

Turning the Knapsack Problem into a Continuous Optimization using a Quantum-Inspired Variational Method

A. Troyano Olivas^{1,2}, Hans H. Brunner¹, Fred Fung¹, Momtchil Peev¹, Vicente Martin²

¹ Munich Research Center, Huawei Technologies Duesseldorf GmbH, Munich, Germany

² Center for Computational Simulation, Universidad Politécnica de Madrid, Madrid, Spain

alvaro.troyano@huawei.com

Abstract

We encode the binary knapsack problem as the minimization of a diagonal quantum Hamiltonian, where the linear constraint is implemented through a Fourier-approximated projector onto feasible bitstrings. By working on the manifold of matrix product states (MPS)—implemented natively in PyTorch with automatic differentiation—we turn the combinatorial search into a tractable continuous gradient-based optimization. A Rényi-entropy regularizer ensures that the only critical points of the objective are the computational basis states, that the ground state is a strict local minimum, and that all other basis states are saddle points. On flat knapsack instances the solver achieves solutions within 2% of optimality for problems up to $n = 15\,000$, with a time that scales linearly in n and remains computationally feasible on a single CPU. Theoretical analysis and principal component analysis of the trajectory confirm that the optimization landscape is free of spurious local minima.

1 Introduction

The *knapsack problem* (in its minimization form) is an NP-hard combinatorial optimization problem:

$$\min_{x \in \{0,1\}^n} \sum_{j=1}^n h_j x_j \quad \text{s.t.} \quad \sum_{j=1}^n a_j x_j \leq b, \quad (1)$$

with $h_j < 0$, $a_j > 0$, and $b > 0$. Classical exact methods (dynamic programming, branch-and-bound) become impractical beyond a few hundred variables.

In this work we propose a quantum-inspired continuous relaxation of the problem. We construct a diagonal Hamiltonian H whose expectation on computational basis states reproduces the objective, and we define a projector P (which is exactly the one that appears in [1]) onto the feasible subspace using Fourier synthesis. Working with matrix product states (MPS) [2] allows us to perform gradient-based optimization of the Rayleigh quotient $\langle \psi | PHP | \psi \rangle / \langle \psi | \psi \rangle$ on a low-dimensional manifold. A carefully chosen Rényi-entropy regularizer keeps the benign critical-point structure: only the optimal basis state is a local minimum, all other basis states are saddle points, and the regularizer prevents premature collapse onto suboptimal solutions.

Results show the effectiveness and scalability of this approach, having linear scaling and providing good solutions (roughly $\sim 2\%$ mean relative error). We find that entanglement plays a crucial role in the optimization, suggesting that it is the underlying cause for the benign critical point structure.

2 Method

2.1 Encoding the problem

Let $D = \text{diag}(0, 1)$ act on a single qubit, and let D_j act as D on the j -th qubit and identity elsewhere.

$$H = \sum_{j=1}^n h_j D_j \quad (2)$$

is a diagonal Hamiltonian in the computational basis and satisfies $\langle x | H | x \rangle = \sum_j h_j x_j$ for $x \in \{0, 1\}^n$. The linear constraint is encoded by the operator $A = \sum_j a_j D_j$; its eigenvalues are exactly the values $a^\top x$.

2.2 Projector onto feasible space

The ideal projector P acts as $P|x\rangle = |x\rangle$ if $a^\top x \leq b$, and 0 otherwise. This is the step function $f(\lambda) = \mathbf{1}_{\lambda \leq b}$ applied to A . We approximate f on the interval $[0, \sum a_j]$ by a truncated Fourier series with $2L + 1$ terms:

$$f(\lambda) \approx \sum_{l=-L}^L c_l e^{il\omega\lambda}, \quad c_l = \frac{b}{\sum a_j} \text{sinc}\left(\pi l \frac{b}{\sum a_j}\right) e^{-il\omega b/2}, \quad \omega = \frac{2\pi}{\sum a_j}, \quad (3)$$

to obtain an approximation for the projector:

$$\tilde{P} = \sum_{l=-L}^L c_l \bigotimes_{j=1}^n e^{il\omega a_j D}. \quad (4)$$

Because $[H, \tilde{P}] = 0$, the projected Hamiltonian $\hat{H} = \tilde{P}H\tilde{P} = H\tilde{P}^2$ can be expressed as a Matrix Product Operator (MPO) [3] with bond dimension $\mathcal{O}(L)$ after exploiting the group property $e^{il\omega a D} e^{im\omega a D} = e^{i(l+m)\omega a D}$. Since $\hat{H}$ is a diagonal matrix (computational basis states are eigenvectors),

$$\min_{x \in \{0,1\}^n} \langle x | \hat{H} | x \rangle = \min_{|\phi\rangle} \langle \phi | \hat{H} | \phi \rangle = \min_{|\psi\rangle} \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \quad (5)$$

Problem (5) is an asymptotically correct approximation of equation (1). Note, $\langle \phi | \phi \rangle = 1$ while $\langle \psi | \psi \rangle \in \mathbb{R}$.

2.3 Regularized objective

To avoid premature collapse onto suboptimal basis states we add the negative second Rényi entropy as a regularizer to the Rayleigh quotient of equation (5):

$$\mathcal{L}(\psi) = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} + \frac{1}{\beta} \log \frac{\sum_x |\psi(x)|^4}{(\langle \psi | \psi \rangle)^2}. \quad (6)$$

This is the Helmholtz free energy [4] with inverse temperature β and Rényi entropy $S_2(\psi) = -\log \sum_x |\psi(x)|^4$. For $\beta \rightarrow \infty$ the landscape is convex; as β increases the optimal state traces a continuous path that avoids spurious critical points. In practice we fix $\beta \in [300, 1000]$ for the results shown.

3 Theoretical analysis

The landscape of $\mathcal{L}$ on the full Hilbert space is remarkably simple. Let $E_x = \langle x | \hat{H} | x \rangle$ and denote by $|x^*\rangle$ the ground state of $\hat{H}$. Define the spectral gap $\Delta = \min_{x \neq x^*} (E_x - E_{x^*})$.

Theorem 1 (Critical-point characterization). *All computational basis states are critical points of $\mathcal{L}$. Any critical point with multiple basis state support requires $|E_x - E_y| \leq 2/\beta$ for all x, y in its support. Hence, for $\Delta > 2/\beta$, the only critical points are basis states. A gap larger than $2/\beta$ forbids mixed critical points.*

Theorem 2 (Local convexity at the optimum). *The Hessian of $\mathcal{L}$ at $|x^*\rangle$ has eigenvalues $2(E_x - E_{x^*} - 2/\beta)$. If $\Delta > 2/\beta$, all are positive and $|x^*\rangle$ is a strict local minimum. Every other basis state possesses at least one negative eigenvalue and is a saddle point, guaranteeing a descent direction toward the optimum.*

Theorem 3 (MPS manifold). *For any bond dimension χ , computational basis states are critical points of $\mathcal{L}$ restricted to the MPS manifold. Spurious critical points, where the projected gradient vanishes but the full-space gradient does not, disappear as χ increases. The full-space landscape is recovered for $\chi \geq 2^{\lfloor n/2 \rfloor}$.*

Thus, gradient-based optimization is well behaved: the objective provides negative curvature toward the global optimum from any suboptimal basis state, while the entropy term prevents premature collapse.

4 Experimental results

All experiments use randomly generated “flat” knapsack instances (weights $h_j \sim U[1, 100]$, values a_j proportional to weights with $\pm 5\%$ noise, capacity ratio $\alpha = 0.5$). The MPS solver (implemented in PyTorch

[5], for automatic gradient tracking) uses bond dimension $\chi = 5$, $L = 10$ Fourier modes, and 60 Adam steps (learning rate 6.5×10^{-3}).

4.1 Solution quality and scaling

Figure 1a shows the empirical complexity of the tensor network solver against a Dynamic Programming solver, confirming that the runtime scales linearly with n . For $n \gtrsim 10000$ the MPS solver is faster than dynamic programming ($O(n^2)$). Figure 1b shows the relative error as a function of problem size. The mean error stays around 2% for all n up to 15 000 without degradation. All tests were run with the same hyperparameters, suggesting it could be brought down by tailoring them to each instance.

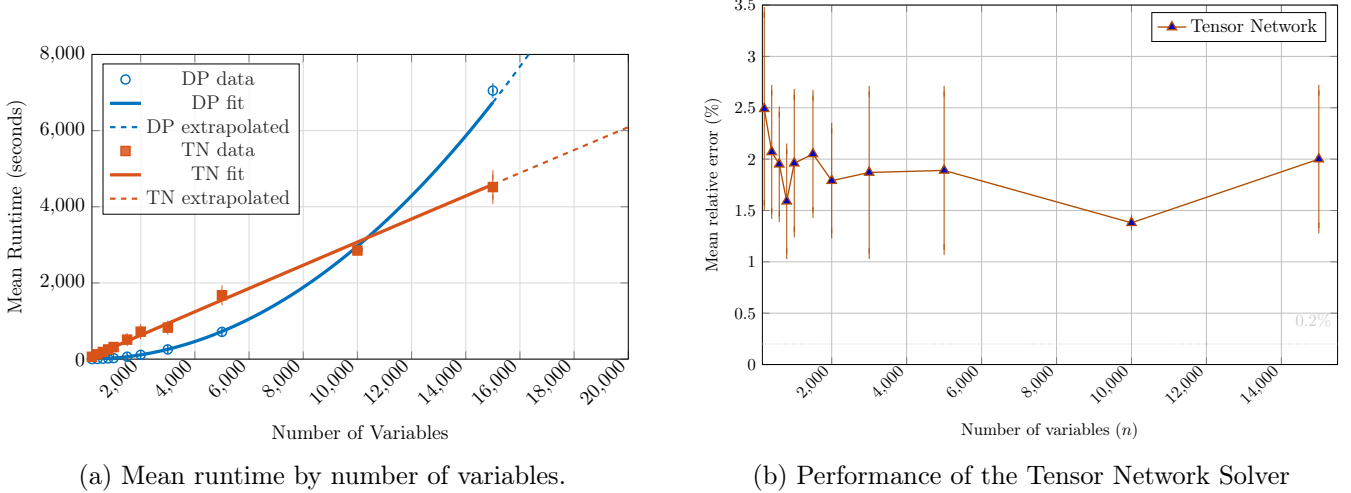


Figure 1: Left: runtime scaling. Right: relative error scaling.

4.2 Landscape diagnostics

Let $\Theta^{(t)}$ be the flattened MPS tensor parameters at step t in the optimization. We collect $\{\Theta^{(0)}, \dots, \Theta^{(T)}\}$, center at the optimum Θ^* , and compute the principal components of the centered trajectory. Figure 2 shows two probes of the optimization landscape along the first two principal component directions of the states in the optimization trajectory. The first one (figure 2a) is the optimization landscape when initializing the solver with the maximally entangled state on bond dimension $\chi = 5$. The second (figure 2b) is the initialization on the uniform superposition state with bond dimension $\chi = 5$.

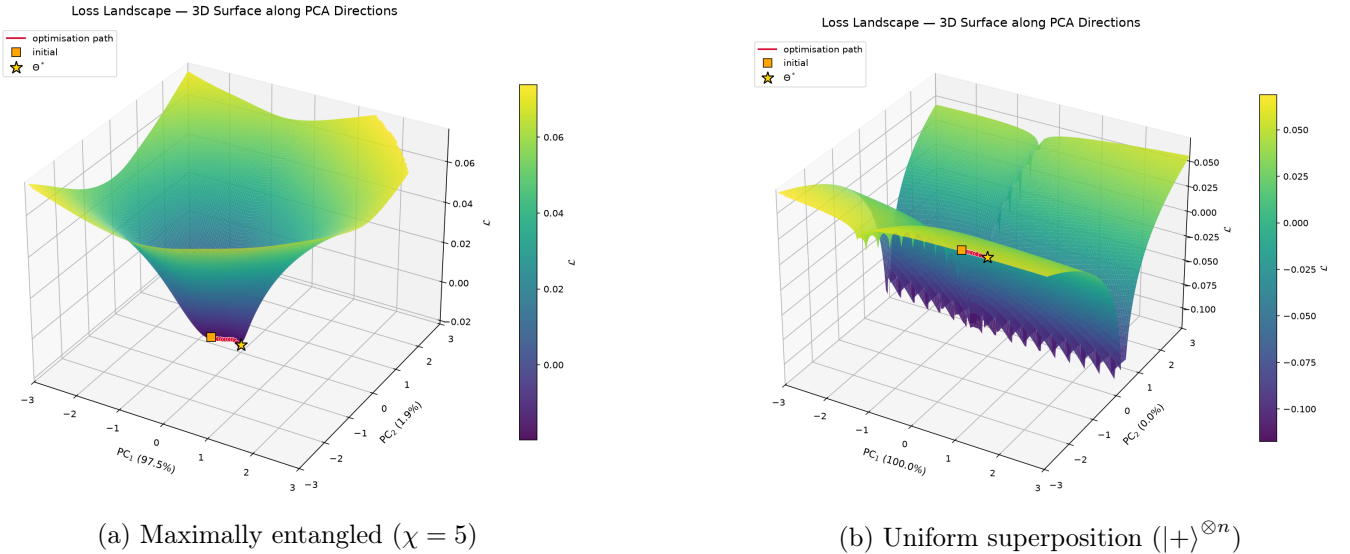


Figure 2: Optimization landscape over PCA directions for two different initial states.

Although these observations agree with Theorems 1–3, they suggest that the real driving force for the results discussed in section 3 is entanglement, rather than bond dimension.

References

- [1] T. Hao, X. Huang, C. Jia, and C. Peng, “A quantum-inspired tensor network algorithm for constrained combinatorial optimization problems,” *Frontiers in Physics*, vol. 10, 2022. [Online]. Available: <http://dx.doi.org/10.3389/fphy.2022.906590>
- [2] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, “Matrix product state representations,” 2007. [Online]. Available: <https://arxiv.org/abs/quant-ph/0608197>
- [3] I. P. McCulloch, “From density-matrix renormalization group to matrix product states,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2007, no. 10, p. P10014–P10014, Oct. 2007. [Online]. Available: <http://dx.doi.org/10.1088/1742-5468/2007/10/P10014>
- [4] R. K. Pathria and P. D. Beale, *Statistical Mechanics*, 3rd ed. Elsevier, 2011.
- [5] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, “PyTorch: An imperative style, high-performance deep learning library,” in *Advances in Neural Information Processing Systems 32*, 2019, pp. 8024–8035.