Convex Semidefinite Tensor Optimization and Quantum Entanglement

Liding Xu Ye-Chao Liu Sebastian Pokutta

Zuse Institute Berlin

EUROOPT 2026

Talk Roadmap

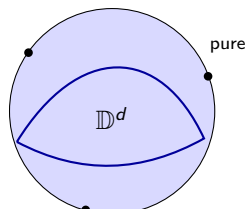
1. Physics background: quantum states, entanglement (quantum supremacy?), robustness of entanglement.
2. Optimization model: positive-semidefinite (PSD) tensors and semidefinite tensor optimization (SDTO)
3. Algorithms, separated:
   - Lifted alternating direction method of multipliers (LADMM)
   - Frank-Wolfe / geometric reconstruction
   - Cutting planes with branch-and-bound linear minimization oracle (LMO)
4. Numerical messages

Quantum States as Density Matrices

A quantum state is represented by a density matrix

$$\rho \in \mathbb{C}^{d \times d}, \qquad \rho \succeq 0, \qquad \mathrm{Tr}(\rho) = 1.$$

- Pure state: $\rho = \psi\psi^\dagger$, so $\mathrm{rank}(\rho) = 1$.
- Mixed state: probabilistic mixture $\rho = \sum_i \lambda_i \psi_i \psi_i^\dagger$.
- For one system, positivity and unit trace are enough.



density matrices
convex state space

Composite Systems: Where Entanglement Appears

For m subsystems with local dimensions $\mathbf{d} = (d_1, \ldots, d_m)$, the Hilbert space is

$$\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_m, \qquad \bar{d} = \prod_{k=1}^{m} d_k.$$

Pure product states

$$\rho = (x_1 x_1^\dagger) \otimes \cdots \otimes (x_m x_m^\dagger).$$

These are states without quantum correlations across the partition.

Entangled pure states

A pure state that cannot be written as a product state. For two qubits, the Bell state

$$\psi = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

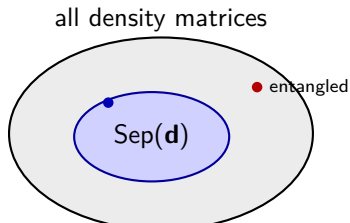
is the canonical example.

Mixed-State Separability

A mixed state is separable if it is a convex mixture of pure product states:

$$\rho \in \mathrm{Sep}(\mathbf{d}) \iff \rho = \sum_i \lambda_i \bigotimes_{k=1}^{m} x_{ik} x_{ik}^{\dagger}, \qquad \lambda_i \geq 0, \quad \sum_i \lambda_i = 1.$$

- Separable: correlations are classically explainable.
- Entangled: no such convex decomposition exists.
- Detecting entanglement in mixed states is central in quantum information and computational physics (Eisert et al., 2010; Nishioka, 2018).



Why Separability Is Hard

Easy for one system

$$m = 1 : \qquad \mathrm{Sep}(d) = \{\rho \succeq 0,\ \mathrm{Tr}(\rho) = 1\}.$$

Projection, eigenvalue tests, and semidefinite programming (SDP) solvers are available.

Hard for multipartite systems

$$m \geq 2 : \qquad \mathrm{Sep}(\mathbf{d}) \subsetneq \mathbb{D}^{\tilde{d}}.$$

Membership testing is strongly nondeterministic-polynomial-time hard (NP-hard), and linear minimization over separable states is NP-hard (Gharibian, 2010; Fang and Fawzi, 2020).

Standard physics criteria give powerful but incomplete tests:

- Positive partial transpose (PPT) is complete only in 2×2 and 2×3 bipartite systems (Peres, 1996; Horodecki, 1997).
- Doherty-Parrilo-Spedalieri (DPS) gives a hierarchy of convex outer approximations (Doherty et al., 2004; Navascués et al., 2009).
- No finite SDP representation exists in general.

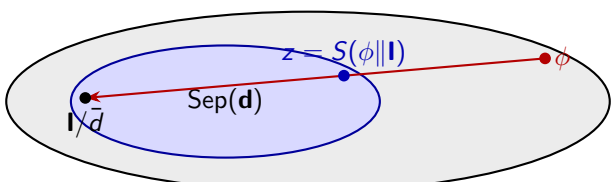
White-Noise Mixing Threshold

We test entanglement robustness by mixing a target state ϕ with white noise (Vidal and Tarrach, 1999):

$$\rho(z;\phi,\mathbf{I}) = (1-z)\phi + z\frac{\mathbf{I}}{\bar{d}} = \left(\frac{\mathbf{I}}{\bar{d}} - \phi\right) z + \phi.$$

The threshold is

$$S(\phi\|\mathbf{I}) = \min\{z \geq 0 : \rho(z;\phi,\mathbf{I}) \in \mathrm{Sep}(\mathbf{d})\}.$$



State of Art

Physics literature has many specialized methods

- (Primal heuristics) Alternating SDP optimization over states, standard in tensor optimization.
- (Primal and dual) Geometric reconstruction and primal-dual geometric reconstruction (PDGR) (Shang and Gühne, 2018; Liu et al., 2025).

Optimization point of view:

- Similar to copositive programming.
- The dimensionality of the separable set grows exponentially with the number of subsystems.

PSD Tensors as Separable States

The paper defines a cone of positive-semidefinite tensors:

$$\mathrm{PSDT}_{\mathbb{F}}(\mathbf{d}) = \left\{ \sum_{i=1}^{r} x_{i1} x_{i1}^{\dagger} \otimes \cdots \otimes x_{im} x_{im}^{\dagger} \right\}.$$

After reshaping, every element is a PSD matrix on the joint Hilbert space. The normalized slice is exactly the separable state set:

$$\mathrm{Sep}(\mathbf{d}) = \{\rho \in \mathrm{PSDT}_{\mathbb{C}}(\mathbf{d}) : \mathrm{Tr}(\rho) = 1\}.$$

Main modeling idea

Replace “is this quantum state separable?” by optimization over a convex cone that has a tensor factorization structure.

Semidefinite Tensor Optimization

The general problem studied in the paper is

$$\text{opt} = \min \left\{ \langle \mathbf{c}, \mathbf{z} \rangle : \mathbf{z} \in \mathcal{Z}, \quad A(\mathbf{z}) + \mathbf{a} \in \text{PSDT}_{\mathbb{F}}(\mathbf{d}) \right\}.$$

Connection to SDP

If $m = 1$, the PSD tensor cone is just a PSD matrix cone, so the model reduces to SDP.

What changes for $m \geq 2$

- no efficient projection,
- no easy membership test,
- cone is not self-dual,
- bounds are more realistic than exact solves.

Dual View: Entanglement Witnesses

Under standard regularity, the SDTO dual is

$$\max\left\{\langle -\mathbf{a}, y^*\rangle : y^* \in \mathrm{PSDT}_{\mathbb{F}}(\mathbf{d})^*, \quad -A^*(y^*) + \mathbf{c} \in \mathcal{Z}^*\right\}.$$

The dual cone can be tested against extreme product rays:

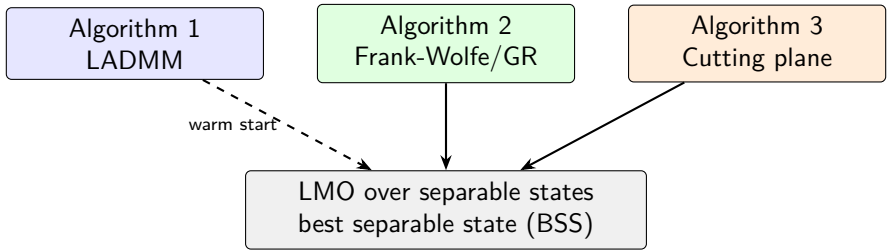
$$\mathrm{PSDT}_{\mathbb{F}}(\mathbf{d})^* = \left\{y^* : \langle y^*, p\rangle \geq 0 \quad \forall p \in \mathrm{ext}(\mathrm{PSDT}_{\mathbb{F}}(\mathbf{d}))\right\}.$$

In physics, y^* is an **entanglement witness**: nonnegative on all separable states, negative on at least one entangled state.

Computational consequence

The dual has infinitely many linear inequalities indexed by pure product states.

Three Algorithmic Routes



- LADMM is the paper's lifted nonconvex heuristic module.
- Frank-Wolfe appears through geometric reconstruction and primal-dual geometric reconstruction (PDGR) baselines.
- Cutting planes are Algorithm 3: the paper's certification/refinement module.
- LMO: (m=1) eignevalue decomposation, (m>=2) homogenous hermitian polynomial optimization.

Algorithm 1: LADMM Starts from a Smooth Lift

Use the factorization map

$$\Psi(\mathbf{x}) = \sum_{i=1}^{r} x_{i1}x_{i1}^{\dagger} \otimes \cdots \otimes x_{im}x_{im}^{\dagger}.$$

The constrained model

$$\min\{f(y) : y \in \mathrm{PSDT}(\mathbf{d}),\ z \in \mathcal{Z},\ A(z) + a = y\}$$

(after change of variables) becomes a lifted nonconvex problem in $\mathbf{x}$:

$$y = \Psi(\mathbf{x}), \qquad A(z) + a = \Psi(\mathbf{x}).$$

Use augmented Lagrangian methods to convert constraints in objective and solve the resulting unconstrained problem with tensor factorization lifting.

Analogy

For $m = 1$, this is the Burer-Monteiro idea for SDP. For $m \geq 2$, the same reparameterization is useful but can introduce spurious local minima.

What LADMM Contributes

Strengths

- Exploits product-state factorization directly.
- Produces high-quality heuristic upper bounds.
- Provides active product components for later refinement.

Limitations

- Nonconvex subproblem is inexact.
- Coupling constraint may remain violated.
- High-accuracy alternating direction method of multipliers (ADMM) convergence can be slow.

In the experiments, LADMM often reaches good objective values, but certification usually requires a cutting-plane crossover or the full iterative-refinement (IR) loop.

Algorithm 2: Frank-Wolfe for Separability Geometry

Frank-Wolfe / conditional gradient solves smooth convex objectives over $\mathrm{Sep}(\mathbf{d})$ using an LMO (Frank and Wolfe, 1956; Pokutta, 2024):

$$s_t \in \arg\min_{s \in \mathrm{Sep}(\mathbf{d})} \langle \nabla f(\rho_t), s \rangle, \qquad \rho_{t+1} = (1-\gamma_t)\rho_t + \gamma_t s_t.$$

In geometric reconstruction, the objective is typically a distance to the separable set:

$$\min_{\rho \in \mathrm{Sep}(\mathbf{d})} \frac{1}{2} \|\rho - \tau\|^2,$$

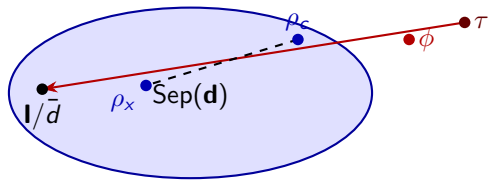
where τ is an extrapolated state on the white-noise line.

LMO meaning in physics

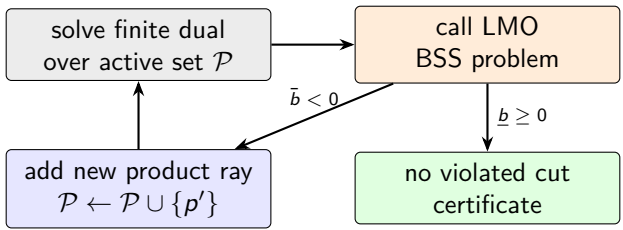
The LMO returns a pure product state that best aligns with the current witness-like search direction.

Geometric Reconstruction Picture

- Start from an extrapolated state outside the separable set.
- Frank-Wolfe returns a contact point $\rho_c \in \mathrm{Sep}(\mathbf{d})$, usually a pure product state from the last LMO call.
- Reconstruction computes ρ_x: an auxiliary point certified inside an inner approximation of $\mathrm{Sep}(\mathbf{d})$.
- If ϕ lies on the segment between ρ_c and ρ_x, then ϕ is certified separable.
- PDGR adds lower-bound estimates through a primal-dual gap mechanism (Shang and Gühne, 2018; Liu et al., 2025).



Algorithm 3: Cutting-Plane Refinement



Maintained lower bound:

$$\mathrm{lb}_{\mathrm{relx}} = \mathrm{ub}_{\mathrm{relx}} + \underline{b},$$

where $\underline{b}$ is the LMO lower bound for the separation subproblem.

The LMO Is the Bottleneck: BSS

The linear minimization oracle is the best separable state problem:

$$\min_{\rho \in \mathsf{Sep}(\mathbf{d})} \langle \chi, \rho \rangle \, .$$

The paper uses a spatial branch-and-bound (sBB) LMO with convex relaxations:

Relaxation ingredients

- DPS and dichotomy DPS outer approximations.
- McCormick inequalities for product terms.
- Tensor Reformulation-Linearization Technique (RLT) inequalities.

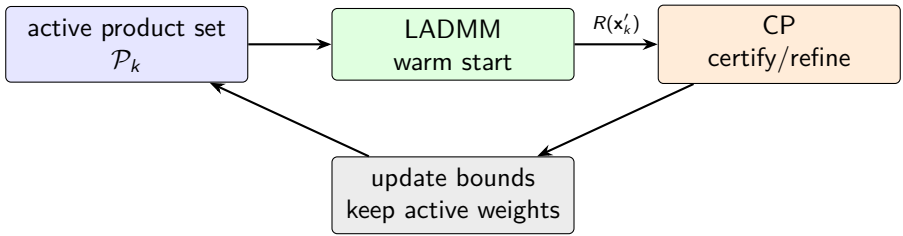
What the oracle returns

- A violated product-state cut, if one is found.
- A lower bound for certification.
- A product-state heuristic when the exact search is too costly.

Practical message

The LMO determines whether cutting planes can certify the current witness; it is also the main scalability bottleneck.

Iterative Refinement: Combining Algorithms 1 and 3



- LADMM supplies useful product components quickly.
- CP converts components into a finite dual relaxation and adds violated cuts.
- The CP subroutine guarantees a non-increasing reliable upper bound.

Experimental Setup

- Instances: GHZ, Dicke, and Cluster qubit states for $m = 3, 4, 5$.
- Objective: compute white-noise threshold upper/lower bounds.
- Implementation: Julia/JuMP with Mosek, Manopt, Ket.jl.
- Time limits: 1 hour for $m = 3$, 2 hours for $m = 4$, 3 hours for $m = 5$.
- Factorization size: $r = 256/512/652$ for $m = 3/4/5$.

Methods compared

SDP-based alternating optimization (Alt-SDP), PDGR/Frank-Wolfe, LADMM, CP, IR, DPS, and DDPS+.

Source code and detailed data are available at

`github.com/ZIB-IOL/TensorOpt4Entanglement`.

Representative Results: Upper Bounds

Instance	Best comparator	LADMM heuristic	CP	IR
GHZ_3	PDGR 0.80002	0.80000	0.80007	**0.80000**
Dicke_3_1	PDGR 0.82208	0.82203	0.82268	**0.82203**
GHZ_4	PDGR 0.88901	0.88889	0.89138	**0.88890**
Cluster_4	Alt-SDP 0.89635	0.88889	0.90164	**0.88890**
Cluster_5	Alt-SDP 0.94712	0.96488	0.95594	**0.94334**
GHZ_5	Alt-SDP 0.94327	0.96762	0.95829	**0.94163**

Message

IR gives the best reliable upper bounds across the reported benchmark suite, especially as m increases.

Scalability Pattern

$m = 3$

Near-exact upper bounds are achievable. LADMM often finds nearly feasible good objectives; IR can close gaps for GHZ_3.

$m = 4$

Runtime grows. CP becomes expensive, but IR improves upper bounds and repairs LADMM feasibility issues.

$m = 5$

Low-rank approximations are needed. IR remains best for upper bounds, but lower-bound quality depends strongly on LMO progress.

Bottleneck

The LMO is still the dominant computational challenge; as m grows, exact product-state separation becomes the limiting step.

Algorithmic Takeaways

Method	Best use	Main limitation
LADMM	fast high-quality heuristic product decompositions	infeasibility and nonconvex local behavior
Frank-Wolfe/PDGR	strong problem-specific bounds for smaller threshold instances	LMO/partition scaling
Cutting plane	reliable refinement and certification route	each cut requires a hard BSS oracle
IR	combines LADMM speed with CP reliability	lower bounds degrade if CP/LMO is under-run

What to Remember

1. Separable quantum states are normalized positive-semidefinite tensors:

$$\mathrm{Sep}(\mathbf{d}) = \{\rho \in \mathrm{PSDT}_{\mathbb{C}}(\mathbf{d}) : \mathrm{Tr}(\rho) = 1\}.$$

2. Eigenvalue decomposition is the LMO for $m = 1$, but for $m \geq 2$ the LMO is NP-hard.
3. We combined several algorithms to make some progress.
4. Open scalability problem: how can we deal with physically relevant instances with $m \geq 6$ subsystems? We can SDP relaxation for around 6-7 subsystems, but SOTA physical applications has 100 subsystems.

Selected References I

Hans J. Briegel and Robert Raussendorf. Persistent entanglement in arrays of interacting particles. *Phys. Rev. Lett.*, 86:910–913, Jan 2001. doi: 10.1103/PhysRevLett.86.910.

R. H. Dicke. Coherence in spontaneous radiation processes. *Phys. Rev.*, 93:99–110, Jan 1954. doi: 10.1103/PhysRev.93.99.

Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. Complete family of separability criteria. *Physical Review A*, 69(2):022308, February 2004. ISSN 1094-1622. doi: 10.1103/physreva.69.022308.

W. Dür and J. I. Cirac. Classification of multiqubit mixed states: Separability and distillability properties. *Phys. Rev. A*, 61:042314, Mar 2000. doi: 10.1103/PhysRevA.61.042314.

J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Rev. Mod. Phys.*, 82:277–306, Feb 2010. doi: 10.1103/RevModPhys.82.277.

Kun Fang and Hamza Fawzi. The sum-of-squares hierarchy on the sphere and applications in quantum information theory. *Mathematical Programming*, 190(1–2):331–360, July 2020. ISSN 1436-4646. doi: 10.1007/s10107-020-01537-7.

Marguerite Frank and Philip Wolfe. An algorithm for quadratic programming. *Nav. Res. Logist. Q.*, 3(1-2):95–110, 1956. doi: https://doi.org/10.1002/nav.3800030109.

S. Gharibian. Strong np-hardness of the quantum separability problem. *Quantum Information and Computation*, 10(3):343–360, March 2010. ISSN 1533-7146. doi: 10.26421/qic10.3-4-11.

Daniel M. Greenberger, Michael A. Horne, Abner Shimony, and Anton Zeilinger. Bell's theorem without inequalities. *American Journal of Physics*, 58(12):1131–1143, December 1990. ISSN 1943-2909. doi: 10.1119/1.16243.

Pawel Horodecki. Separability criterion and inseparable mixed states with positive partial transposition. *Phys. Lett. A*, 232(5):333–339, 1997. ISSN 0375-9601. doi: https://doi.org/10.1016/S0375-9601(97)00416-7.

Ye-Chao Liu, Jannis Halbey, Sebastian Pokutta, and Sébastien Designolle. A unified toolbox for multipartite entanglement certification. *arXiv preprint arXiv:2507.17435*, 2025. doi: 10.48550/arXiv.2507.17435.

Selected References II

Miguel Navascués, Masaki Owari, and Martin B. Plenio. Complete criterion for separability detection. *Phys. Rev. Lett.*, 103:160404, Oct 2009. doi: 10.1103/PhysRevLett.103.160404.

Tatsuma Nishioka. Entanglement entropy: Holography and renormalization group. *Rev. Mod. Phys.*, 90:035007, Sep 2018. doi: 10.1103/RevModPhys.90.035007.

Asher Peres. Separability criterion for density matrices. *Physical Review Letters*, 77(8):1413–1415, August 1996. ISSN 1079-7114. doi: 10.1103/physrevlett.77.1413.

Sebastian Pokutta. The frank-wolfe algorithm: a short introduction. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 126(1):3–35, 2024. doi: 10.1365/s13291-023-00275-x.

Jiangwei Shang and Otfried Gühne. Convex optimization over classes of multiparticle entanglement. *Phys. Rev. Lett.*, 120:050506, Feb 2018. doi: 10.1103/PhysRevLett.120.050506.

Guifré Vidal and Rolf Tarrach. Robustness of entanglement. *Phys. Rev. A*, 59:141–155, Jan 1999. doi: 10.1103/PhysRevA.59.141.