$ for the list representing O_j:

$$O_j = \sum_{(c,P) \in \mathcal{L}_j} cP.$$

Initially $\mathcal{L}_m$ contains the terms of the requested output observable $O_m = O$. Process the gates for $j = m, m-1, \dots, 1$. At each step construct a new list for $O_{j-1} = U_j^\dagger O_j U_j$ by applying the appropriate rule to every entry of the old list $\mathcal{L}_j$. Write

$$P = P_1 \otimes \cdots \otimes P_N, \qquad P_i \in \{I, X, Y, Z\},$$

and let $P[i \leftarrow R]$ denote the string obtained by replacing its i-th factor with R.

1. If $U_j = C$ is a Clifford unitary, then for every $(c, P) \in \mathcal{L}_j$, replace it with $(c\eta, P')$, where

   $$C^\dagger P C = \eta P', \qquad \eta \in \{+1, -1\}.$$

   Here P' is a canonical tensor product of I, X, Y, Z.

2. If $U_j = T_i$ for some $i \in [N] = \{1, \dots, N\}$, then for every $(c, P) \in \mathcal{L}_j$, inspect the local factor P_i:

   - if $P_i = I$ or $P_i = Z$, keep (c, P) unchanged;

- if $P_i = X$, replace it with the two entries

$$\left(\frac{c}{\sqrt{2}}, P\right), \qquad \left(-\frac{c}{\sqrt{2}}, P[i \leftarrow Y]\right);$$

- if $P_i = Y$, replace it with the two entries

$$\left(\frac{c}{\sqrt{2}}, P[i \leftarrow X]\right), \qquad \left(\frac{c}{\sqrt{2}}, P\right).$$

3. After this update, terms with the same Pauli string may be merged by adding their coefficients; zero coefficients may be discarded. The resulting list is $\mathcal{L}_{j-1}$.
4. After all gates have been processed, evaluate every Pauli in $\mathcal{L}_0$ on $|\phi\rangle$ using Proposition 2.1, multiply by its coefficient, and sum.

The number of T-gates is the *T-count*. The branching rule gives a clean parameterized result.

Theorem 2.2 (Exact Pauli expectations for Clifford+T)**.** *Let U be an N-qubit circuit over $\{H, S, \mathrm{CNOT}, T\}$, of size m and T-count t. Given a stabilizer input $|\phi\rangle$ and a Hermitian Pauli P, the expectation*

$$\langle\phi|U^\dagger P U|\phi\rangle$$

can be computed exactly in

$$2^t \operatorname{poly}(N, m, t)$$

time. In particular, the computation is polynomial-time when $m = \operatorname{poly}(N)$ and $t = O(\log N)$.

Proof. For the requested Pauli observable $O_m = P$, the initial list $\mathcal{L}_m = \{(1, P)\}$ has size one. A Clifford gate replaces every entry with exactly one entry, so the list size stays the same before optional merging. A T-gate replaces every entry with at most two entries, so the list size increases by a factor of at most two. Merging or cancellation can only reduce the size. Since there are t T-gates in total, every intermediate list has size at most 2^t.

There are m gate updates, each acting on at most 2^t entries. A Pauli label uses $O(N)$ bits, and each update takes polynomial time. The coefficients lie in $\mathbb{Q}(\sqrt{2})$ and have bit length polynomial in t, so exact arithmetic is polynomial per entry as well. Finally, Proposition 2.1 evaluates each of the at most 2^t final entries in polynomial time. This gives the stated $2^t \operatorname{poly}(N, m, t)$ time bound. □

Remark 2.3. *The 2^t bound is a worst-case bound. A T-gate does not branch when the current local Pauli is I or Z, and entries in the list can later merge or cancel.*

2.5 Worked two-qubit example

Let the input be $|00\rangle$, let

$$U = \mathrm{CNOT}_{1\to 2}\, H_1 T_1 H_1,$$

where the rightmost gate acts first, and request the output observable Z_2. Backward propagation gives

$$Z_2 \xmapsto{\mathrm{CNOT}_{1\to2}} Z_1 Z_2 \xmapsto{H_1} X_1 Z_2 \xmapsto{T_1} \frac{X_1 Z_2 - Y_1 Z_2}{\sqrt{2}} \xmapsto{H_1} \frac{Z_1 Z_2 + Y_1 Z_2}{\sqrt{2}}.$$

Since

$$\langle 00|Z_1 Z_2|00\rangle = 1, \qquad \langle 00|Y_1 Z_2|00\rangle = 0,$$

we obtain

$$\langle 00|U^\dagger Z_2 U|00\rangle = \frac{1}{\sqrt{2}}.$$

Consequently,

$$\Pr[Y_2 = 0] = \frac{1 + \langle Z_2\rangle}{2} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right).$$

2.6 Small output marginals and the limitation of the method

For measured qubits $R = \{j_1, \dots, j_k\}$ and $y \in \{0,1\}^k$, the projector onto outcome y is

$$\Pi_{R,y} = \prod_{\ell=1}^{k} \frac{I + (-1)^{y_\ell} Z_{j_\ell}}{2} = 2^{-k} \sum_{A \subseteq \{1,\dots,k\}} (-1)^{\sum_{\ell\in A} y_\ell} \prod_{\ell\in A} Z_{j_\ell}.$$

Corollary 2.4 (Exact simulation of a small output marginal)**.** *A specified probability* $\Pr[Y_R = y]$ *for the circuit in Theorem 2.2 can be computed exactly in*

$$2^{t+k}\,\mathrm{poly}(N, m, t)$$

time.

Proof. The projector contains 2^k Pauli terms. Apply Theorem 2.2 to each term and add the results. □

The theorem evaluates selected observables; it does not automatically give a weak simulator for all N output bits. A full output projector contains 2^N Pauli terms, and sequential sampling introduces projectors for the previously sampled bits. A different representation may be preferable when the objective is to store or sample the complete state.

More sophisticated stabilizer-decomposition methods can improve the dependence on t for other simulation tasks [2]. Approximate Pauli propagation can also truncate the expanding operator representation [3]. Here we keep the elementary exact theorem because its proof makes the controlling parameter transparent.

3 A second compression principle: limited entanglement

Pauli propagation compresses one chosen observable. We now seek a compact description of the complete state. Fix an ordering of n sites, each of local dimension d. Matrix product states exploit limited correlations across the consecutive cuts

$$1 \dots k \mid k+1 \dots n.$$

The ordering matters: the same state can have small ranks for one ordering and large ranks for another. A broad introduction to the representation and its algorithms can be found in [4].

4 Schmidt decomposition

Definition 4.1 (Schmidt decomposition)**.** *Let $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. A Schmidt decomposition is*

$$|\psi\rangle = \sum_{\alpha=1}^{r} s_\alpha |u_\alpha\rangle_A |v_\alpha\rangle_B,$$

where $s_\alpha > 0$, the two families of Schmidt vectors are orthonormal, and $\sum_\alpha s_\alpha^2 = 1$. The number r is the Schmidt rank.

To obtain it, write

$$|\psi\rangle = \sum_{a,b} M_{ab} |a\rangle_A |b\rangle_B$$

and apply a singular-value decomposition $M = U\Sigma V^\dagger$. There are exactly $\mathrm{rank}(M)$ nonzero singular values, and each gives one term in the Schmidt decomposition. Thus the number of terms is $r = \mathrm{rank}(M)$. These nonzero singular values are the s_α, and the Schmidt vectors are

$$|u_\alpha\rangle = \sum_a U_{a\alpha} |a\rangle, \qquad |v_\alpha\rangle = \sum_b \overline{V_{b\alpha}} |b\rangle.$$

In particular, the Schmidt rank equals the matrix rank of M.

A product state has Schmidt rank 1. A Bell pair has Schmidt coefficients $(1/\sqrt{2}, 1/\sqrt{2})$ and rank 2. The n-qubit GHZ state

$$|\mathrm{GHZ}_n\rangle = \frac{|0^n\rangle + |1^n\rangle}{\sqrt{2}}$$

also has Schmidt rank 2 across every nontrivial consecutive cut. Thus a globally entangled state can still have uniformly small Schmidt ranks.

Proposition 4.2 (Optimal truncation across one cut)**.** *Suppose $s_1 \geq s_2 \geq \cdots \geq s_r > 0$, and let $1 \leq \chi \leq r$. Among unnormalized vectors of Schmidt rank at most χ, a best approximation to $|\psi\rangle$ in Euclidean norm is*

$$|\widetilde{\psi}_\chi\rangle = \sum_{\alpha=1}^{\chi} s_\alpha |u_\alpha\rangle |v_\alpha\rangle,$$

and its squared error is

$$\big\| |\psi\rangle - |\widetilde{\psi}_\chi\rangle \big\|^2 = \sum_{\alpha > \chi} s_\alpha^2.$$

Proof. Let $|\varphi\rangle$ be any competing vector of Schmidt rank at most χ. Its left Schmidt vectors span a subspace $V \subseteq \mathcal{H}_A$ of dimension $q \leq \chi$, so $|\varphi\rangle \in V \otimes \mathcal{H}_B$. Let Π_V be the orthogonal projector onto V. Projecting the error onto the orthogonal complement of V cannot increase its norm, and annihilates $|\varphi\rangle$. Hence

$$\begin{aligned}
\big\| |\psi\rangle - |\varphi\rangle \big\|^2 &\geq \big\| ((I - \Pi_V) \otimes I) |\psi\rangle \big\|^2 \\
&= \sum_{\alpha=1}^{r} s_\alpha^2 \big\| (I - \Pi_V) |u_\alpha\rangle \big\|^2 \\
&= 1 - \sum_{\alpha=1}^{r} s_\alpha^2 w_\alpha, \qquad w_\alpha = \big\| \Pi_V |u_\alpha\rangle \big\|^2.
\end{aligned}$$

The second equality uses orthogonality of the right Schmidt vectors.

We have $0 \le w_\alpha \le 1$ and $\sum_\alpha w_\alpha \le q \le \chi$. To justify the latter inequality, choose an orthonormal basis $|e_1\rangle, \dots, |e_q\rangle$ of V. Then

$$\sum_{\alpha=1}^{r} w_\alpha = \sum_{j=1}^{q} \sum_{\alpha=1}^{r} |\langle e_j \mid u_\alpha \rangle|^2 \le q,$$

because the $|u_\alpha\rangle$'s are orthonormal.

These constraints imply that the captured squared norm is largest when all the available weight is placed on the χ largest coefficients. Explicitly, since $s_\alpha^2 \le s_\chi^2$ for $\alpha > \chi$,

$$\begin{aligned}\sum_{\alpha=1}^{r} s_\alpha^2 w_\alpha &\le \sum_{\alpha=1}^{\chi} s_\alpha^2 w_\alpha + s_\chi^2 \sum_{\alpha>\chi} w_\alpha \\ &\le \sum_{\alpha=1}^{\chi} \big(s_\alpha^2 w_\alpha + s_\chi^2 (1 - w_\alpha)\big) \\ &\le \sum_{\alpha=1}^{\chi} s_\alpha^2.\end{aligned}$$

Thus every competitor has squared error at least $\sum_{\alpha>\chi} s_\alpha^2$. The truncated vector in the statement attains this bound: its error is precisely the sum of the discarded, orthogonal Schmidt terms. □

The discarded weight therefore gives a concrete local measure of compressibility. If Schmidt coefficients are tied at the cutoff, the optimal approximation need not be unique.

5 Matrix product states

Definition 5.1 (Open-boundary matrix product state). *An open-boundary n-qudit matrix product state (MPS) $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ is a state whose amplitudes have the form*

$$\psi_{i_1 \dots i_n} = A^{[1]i_1} A^{[2]i_2} \cdots A^{[n]i_n},$$

or equivalently

$$|\psi\rangle = \sum_{i_1, \dots, i_n = 1}^{d} A^{[1]i_1} A^{[2]i_2} \cdots A^{[n]i_n} |i_1 \cdots i_n\rangle.$$

Here $A^{[k]i_k}$ is a $\chi_{k-1} \times \chi_k$ matrix, with $\chi_0 = \chi_n = 1$, so each complete matrix product is a scalar. The integers χ_k are the bond dimensions.

Remark (MPS as a tensor network). *A matrix product state is a special case of a* tensor network*: nodes represent tensors, and links joining nodes represent* contractions*, that is, summation over a shared index. An open link represents an index that is not summed over. For an MPS, the nodes form a chain. The tensor at site k has entries $A^{[k]i_k}_{\alpha_{k-1},\alpha_k}$; its bond indices connect neighboring sites, while its physical index i_k remains open.*

Horizontal links: contracted bond indices $\alpha_k \in \{1, \dots, \chi_k\}$.
Upward open links: physical indices $i_k \in \{1, \dots, d\}$.

The matrix product in the definition is precisely the contraction of all bond indices:

$$\psi_{i_1\dots i_n} = \sum_{\alpha_1=1}^{\chi_1} \cdots \sum_{\alpha_{n-1}=1}^{\chi_{n-1}} A^{[1]i_1}_{1,\alpha_1} A^{[2]i_2}_{\alpha_1,\alpha_2} \cdots A^{[n]i_n}_{\alpha_{n-1},1}.$$

The open physical indices label the amplitudes of the resulting state.

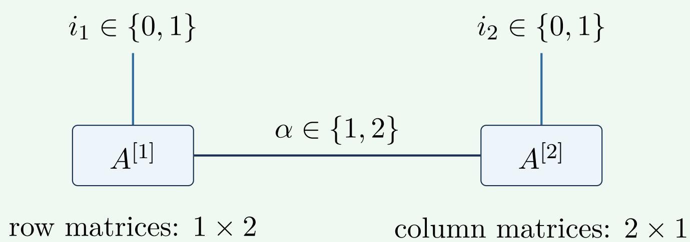
A two-qubit example: the Bell state. For qubits, label the physical basis by $0, 1$ rather than $1, 2$. We represent

$$|\Phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

with bond dimensions $\chi_0 = 1$, $\chi_1 = 2$, and $\chi_2 = 1$. Choose the two row matrices at site 1 and the two column matrices at site 2 as follows:

$$A^{[1]0} = \frac{1}{\sqrt{2}}(1 \;\; 0), \qquad A^{[1]1} = \frac{1}{\sqrt{2}}(0 \;\; 1), \qquad A^{[2]0} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad A^{[2]1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The tensor $A^{[1]}$ therefore consists of two 1×2 matrices, one for each value of i_1; similarly, $A^{[2]}$ consists of two 2×1 matrices, one for each value of i_2.



For fixed i_1, i_2, multiply the selected row and column. This sums over the shared bond index α, but not over the physical indices:

$$\psi_{i_1 i_2} = A^{[1]i_1} A^{[2]i_2} = \sum_{\alpha=1}^{2} A^{[1]i_1}_{1,\alpha} A^{[2]i_2}_{\alpha,1}.$$

Indeed, the four products are

$$\psi_{00} = A^{[1]0} A^{[2]0} = \frac{1}{\sqrt{2}}, \qquad \psi_{01} = A^{[1]0} A^{[2]1} = 0,$$

$$\psi_{10} = A^{[1]1} A^{[2]0} = 0, \qquad \psi_{11} = A^{[1]1} A^{[2]1} = \frac{1}{\sqrt{2}}.$$

Thus these tensors give exactly the Bell state, directly from the MPS definition, without using the bond-dimension theorem below.

If every $\chi_k \le \chi$, the tensors contain at most

$$\sum_{k=1}^{n} d\, \chi_{k-1} \chi_k \le n d \chi^2$$

complex numbers, rather than the d^n amplitudes of a general state. How small can we make the bond dimensions? The answer is determined by the Schmidt ranks.

Theorem 5.2 (Bond dimension equals Schmidt rank)**.** *For a state $|\psi\rangle$ on an ordered chain, let r_k be its Schmidt rank across the cut $1 \dots k \mid k+1 \dots n$. Every MPS representation of $|\psi\rangle$ with bond dimensions χ_k satisfies $r_k \leq \chi_k$. Moreover, there is an exact MPS representation with $\chi_k = r_k$ simultaneously at every cut.*

Theorem 5.2 will be proved in Lecture 5. There we will construct the tensors using successive SVDs and then explain how to compute with and update an MPS.

6 Optional training exercises

1. Propagate X backwards through two consecutive T-gates. Show explicitly that the four formal contributions merge and cancel to give $-Y$, consistently with $T^2 = S$ and $S^\dagger X S = -Y$.
2. Compute the Schmidt decompositions of a product state, a Bell state, and $|\mathrm{GHZ}_n\rangle$ across every consecutive cut.

7 Assigned exercise: Schmidt decomposition of the W state

For $n \geq 2$, define

$$|W_n\rangle = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |0^{j-1} 1\, 0^{n-j}\rangle,$$

and set $|W_1\rangle = |1\rangle$. This exercise uses only Schmidt decomposition and optimal truncation; no result about MPS is needed.

(a) For $1 \leq k < n$, prove the decomposition

$$|W_n\rangle = \sqrt{\frac{k}{n}}\, |W_k\rangle |0^{n-k}\rangle + \sqrt{\frac{n-k}{n}}\, |0^k\rangle |W_{n-k}\rangle,$$

and show that it is a Schmidt decomposition.

(b) Determine the Schmidt rank and the two Schmidt coefficients across each nontrivial consecutive cut.

(c) For a fixed cut k, give a best *unnormalized* approximation to $|W_n\rangle$ of Schmidt rank at most one, and compute the minimum squared Euclidean error. Justify optimality using Proposition 4.2. What happens when n is even and $k = n/2$?

References

[1] P. Oscar Boykin, Tal Mor, Matthew Pulver, Vwani Roychowdhury, and Farrokh Vatan, *On Universal and Fault-Tolerant Quantum Computing*, 1999.
https://arxiv.org/abs/quant-ph/9906054

[2] Sergey Bravyi and David Gosset, *Improved Classical Simulation of Quantum Circuits Dominated by Clifford Gates*, Physical Review Letters 116, 250501 (2016).
https://arxiv.org/abs/1601.07601

[3] Manuel S. Rudolph, Tyson Jones, Yanting Teng, Armando Angrisani, and Zoë Holmes, *Pauli Propagation: A Computational Framework for Simulating Quantum Systems*, 2025.
https://arxiv.org/abs/2505.21606

[4] Ulrich Schollwöck, *The density-matrix renormalization group in the age of matrix product states*, Annals of Physics 326, 96–192 (2011).
https://arxiv.org/abs/1008.3477