Tensor Networks

3 hours · 100 points

Answer all questions. This is a closed-book test; a calculator is allowed. Parts I and II need no justification. In Part III, show your reasoning: most of the credit is for the idea, not the arithmetic. Logarithms are natural. Unless stated otherwise, every index of a tensor network has dimension χ.

(12 points) Part I: True/false questions

Mark each statement True (T) or False (F). Each question is worth 2 points.

1. ( ) Slicing k bonds of dimension 2 splits a contraction into 2^k independent smaller contractions whose results are summed. It lowers the peak memory and never increases the total number of operations.
2. ( ) If every tensor in a network has nonnegative entries, there is no sign problem, so the network can always be contracted exactly in time polynomial in the number of tensors.
3. ( ) The left-canonical form of an MPS is unique: two left-canonical MPS of the same state with the same minimal bond dimensions have identical tensors.
4. ( ) Consider a local Hamiltonian on an open chain of N sites, with bounded local terms, a unique ground state and a gap bounded below by a constant independent of N. Its ground state can be approximated to fidelity $1-\varepsilon$ by an MPS whose bond dimension grows only polynomially in N and $1/\varepsilon$.
5. ( ) Reverse-mode automatic differentiation gives the gradient of a real scalar loss with respect to all input tensors of a contraction in a number of operations bounded by a small constant times that of the forward contraction, however many inputs there are.
6. ( ) Exchanging two legs of a bosonic tensor is a transpose of the array. In a fermionic ($\mathbb{Z}_2$-graded) tensor network, exchanging two adjacent legs also multiplies by -1 every block in which both legs carry odd parity.

(18 points) Part II: Fill in the blanks

Each question is worth 3 points.

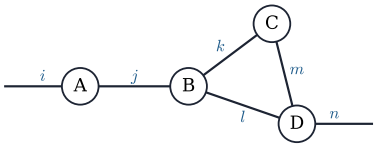
7. In the bulk of a long open chain, the minimal MPO bond dimension of the transverse-field Ising Hamiltonian $H = -\sum_i Z_i Z_{i+1} - g\sum_i X_i$ is _____. For the Heisenberg chain $H = \sum_i \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1}$ it is _____.
8. Across a cut into k and $N-k$ qubits ($1 \le k < N$), the Schmidt rank of the GHZ state $(|0...0\rangle + |1...1\rangle)/\sqrt{2}$ is _____. For the W state $(|10...0\rangle + |01...0\rangle + ... + |0...01\rangle)/\sqrt{N}$ it is _____.
9. A rank-χ real $n \times n$ matrix is written as $M = AB$ with A of size $n \times \chi$ and B of size $\chi \times n$. This uses $2n\chi$ numbers, but the set of such matrices has dimension _____. An open MPS with N sites, physical dimension d and all $N-1$ bonds of dimension $\chi \le d$ has, in the same sense, _____ redundant parameters.
10. An MPS of bond dimension χ has half-chain entanglement entropy at most _____. For the critical transverse-field Ising chain (central charge $c = 1/2$) with open ends, this bound alone forces χ to grow at least as N^p with $p =$ _____.
11. The norm $\langle\psi|\psi\rangle$ of an open MPS (N sites, physical dimension d, bond dimension χ) is contracted rung by rung from the left in the best pairwise order. The total cost scales as $O(N \cdot$ _____). Contracting all ket tensors first instead creates an intermediate with _____ entries.
12. Fibonacci anyons have one nontrivial particle τ with $\tau \times \tau = 1 + \tau$. Six τ anyons can fuse to the vacuum in _____ linearly independent ways, and to a single τ in _____ ways.

(70 points) Part III: Problems. Show all your work and justify your answers.

13. (14 points) Einsum and contraction order. A tensor network is a set of tensors with labelled indices. Indices shared by two tensors are summed; open indices form the output. The figure shows the network `ij,jkl,km,lmn->in`.

(a) (4 pts) Write the output T_{in} as an explicit sum. The cost of one pairwise contraction is the product of the sizes of all distinct indices of both tensors. Find the total cost and the number of entries of the largest intermediate for the orders $((AB)C)D$ and $((CD)B)A$. Which pairs should never be contracted first, and why?

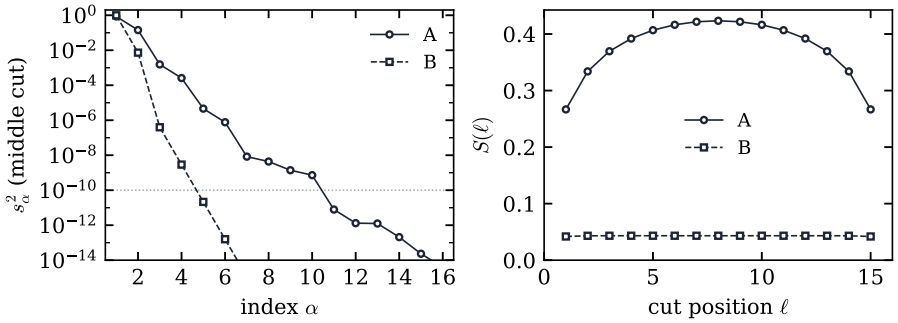
(b) (4 pts) Any contraction order is a binary tree whose leaves are the tensors. Explain why exhaustive search over trees is hopeless for a network of 500 tensors. Name one heuristic used in practice and describe it in one or two sentences. The optimal largest intermediate of a network is governed by a graph quantity: which one?



(c) (3 pts) For the $L \times L$ grid with bond dimension χ, argue that every contraction order has an intermediate with at least χ^L entries, up to constant factors in the exponent. Hint: follow the moment when half the tensors have been absorbed.

(d) (3 pts) The intermediate of (c) does not fit in memory. Explain how slicing s bonds lowers the peak memory, what it does to the total operation count, and why the sliced contractions run well on many GPUs.

14. (14 points) Matrix product states. An MPS of N sites, physical dimension d and bond dimension χ writes $\psi_{s_1 \ldots s_N} = A_1^{s_1} A_2^{s_2} \ldots A_N^{s_N}$ where A_k^s is a $\chi_{k-1} \times \chi_k$ matrix, $\chi_0 = \chi_N = 1$ and $\chi_k \leq \chi$. The figure shows the ground state of the transverse-field Ising chain $H = -\sum_i Z_i Z_{i+1} - g \sum_i X_i$ with $N = 16$ open sites, computed exactly at two values of g: the Schmidt weights s_α^2 at the middle cut (left) and the entanglement entropy $S(\ell) = -\sum_\alpha s_\alpha^2 \ln s_\alpha^2$ at every cut (right).



(a) (3 pts) State the left-canonical condition on A_k as an equation and draw it as a diagram. In mixed-canonical form, sites $1 \ldots k$ are left-canonical, sites $k+1 \ldots N$ are right-canonical, and a centre matrix C sits on bond k. Why are the singular values of C the Schmidt coefficients of that cut?

(b) (4 pts) Show that truncating the centre to its χ largest singular values gives a state ψ_χ with $\|\psi - \psi_\chi\|^2 = \varepsilon = \sum_{\alpha > \chi} s_\alpha^2$ (fidelity $1 - \varepsilon$ after normalizing), and that no state of bond dimension χ on that cut does better. One curve is $g = 1$ (the critical point) and the other is $g = 3$. Which is which? Using the left panel, estimate the χ needed for $\varepsilon < 10^{-10}$ in each case.

(c) (4 pts) Time-evolving block decimation (TEBD) splits H into even and odd bonds. Write the first-order Trotter step and state the order of its error per unit time. How is a two-site gate applied to an MPS, and what does each application cost in χ and d? Why does imaginary-time evolution $e^{-\tau H}|\psi\rangle$, renormalized as it goes, project onto the ground state, and what must hold for $|\psi\rangle$?

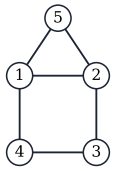
(d) (3 pts) DMRG minimizes $\langle\psi|H|\psi\rangle / \langle\psi|\psi\rangle$ over the MPS one or two sites at a time. Each local step is an eigenvalue problem: for which matrix, and why is a Lanczos solver used rather than full diagonalization? Ground states in two dimensions obey an area law $S \sim W$ on a cylinder of width W. How must χ scale with W?

15. (14 points) Differentiable and generic tensor networks. Contraction uses only multiplication and addition. Programs built from it can be differentiated, and the two operations can be replaced by those of another semiring.

(a) (4 pts) Let $C = AB$, i.e. `ij,jk->ik`, for real matrices A and B, and let $\mathcal{L}$ be a real scalar computed from C. Derive $\overline{A} = \partial\mathcal{L}/\partial A$ and $\overline{B}$ from $\overline{C}$, and write both as einsums. State the general rule for the adjoint of one input of an einsum. What must reverse mode keep in memory, and how does checkpointing reduce it?

(b) (4 pts) The backward rule of the symmetric eigendecomposition $A = U\Lambda U^T$ contains $F_{ij} = 1/(\lambda_j - \lambda_i)$ for $i \neq j$. When does it blow up? Fixed-point methods such as the corner-transfer-matrix renormalization group (CTMRG) run such decompositions inside an iteration $x_{t+1} = f(x_t, \theta)$ until a fixed point $x^* = f(x^*, \theta)$. Derive $dx^*/d\theta$ by implicit differentiation. Compare its memory cost with differentiating through all the unrolled iterations.

(c) (6 pts) An independent set of a graph is a set of vertices with no edge inside it. Give each vertex a variable $x_v \in \{0, 1\}$, a vertex tensor $W_v = (1, w_v)$ and each edge a matrix $B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, indexed by the two endpoint variables. The network is $\sum_{\{x_v\}} \prod_v W_v(x_v) \prod_{(u,v)\in E} B(x_u, x_v)$; a vertex variable is shared by all tensors that touch it. (i) What does the ordinary contraction compute when all $w_v = z$, a formal variable? Do it for the "house" graph in the figure. (ii) Now use tropical numbers, with $W_v = (0, w_v)$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & -\infty \end{bmatrix}$. What does the contraction return? Call it $E_{\max}(w)$ and give its value for the house graph with all $w_v = 1$. (iii) Show that $\partial E_{\max}/\partial w_v$ is the value of x_v in the optimal set, when that set is unique. (iv) Rydberg atoms in tweezers cannot both be excited within the blockade radius. Explain how this physics, at positive detuning, relates to (ii).

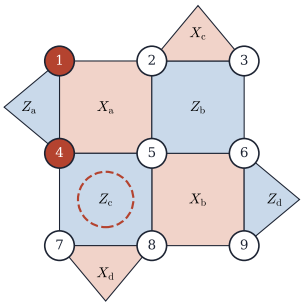


16. (14 points) Quantum circuits, channels and error correction. Quantum states, gates, channels and the decoding of codes are all tensor networks.

(a) (3 pts) Write the amplitude $\langle x|U|0^n\rangle$ of an n-qubit circuit as a closed tensor network: what are its tensors? Why can a single amplitude be far cheaper than the full state vector? Draw the network of an expectation value $\langle\psi|O|\psi\rangle$.

(b) (4 pts) Amplitude damping has Kraus operators $K_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}$ and $K_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}$. Check that the channel preserves trace. Compute its Pauli transfer matrix $R_{ij} = \mathrm{tr}\big(P_i \mathcal{E}(P_j)\big)/2$, with $P \in (I, X, Y, Z)$. Which entry shows that the channel is not unital, and what does it do physically?

(c) (7 pts) The figure shows the rotated $d = 3$ surface code: 9 data qubits, four Z-type checks $Z_\mathrm{a}...Z_\mathrm{d}$ (blue) and four X-type checks $X_\mathrm{a}...X_\mathrm{d}$ (red). (i) Count the independent stabilizers and the logical qubits. Why does each check commute with every check of the other type? (ii) X errors hit qubits 1 and 4 (filled). Explain why only the circled check Z_c flags, and why Z_a does not. Find a minimum-weight correction and the net operator after correcting. Did decoding succeed? (iii) Maximum-likelihood decoding instead compares the total probability of each logical class, $P(L \mid s) \propto \sum_{S \in \mathcal{S}} P(E_0 LS)$, a sum over the stabilizer group. Why is it better than choosing the single most likely error? Why is this sum a tensor-network contraction? For depolarizing noise, why does one need approximate (boundary-MPS) contraction at large d?



17. (14 points) Symmetries and tensor categories. A tensor that commutes with a symmetry is block sparse. For non-abelian symmetries and anyons, each block is labelled by a fusion tree.

(a) (4 pts) In the XXZ chain the total S^z is conserved. Let h be the two-site XXZ bond operator. Show which entries of the gate $e^{-ih\delta t}$ can be nonzero in the basis $(\uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow)$, and give the block sizes. For an MPS tensor $A^s_{\alpha\beta}$ with charge q_α flowing in on the left bond, q_s in on the physical leg and q_β out on the right bond, state the selection rule. A bond of $\chi = 300$ splits into three sectors of size 100. By what factor does block structure speed up its SVD?

(b) (5 pts) For SU(2), the Wigner–Eckart theorem factorizes an invariant tensor into a reduced tensor times Clebsch–Gordan coefficients. (i) How many independent SU(2)-invariant tensors have four spin-1/2 legs? (ii) Show that $\boldsymbol{S}_1 \cdot \boldsymbol{S}_2$ ($\hbar = 1$) on two spin-1/2 sites is fixed by two numbers, and give them. (iii) A bond space has multiplicities $m_0 = 50$, $m_1 = 50$ and $m_2 = 20$ for spins $j = 0, 1, 2$. Give its dimension. For an SU(2)-invariant square matrix acting on this space, give the number of entries actually stored and estimate the speedup of its SVD.

(c) (5 pts) The figure shows an F-move, the change of basis between two fusion trees with the same outer legs. (i) Why is a leg permutation of a non-abelian symmetric tensor not just a transpose of its blocks? (ii) For Fibonacci anyons, $F_\tau^{\tau\tau\tau} = \begin{bmatrix} \varphi^{-1} & \varphi^{-1/2} \\ \varphi^{-1/2} & -\varphi^{-1} \end{bmatrix}$ with $\varphi = \left(1 + \sqrt{5}\right)/2$. Show that it is unitary and equal to its own inverse. (iii) Why is this framework, with F-symbols for recoupling and R-symbols for braiding, useful for anyonic chains and topological phases? Which of the two matters for fermionic PEPS in two dimensions?

