

Lecture 7: Variational and convex methods for ground states

Computational aspects of quantum simulation
ENS de Lyon

2026

Abstract

This lecture studies methods for estimating the ground energy and finding ground states of a local Hamiltonian. Variational methods construct states and therefore give upper bounds on the ground energy. Convex relaxations discard some global consistency conditions and give lower bounds. We develop a hierarchy based on expectation values of Pauli operators and its dual interpretation as a sum-of-squares certificate. A two-qubit example shows how an upper and a lower bound can meet and certify the exact answer.

Main message. For a trial state $|\psi\rangle$ and a valid convex lower bound E_{low} ,

$$E_{\text{low}} \leq E_0 \leq \langle \psi | H | \psi \rangle.$$

The right inequality is constructive: it comes with a state. The left inequality is a certificate: it applies to every state. When the two values are close, the ground energy has been determined accurately.

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

- distinguish variational upper bounds from certified lower bounds;
- formulate a basic semidefinite relaxation of a classical or quantum ground-energy problem;
- construct a Pauli moment matrix and explain why it is positive semidefinite;
- read a sum-of-squares identity as a certificate of a lower bound.

Contents

1	Exercise from Lecture 6: a classical constraint Hamiltonian	2
1.1	Statement	2
1.2	Solution	2
2	Variational methods: constructing upper bounds	3
2.1	Product states and MPS	3
2.2	A parameterized-circuit ansatz	3
2.3	Worked example: a product ansatz can miss the ground state	4
3	Convex relaxations: certifying lower bounds	5
3.1	The exact but exponentially large convex problem	5
3.2	A first example: the vector relaxation of MaxCut	5

4	Pauli moment matrices and sums of squares	6
4.1	The moment matrix of a state	6
4.2	The moment relaxation at level r	7
4.3	The main theorem	7
4.4	A complete two-qubit certificate	10

1 Exercise from Lecture 6: a classical constraint Hamiltonian

1.1 Statement

Let $G = (V, E)$ be a finite undirected graph, with one Boolean variable x_i and one qubit at each vertex. Each edge $\{i, j\}$ imposes the constraint $x_i \neq x_j$. Define

$$h_{ij} = \frac{I + Z_i Z_j}{2}, \quad H_G = \sum_{\{i,j\} \in E} h_{ij}.$$

Show that h_{ij} is a projector and that its value on $|x\rangle$ is one exactly when the constraint is violated. Prove that

$$E_0(H_G) = |E| - \text{MaxCut}(G),$$

and explain why a quantum superposition cannot have energy below the best classical assignment.

1.2 Solution

Put $P_{ij} = Z_i Z_j$. Since $P_{ij}^2 = I$,

$$h_{ij}^2 = \frac{(I + P_{ij})^2}{4} = \frac{I + P_{ij}}{2} = h_{ij}.$$

Thus h_{ij} is the projector onto the $+1$ eigenspace of $Z_i Z_j$. On a computational-basis vector,

$$Z_i Z_j |x\rangle = (-1)^{x_i + x_j} |x\rangle.$$

Consequently,

$$\langle x | h_{ij} | x \rangle = \begin{cases} 1, & x_i = x_j, \\ 0, & x_i \neq x_j. \end{cases}$$

The energy of $|x\rangle$ is therefore the number of edges that do not cross the cut defined by x . Minimizing this number is the same as maximizing the number of crossing edges, and hence

$$\min_x \langle x | H_G | x \rangle = |E| - \text{MaxCut}(G).$$

Finally, H_G is diagonal in the computational basis. If $|\psi\rangle = \sum_x \alpha_x |x\rangle$ is normalized, then

$$\langle \psi | H_G | \psi \rangle = \sum_x |\alpha_x|^2 \langle x | H_G | x \rangle.$$

This is a convex combination of classical energies, so it cannot be below their minimum. The ground space is the span of the basis states describing maximum cuts.

2 Variational methods: constructing upper bounds

Let H be an n -qubit Hamiltonian and $E_0 = \lambda_{\min}(H)$. Lecture 6 established the variational identity

$$E_0 = \min_{\|\psi\|=1} \langle \psi | H | \psi \rangle.$$

It has an immediate algorithmic consequence: every normalized state that we can construct gives a rigorous upper bound

$$E_0 \leq E(\psi) := \langle \psi | H | \psi \rangle.$$

A variational method selects a family of states $\{|\psi(\theta)\rangle : \theta \in \Theta\}$, evaluates

$$E(\theta) = \langle \psi(\theta) | H | \psi(\theta) \rangle,$$

and searches for parameters of small energy.

There are two logically distinct issues:

1. *Expressiveness*: does the family contain a good approximation to a ground state?
2. *Optimization*: can we find good parameters in that family?

Even if the parameter search is nonconvex and stops at a local minimum, the energy of the state actually found remains a valid upper bound. What may be lost is the guarantee that it is the best upper bound available within the family.

2.1 Product states and MPS

A product-state ansatz has the form

$$|\psi(\theta)\rangle = |\psi_1(\theta_1)\rangle \otimes \cdots \otimes |\psi_n(\theta_n)\rangle.$$

For a local Hamiltonian, its energy is obtained from small local expectation values. Product states are easy to store and evaluate, but they cannot represent entanglement.

An MPS of bond dimension χ is a richer variational family. Lecture 5 showed how to compute local expectation values efficiently when χ is moderate. Bond dimension one gives product states, and increasing χ can only improve the optimal variational energy. Nevertheless, optimizing the tensor entries is still a nonconvex problem. Algorithms used in practice optimize one tensor, or a small block of tensors, while keeping the others fixed and sweep repeatedly through the chain.

2.2 A parameterized-circuit ansatz

Variational states can also be prepared by a parameterized quantum circuit. For the constraint Hamiltonian H_G , define the mixing Hamiltonian

$$B = \sum_{i \in V} X_i$$

and, for an integer $p \geq 1$, consider

$$|\psi_p(\boldsymbol{\beta}, \boldsymbol{\gamma})\rangle = e^{-i\beta_p B} e^{-i\gamma_p H_G} \cdots e^{-i\beta_1 B} e^{-i\gamma_1 H_G} |+\rangle^{\otimes n}. \quad (1)$$

This is the alternating-operator ansatz underlying the quantum approximate optimization algorithm (QAOA). The angles $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)$ and $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_p)$ are chosen to minimize

$$E_p(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \langle \psi_p(\boldsymbol{\beta}, \boldsymbol{\gamma}) | H_G | \psi_p(\boldsymbol{\beta}, \boldsymbol{\gamma}) \rangle.$$

The circuit is explicit and local. All terms in B commute, and all terms in the diagonal Hamiltonian H_G commute, so one alternating layer decomposes as

$$e^{-i\beta B} = \prod_{i \in V} e^{-i\beta X_i}, \quad e^{-i\gamma H_G} = \prod_{\{i,j\} \in E} e^{-i\gamma h_{ij}}.$$

The energy can be estimated by measuring in the computational basis and counting violated constraints. At depth zero the state is uniform over all assignments and has expected energy $|E|/2$. Increasing p enlarges the family, since an additional layer can be made trivial by setting its two angles to zero. Thus the best variational energy cannot increase with p , although finding the best angles remains a nonconvex optimization problem and no small value of p is guaranteed to find the ground state.

2.3 Worked example: a product ansatz can miss the ground state

Consider

$$H_{\text{Bell}} = -X_1 X_2 - Z_1 Z_2.$$

Let $|u\rangle|v\rangle$ be a product state, and denote its two Bloch vectors by $u, v \in \mathbb{R}^3$. Thus $u_x = \langle u|X|u\rangle$, $u_z = \langle u|Z|u\rangle$, and similarly for v . To see explicitly why a pure-state Bloch vector has norm one, write

$$|u\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1.$$

A direct calculation gives

$$u_x = 2 \operatorname{Re}(\bar{\alpha}\beta), \quad u_y = 2 \operatorname{Im}(\bar{\alpha}\beta), \quad u_z = |\alpha|^2 - |\beta|^2.$$

Consequently,

$$\begin{aligned} u_x^2 + u_y^2 + u_z^2 &= 4|\alpha|^2|\beta|^2 + (|\alpha|^2 - |\beta|^2)^2 \\ &= (|\alpha|^2 + |\beta|^2)^2 = 1. \end{aligned}$$

In particular,

$$u_x^2 + u_z^2 = 1 - u_y^2 \leq 1,$$

and the same argument gives $v_x^2 + v_z^2 \leq 1$. Its energy is

$$\langle u|\langle v|H_{\text{Bell}}|u\rangle|v\rangle = -(u_x v_x + u_z v_z).$$

By Cauchy–Schwarz and the inequalities just proved,

$$u_x v_x + u_z v_z \leq \sqrt{u_x^2 + u_z^2} \sqrt{v_x^2 + v_z^2} \leq 1.$$

The best product-state energy is therefore -1 , attained for example by $|+\rangle|+\rangle$.

On the other hand, the Bell state

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

satisfies

$$X_1 X_2 |\Phi^+\rangle = |\Phi^+\rangle, \quad Z_1 Z_2 |\Phi^+\rangle = |\Phi^+\rangle.$$

It has energy -2 . Thus

$$E_0(H_{\text{Bell}}) \leq -2,$$

whereas the product-state ansatz only proves $E_0(H_{\text{Bell}}) \leq -1$. A two-qubit MPS of bond dimension two contains $|\Phi^+\rangle$ and recovers the better upper bound.

3 Convex relaxations: certifying lower bounds

A trial state only proves that some state has small energy. A lower bound must prove that *every* normalized state has energy at least a given number. This is the role of convex relaxations.

3.1 The exact but exponentially large convex problem

A density matrix is a matrix $\rho \succeq 0$ with $\text{Tr } \rho = 1$. Pure states correspond to $\rho = |\psi\rangle\langle\psi|$, and their energy is $\text{Tr}(H\rho)$. We can therefore write

$$E_0 = \min\{\text{Tr}(H\rho) : \rho \succeq 0, \text{Tr } \rho = 1\}. \quad (2)$$

Indeed, in an eigenbasis of H , the objective is a convex combination of its eigenvalues. Equality is attained by any density matrix supported on the ground space.

Problem (2) is a *semidefinite program*: the objective and equality constraints are linear, and the matrix variable is required to be positive semidefinite. It is convex, but ρ has dimension 2^n . Convexity alone does not make the formulation efficient.

A relaxation keeps a polynomial number of necessary conditions satisfied by every state and drops the remaining global consistency conditions. The feasible set becomes larger. Because we are minimizing, the optimum can only decrease, producing a lower bound on E_0 .

3.2 A first example: the vector relaxation of MaxCut

Return to the constraint Hamiltonian H_G . Write $z_i = (-1)^{x_i} \in \{-1, 1\}$. Its classical energy is

$$\frac{1}{2} \sum_{\{i,j\} \in E} (1 + z_i z_j).$$

Introduce the Gram matrix $\Gamma = zz^T$. It satisfies

$$\Gamma \succeq 0, \quad \Gamma_{ii} = 1, \quad \text{rank}(\Gamma) = 1.$$

Conversely, a real rank-one matrix with these properties comes from such a vector z , up to an irrelevant global sign. Thus the classical problem can be written exactly as

$$E_0(H_G) = \min \left\{ \frac{1}{2} \sum_{\{i,j\} \in E} (1 + \Gamma_{ij}) : \Gamma \succeq 0, \Gamma_{ii} = 1, \text{rank}(\Gamma) = 1 \right\}.$$

The rank constraint is nonconvex. Dropping it gives the SDP

$$E_{\text{sdp}}(G) = \min \left\{ \frac{1}{2} \sum_{\{i,j\} \in E} (1 + \Gamma_{ij}) : \Gamma \succeq 0, \Gamma_{ii} = 1 \right\}. \quad (3)$$

Every classical assignment remains feasible, so

$$E_{\text{sdp}}(G) \leq E_0(H_G).$$

There is a useful geometric interpretation. The constraints $\Gamma \succeq 0$ and $\Gamma_{ii} = 1$ mean that Γ is a Gram matrix:

$$\Gamma_{ij} = \langle v_i, v_j \rangle$$

for some unit vectors v_i . The discrete choices $z_i \in \{-1, 1\}$ have been replaced by arbitrary unit vectors.

The triangle. For the triangle, every cut leaves at least one edge uncut, so $E_0(H_{K_3}) = 1$. In the vector relaxation,

$$0 \leq \|v_1 + v_2 + v_3\|^2 = 3 + 2 \sum_{1 \leq i < j \leq 3} \langle v_i, v_j \rangle.$$

It follows that the relaxed objective is at least

$$\frac{3}{2} + \frac{1}{2} \left(-\frac{3}{2} \right) = \frac{3}{4}.$$

Three coplanar vectors separated by angles of 120° attain this value. Therefore

$$E_{\text{sdp}}(K_3) = \frac{3}{4} < 1 = E_0(H_{K_3}).$$

The SDP gives a valid lower bound, but the relaxation is not exact.

4 Pauli moment matrices and sums of squares

The vector relaxation retains pairwise products $z_i z_j$ and enforces the positivity of their Gram matrix. The quantum analogue retains expectation values of products of simple operators.

Let

$$\mathcal{P}_n = \{I, X, Y, Z\}^{\otimes n}$$

be the set of Hermitian n -qubit Pauli strings, and let $\text{wt}(P)$ be the number of nonidentity tensor factors in P . For an integer r , put

$$\mathcal{B}_r = \{P \in \mathcal{P}_n : \text{wt}(P) \leq r\}, \quad \mathcal{V}_r = \text{span}_{\mathbb{C}}(\mathcal{B}_r).$$

4.1 The moment matrix of a state

For a density matrix ρ , define a matrix indexed by $P, Q \in \mathcal{B}_r$ by

$$\Gamma^{(r)}(\rho)_{P,Q} = \text{Tr}(\rho P^\dagger Q) = \text{Tr}(\rho P Q). \quad (4)$$

Although P and Q are Hermitian, their product need not be Hermitian, so the entries of $\Gamma^{(r)}$ may be complex. The matrix itself is Hermitian because

$$\Gamma_{Q,P}^{(r)} = \overline{\Gamma_{P,Q}^{(r)}}.$$

More importantly, it is positive semidefinite. For coefficients $(c_P)_{P \in \mathcal{B}_r}$, set $A = \sum_P c_P P$. Then

$$\sum_{P,Q} \overline{c_P} \Gamma_{P,Q}^{(r)} c_Q = \text{Tr}(\rho A^\dagger A) \geq 0. \quad (5)$$

The entries are not independent. Every product of Hermitian Pauli strings has a unique normal form

$$PQ = \omega(P, Q) R(P, Q), \quad \omega(P, Q) \in \{1, -1, i, -i\}, \quad R(P, Q) \in \mathcal{P}_n.$$

The phase and the Hermitian Pauli string are unique because $\mathcal{P}_n = \{I, X, Y, Z\}^{\otimes n}$ contains exactly one representative of each Pauli operator up to phase. Consequently,

$$PQ = P'Q' \iff \omega(P, Q) = \omega(P', Q') \text{ and } R(P, Q) = R(P', Q').$$

This gives a systematic way to determine which moment-matrix entries must be identified. In fact,

$$\Gamma_{P,Q}^{(r)} = \omega(P, Q) \operatorname{Tr}(\rho R(P, Q)).$$

For example,

$$\Gamma_{X_j, Z_j}^{(r)} = -i\Gamma_{I, Y_j}^{(r)}, \quad \Gamma_{Z_j, X_j}^{(r)} = i\Gamma_{I, Y_j}^{(r)}, \quad \Gamma_{X_j, X_j}^{(r)} = \Gamma_{I, I}^{(r)} = 1,$$

and, for $i \neq j$,

$$\Gamma_{X_i, Z_j}^{(r)} = \Gamma_{Z_j, X_i}^{(r)}.$$

Additional equalities can hold for a particular state, but the equalities above are forced solely by the Pauli multiplication rules.

4.2 The moment relaxation at level r

Introduce a real variable y_R for every Hermitian Pauli string of weight at most $2r$, intended to represent $\operatorname{Tr}(\rho R)$, and impose $y_I = 1$. Define $\Gamma^{(r)}(y)$ by Pauli multiplication:

$$\Gamma^{(r)}(y)_{P,Q} = \omega(P, Q) y_{R(P,Q)} \quad \text{when} \quad PQ = \omega(P, Q) R(P, Q).$$

This definition automatically identifies entries that represent the same operator, including the phase arising from multiplication.

Suppose

$$H = \sum_R h_R R$$

is k -local, with real coefficients h_R , and choose r so that $2r \geq k$. The level- r Pauli moment relaxation is

$$E_{\text{sdp}}^{(r)}(H) = \min_y \left\{ \sum_R h_R y_R : y_I = 1, \Gamma^{(r)}(y) \succeq 0 \right\}. \quad (6)$$

This is an SDP: the objective is linear in y , and every entry of the matrix is linear in y .

4.3 The main theorem

Theorem 4.1 (Pauli moment and sum-of-squares hierarchy). *Let H be a k -local Hamiltonian on n qubits. For $2r \geq k$, let $E_{\text{sdp}}^{(r)}(H)$ be defined by (6). Then*

$$E_{\text{sdp}}^{(r)}(H) \leq E_{\text{sdp}}^{(r+1)}(H) \leq E_0(H), \quad E_{\text{sdp}}^{(n)}(H) = E_0(H).$$

Moreover, the same number $E_{\text{sdp}}^{(r)}(H)$ is the largest $\lambda \in \mathbb{R}$ for which there are operators $A_1, \dots, A_s \in \mathcal{V}_r$ satisfying the operator identity

$$H - \lambda I = \sum_{a=1}^s A_a^\dagger A_a. \quad (7)$$

Thus the primal SDP searches over locally consistent moments, while the dual SDP searches for a sum-of-squares certificate of a lower bound.

Proof. Every level gives a lower bound. Let ρ be any density matrix and put $y_R = \operatorname{Tr}(\rho R)$. Equations (4) and (5) show that y is feasible for the relaxation, and its objective value is

$$\sum_R h_R y_R = \operatorname{Tr}(\rho H).$$

In particular, the moments of a ground state are feasible and have objective value $E_0(H)$. Since the SDP minimizes over a set that may also contain moments not arising from a state,

$$E_{\text{sdp}}^{(r)}(H) \leq E_0(H).$$

The bounds increase with the level. A feasible level- $(r+1)$ collection of moments restricts to a feasible level- r collection: take the principal submatrix indexed by $\mathcal{B}_r$. Hence the level- $(r+1)$ feasible set has no more freedom on the moments relevant to the objective, and

$$E_{\text{sdp}}^{(r)}(H) \leq E_{\text{sdp}}^{(r+1)}(H).$$

The full level is exact. At level n , all Pauli strings occur. A feasible collection y defines a linear functional L on all $2^n \times 2^n$ matrices by $L(R) = y_R$ on Pauli strings and linear extension. Positivity of the full moment matrix says

$$L(A^\dagger A) \geq 0 \quad \text{for every matrix } A.$$

Define

$$\rho = \frac{1}{2^n} \sum_{R \in \mathcal{P}_n} y_R R.$$

The Pauli strings are orthogonal for the Hilbert–Schmidt inner product:

$$\text{Tr}(RS) = 2^n \delta_{R,S} \quad (R, S \in \mathcal{P}_n).$$

Therefore, for every Pauli string S ,

$$\text{Tr}(\rho S) = \frac{1}{2^n} \sum_{R \in \mathcal{P}_n} y_R \text{Tr}(RS) = y_S = L(S).$$

Orthogonality is what recovers the coefficient y_S . We now use the second fact: the Pauli strings form a basis of the full matrix algebra. If $A = \sum_S a_S S$ is an arbitrary matrix, linearity gives

$$\text{Tr}(\rho A) = \sum_S a_S \text{Tr}(\rho S) = \sum_S a_S L(S) = L(A).$$

In particular, $\text{Tr} \rho = L(I) = y_I = 1$. If $|v\rangle$ is any normalized vector, then $|v\rangle\langle v|$ is positive and equals its own square, so

$$\langle v | \rho | v \rangle = L(|v\rangle\langle v|) \geq 0.$$

Thus $\rho \succeq 0$: the full moments come from an actual density matrix. The exact formulation (2) then gives $E_{\text{sdp}}^{(n)}(H) \geq E_0(H)$, proving equality.

The dual certificate and strong duality. It is useful first to prove the easy direction, usually called *weak duality*. A feasible collection y defines a linear functional L_y by $L_y(R) = y_R$ on Pauli strings and linear extension. Its moment matrix satisfies

$$\Gamma^{(r)}(y)_{P,Q} = L_y(P^\dagger Q).$$

Suppose that, for some real λ and some matrix $M \succeq 0$,

$$H - \lambda I = \sum_{P,Q \in \mathcal{B}_r} M_{P,Q} P^\dagger Q. \tag{8}$$

Factor M as a Gram matrix,

$$M_{P,Q} = \sum_a \overline{c_{a,P}} c_{a,Q}.$$

Applying L_y to (8) gives

$$\begin{aligned} L_y(H) - \lambda &= \sum_{P,Q} M_{P,Q} \Gamma^{(r)}(y)_{P,Q} \\ &= \sum_a c_a^\dagger \Gamma^{(r)}(y) c_a \geq 0. \end{aligned}$$

Thus $\lambda \leq L_y(H)$ for every feasible y , and therefore

$$\lambda \leq E_{\text{sdp}}^{(r)}(H) \leq E_0(H).$$

This already explains why every feasible dual point is a certified lower bound.

The Gram factorization also explains the sum-of-squares form. Define

$$A_a = \sum_{P \in \mathcal{B}_r} c_{a,P} P \in \mathcal{V}_r.$$

Then the right-hand side of (8) is

$$\sum_a A_a^\dagger A_a.$$

Conversely, expanding any sum of squares with $A_a \in \mathcal{V}_r$ produces a positive semidefinite Gram matrix M . Therefore the SDP dual of (6) can be written explicitly as

$$\begin{aligned} &\text{maximize} \quad \lambda \\ &\text{subject to} \quad M \succeq 0, \\ &\quad H - \lambda I = \sum_{P,Q \in \mathcal{B}_r} M_{P,Q} P^\dagger Q. \end{aligned} \tag{9}$$

The scalar λ is the dual variable for the normalization constraint $y_I = 1$; equality of the coefficient of each Pauli string in the last line is the dual constraint associated with the corresponding moment variable y_R .

Weak duality only shows that the optimum of (9) is at most $E_{\text{sdp}}^{(r)}(H)$. We use strong SDP duality to show that there is no gap. Take the moments of the maximally mixed state $\rho_* = I/2^n$. They satisfy

$$y_I = 1, \quad y_R = 0 \quad (R \neq I),$$

and their moment matrix is

$$\Gamma^{(r)}(\rho_*)_{P,Q} = \frac{1}{2^n} \text{Tr}(PQ) = \delta_{P,Q}.$$

Hence $\Gamma^{(r)}(\rho_*) = I \succ 0$: the primal SDP has a strictly feasible point. Its optimum is finite: every diagonal entry of a feasible moment matrix equals one, so positivity implies $|\Gamma_{P,Q}^{(r)}| \leq 1$, and every moment occurring in the objective appears as such an entry because $2r \geq k$. Slater's condition for semidefinite programs therefore implies that the primal and dual optimal values are equal and that the dual optimum is attained. This is exactly where strong duality is used. It follows that $E_{\text{sdp}}^{(r)}(H)$ is the largest λ admitting the sum-of-squares identity (7). $\square$

4.4 A complete two-qubit certificate

Return to

$$H_{\text{Bell}} = -X_1 X_2 - Z_1 Z_2.$$

The variational calculation gave the upper bound $E_0 \leq -2$. For a matching lower bound, use $X_j^2 = Z_j^2 = I$ and the fact that operators on different qubits commute:

$$\begin{aligned}\frac{1}{2}(X_1 - X_2)^2 &= I - X_1 X_2, \\ \frac{1}{2}(Z_1 - Z_2)^2 &= I - Z_1 Z_2.\end{aligned}$$

Adding the two identities gives

$$H_{\text{Bell}} + 2I = \frac{1}{2}(X_1 - X_2)^2 + \frac{1}{2}(Z_1 - Z_2)^2 \succeq 0. \quad (10)$$

Hence $E_0 \geq -2$. Together with the Bell-state upper bound,

$$E_0(H_{\text{Bell}}) = -2.$$

The operators $(X_1 - X_2)/\sqrt{2}$ and $(Z_1 - Z_2)/\sqrt{2}$ are linear combinations of weight-one Paulis. Thus (10) is already a level-one dual certificate, and the level-one relaxation is exact for this example.

The Bell state proves that an energy of -2 is achievable, while the sum-of-squares identity proves that no state can have lower energy. Neither direction alone establishes the exact answer.