

# Sharp partial-transpose geometry of two-qubit states, with the relative entropy of entanglement of Werner states as an application

scientific-author-45

Corresponding author, GeWu project owner

September 9, 2026

## Abstract

How much can the partial transpose of a two-qubit operator move its spectrum, measured against the trace norm of the operator itself? We answer this exactly for traceless Hermitian operators: the operator norm of the partial transpose is at most  $(1 + \sqrt{2})/4 \approx 0.6036$  times the trace norm, the constant is optimal, and an explicit extremiser attains it. Everything else follows from that single inequality. The Peres–Horodecki criterion becomes quantitative: if a state  $\rho$  is compared with any positive-partial-transpose (PPT) state  $\sigma$  whose partial transpose has smallest eigenvalue  $g$ , then the smallest eigenvalue of  $\rho^\Gamma$  is at least  $g - \frac{1+\sqrt{2}}{4}\|\rho - \sigma\|_1$ , and the negativity of  $\rho$  is at most  $[\frac{1+\sqrt{2}}{4}\|\rho - \sigma\|_1 - g]_+$ . Each separable Werner state with singlet fraction  $\lambda$  is the centre of a certified separable ball of trace-norm radius  $4(\sqrt{2} - 1)\eta(\lambda)$ , where  $\eta(\lambda)$  is its partial-transpose gap; this improves the elementary radius  $\eta(\lambda)$  and replaces a doubled radius that we show to be false by an explicit counterexample. The same constant makes the distance from  $\rho_\lambda$  to the nearest non-PPT state exactly  $\eta(\lambda)/K$  on  $[\lambda_c, \frac{1}{4}]$ , with  $\lambda_c = (3\sqrt{2} - 4)/2$  and the endpoint value  $d(\frac{1}{4}) = \sqrt{2} - 1$ , and an explicit family gives the exact value  $\frac{2}{3}(2\sqrt{(1-\lambda)(1+2\lambda)} - (2\lambda+1))$  on  $[\frac{1}{4}, \frac{1}{2}]$ ; below  $\lambda_c$  we prove that the Weyl bound is strict,  $d(\lambda) > \eta(\lambda)/K$ , and certify upper bounds from two explicit families, but the exact value is open. As an application of the same PPT and twirling data we re-derive the classical Werner value of the relative entropy of entanglement,  $E_R(\rho_\lambda) = \ln 2 - H_{\text{nat}}(\lambda)$  for  $\lambda > \frac{1}{2}$  and 0 otherwise, classify its closest separable states completely, and prove for every two-qubit state the two-sided bound  $E_R(\rho) \leq F(\lambda(\rho)) + D(\rho\|\rho_{\lambda(\rho)})$  above and  $F(\lambda(\rho))$  below, where  $\lambda(\rho)$  is the singlet fraction and the correction is controlled uniformly by a trace-norm modulus. A closed-form rule for the closest separable state of an arbitrary two-qubit state is not claimed and remains open.

## 1 Introduction

A two-qubit state is separable exactly when its partial transpose is positive: this is the Peres–Horodecki criterion [1, 2], and on  $\mathbb{C}^2 \otimes \mathbb{C}^2$  it is an if-and-only-if, one of the very few dimensions where a structural optimisation over separable states collapses to a linear-algebra test. The criterion is, however, qualitative. It says that the separable set and the positive-partial-transpose (PPT) set coincide; it says nothing about *how far* a given state sits from that common boundary, nothing about how the boundary moves when the state is perturbed, and nothing about how large a neighbourhood of a given separable state is again separable. Every quantitative question of this kind—the size of separable balls, the continuity of entanglement measures, the distance to the nearest entangled state—needs a *norm* statement about the partial transpose, and the obvious candidate statements are either false or far from sharp.

This paper supplies the missing norm statement and then spends it.

## The sharp constant

The central object is the linear involution  $\Gamma$  that transposes one tensor factor. It is unital and trace preserving, so it leaves the trace of every operator alone, but it is *not* an isometry for the trace norm: we exhibit in Section 5 a  $4 \times 4$  Hermitian matrix  $X$  with  $\|X\|_1 = 4$  and  $\|X^\Gamma\|_1 = 2\sqrt{2}$ . Any bound of the form  $\|X^\Gamma\|_{\text{op}} \leq c \|X\|_1$  therefore has to be proved, not inherited, and the natural restriction to traceless operators—which is where all differences of states live—turns out to have an exact answer.

*For every traceless Hermitian operator  $X$  on  $\mathbb{C}^2 \otimes \mathbb{C}^2$ ,*

$$\|X^\Gamma\|_{\text{op}} \leq K \|X\|_1, \quad K = \frac{1 + \sqrt{2}}{4} = 0.6035533905932737\dots, \quad (1)$$

*equivalently  $\lambda_{\min}(X^\Gamma) \geq -K \|X\|_1$ . The constant  $K$  is optimal: equality holds for*

$$X_* = \frac{1}{2}(|\phi^-\rangle\langle\phi^-| - |01\rangle\langle 01|), \quad \|X_*\|_1 = 1, \quad \|X_*^\Gamma\|_{\text{op}} = K,$$

*and for this operator  $\lambda_{\min}(X_*^\Gamma) = -K$ , so the bound is attained at the negative end of the spectrum.*

The proof is short and is the reason the rest of the paper works at all. Conjugating a pure state by local unitaries conjugates its partial transpose by local unitaries, so the spectrum of  $(|\psi\rangle\langle\psi|)^\Gamma$  depends only on the two Schmidt weights of  $\psi$ ; it is  $\{t, 1-t, \pm\sqrt{t(1-t)}\}$ . Because  $X$  is traceless, one may subtract any multiple of the identity from  $(|\psi\rangle\langle\psi|)^\Gamma$  inside  $\text{Tr}[X(|\psi\rangle\langle\psi|)^\Gamma]$ , and subtracting the midpoint of the spectrum turns the optimisation into the half-width of that spectrum,  $\frac{1}{2}(t + \sqrt{t(1-t)})$ , which is maximised at  $t_* = (2 + \sqrt{2})/4$ , where it equals  $K = (1 + \sqrt{2})/4$ . Tracelessness is what makes the centring legal, and it is the only place where the constant 1 degrades to  $K$ : for a positive operator such as  $|00\rangle\langle 00|$  the left side is already 1.

## What the constant buys

**Quantitative Peres–Horodecki.** Let  $\sigma$  be any PPT state and  $g = \lambda_{\min}(\sigma^\Gamma) \geq 0$ . Weyl’s inequality plus (1) give, for every two-qubit state  $\rho$ ,

$$\lambda_{\min}(\rho^\Gamma) \geq g - K \|\rho - \sigma\|_1,$$

so  $\rho$  is PPT—hence separable—as soon as  $\|\rho - \sigma\|_1 < g/K$ . Because a two-qubit partial transpose carries at most one negative eigenvalue [3, 4], the negativity is  $N(\rho) = [-\lambda_{\min}(\rho^\Gamma)]_+$  and the same estimate reads

$$N(\rho) \leq [K \|\rho - \sigma\|_1 - g]_+.$$

Applied to the separable Werner state  $\rho_\lambda$  with  $\eta(\lambda) = \min\{\frac{1-2\lambda}{2}, \frac{1+2\lambda}{6}\}$  this yields a certified separable ball of trace-norm radius  $\eta(\lambda)/K = 4(\sqrt{2} - 1)\eta(\lambda) \approx 1.657 \eta(\lambda)$  around *every* separable Werner state. The elementary Weyl argument gives radius  $\eta(\lambda)$ ; a doubled radius  $2\eta(\lambda)$ , which would follow from the false claim that  $\Gamma$  is a trace-norm isometry, is refuted by an explicit NPT state at distance 0.4440 from  $\rho_{0.24}$ , less than  $2\eta(0.24) = 0.4933$  but more than the correct radius  $\eta(0.24)/K = 0.4087$ . Section 5 states the counterexample and the corrected radius.

**Distance to the nearest entangled state.** Define, for the separable Werner state  $\rho_\lambda$  with  $\lambda < \frac{1}{2}$ ,

$$d(\lambda) = \inf\{\|\sigma - \rho_\lambda\|_1 : \sigma \text{ a two-qubit state, } \sigma \text{ NPT}\},$$

the trace norm of the difference being twice the usual trace distance. Weyl plus (1) lower bound it by  $\eta(\lambda)/K$ , and on  $[\lambda_c, \frac{1}{4}]$  that lower bound is attained by the explicit ray  $\sigma(\lambda) = \rho_\lambda + (\eta(\lambda)/K)X_*$ , where  $\lambda_c = (3\sqrt{2} - 4)/2 = 0.12132034\dots$  is the largest  $\lambda$  for which that ray is still positive semidefinite. Hence

$$d(\lambda) = \frac{\eta(\lambda)}{K} = 4(\sqrt{2} - 1) \frac{1 + 2\lambda}{6} \quad \text{on } [\lambda_c, \tfrac{1}{4}], \quad d(\tfrac{1}{4}) = \sqrt{2} - 1, \quad d(\lambda_c) = 6 - 4\sqrt{2}.$$

On the upper interval  $[\frac{1}{4}, \frac{1}{2})$  the distance is also exact,

$$d(\lambda) = d_{\text{up}}(\lambda) = \frac{2}{3} \left( 2\sqrt{(1-\lambda)(1+2\lambda)} - (2\lambda + 1) \right),$$

attained by  $\rho_\lambda + c(\lambda)(|\psi^-\rangle\langle\psi^-| - |11\rangle\langle 11|)$  with  $c(\lambda) = \frac{2}{3}(\sqrt{(1-\lambda)(1+2\lambda)} - \frac{2\lambda+1}{2})$ . Below  $\lambda_c$  the ray leaves the state space and the exact value is *open*: this paper proves the certified upper bounds  $d(\lambda) \leq \delta_{\text{low}}(\lambda)$  on  $(0, \lambda_0]$  and  $d(\lambda) \leq d_{\text{iso}}(\lambda)$  on  $[0, \frac{1}{4}]$  from explicit double-corner states, with  $\delta_{\text{low}}(0) = 2/(3 + 2\sqrt{3}) = 0.30940\dots$ , the crossover  $\lambda_0 = (7 - 3\sqrt{5})/4 = 0.072949\dots$  where  $\delta_{\text{low}} = d_{\text{iso}}$ , and a strict inequality  $d(\lambda) > \eta(\lambda)/K$  on  $(0, \lambda_c)$ , proved in Appendix A.4: a Schur complement reduces the positivity of any equality candidate to the single condition  $F(\lambda) \leq 1$  for an explicit rational function, and the lower root of  $F(\lambda) = 1$  is exactly  $\lambda_c$ . The picture is Figure 1.

**Werner relative entropy of entanglement and a bound for every state.** With  $F(\lambda) = E_R(\rho_\lambda)$  written out, twirling reduces the defining minimisation to one variable and gives the classical value

$$F(\lambda) = \begin{cases} 0, & 0 \leq \lambda \leq \frac{1}{2}, \\ \ln 2 - H_{\text{nat}}(\lambda), & \frac{1}{2} < \lambda \leq 1, \end{cases} \quad H_{\text{nat}}(\lambda) = -\lambda \ln \lambda - (1 - \lambda) \ln(1 - \lambda),$$

so  $E_R$  equals one ebit at the Bell point and  $0.1308120359 \text{ nats} = 0.1887218755 \text{ ebits}$  at  $\lambda = \frac{3}{4}$ . The closest separable state is unique for every  $0 < \lambda < 1$ , equals  $\rho_\lambda$  for  $\lambda \leq \frac{1}{2}$  and  $\sigma_* = \frac{1}{2}P_- + \frac{1}{6}P_{\text{sym}}$  for  $\frac{1}{2} < \lambda < 1$ , and becomes a two-parameter family exactly at the Bell point  $\lambda = 1$ , where the closest separable states are precisely the separable  $\sigma$  with  $\sigma|\psi^-\rangle = |\psi^-\rangle/2$ . This part of the paper is an application, not a novelty claim: it reproduces results of [5, 6]. What is new is the use to which it is put. For *every* two-qubit state  $\rho$ , with singlet fraction  $\lambda(\rho) = \text{Tr}[P_- \rho]$ ,

$$\boxed{F(\lambda(\rho)) \leq E_R(\rho) \leq F(\lambda(\rho)) + D(\rho \| \rho_{\lambda(\rho)})}$$

and the correction is an entropy deficit,  $D(\rho \| \rho_{\lambda(\rho)}) = F(\lambda(\rho)) + [S(\rho_\lambda) - S(\rho)]$ , controlled uniformly by the Fannes–Audenaert modulus  $T \ln 3 + h(T)$  with  $T = \frac{1}{2} \|\rho - \rho_\lambda\|_1$  [7]. The two bounds coincide on the Werner line itself, so the sandwich is sharp there; and on the entangled half of that line  $E_R$  is Fréchet differentiable with derivative  $H \mapsto \text{Tr}[H(\ln \rho_\lambda - \ln \sigma_*)] = \text{Tr}[P_- H] \ln \frac{\lambda}{1-\lambda}$ .

## Scope

The source problem from which this study started asks for a closed-form rule that returns the closest separable state of an *arbitrary* two-qubit state [8, 9]. That question is *not* answered here, and we do not claim it. What is answered is narrower and complete on its own terms: the sharp norm of the partial transpose on traceless two-qubit operators, the continuity law and negativity bound it implies, the exact nearest-NPT distance on  $[\lambda_c, \frac{1}{2})$  with certified bounds below  $\lambda_c$ , and a two-sided singlet-fraction sandwich for all two-qubit states with an explicit modulus and a first-order expansion. Section 10 lists the four open directions this leaves.

Section 2 separates these results from the closest prior work; Section 3 fixes notation; Sections 4–7 carry the proofs; Section 8 is the all-state bound; Section 9 reports the numerical certificates; and Appendices A–C hold the derivations, the explicit boundary states, and the protocols.

## 2 Related work

Positivity of the partial transpose as a necessary separability test is Peres’ [1], and its sufficiency on  $\mathbb{C}^2 \otimes \mathbb{C}^2$  and  $\mathbb{C}^2 \otimes \mathbb{C}^3$  is the Horodeckis’ [2]. Both are statements about the *sign* of  $\lambda_{\min}(\rho^\Gamma)$ . The inequality (1) is a statement about its *magnitude* relative to the trace norm of a perturbation, and it is the input that converts the criterion into the continuity law of Theorem 5.1. The distinction matters: the criterion alone certifies separability only at a point where  $\lambda_{\min}(\rho^\Gamma)$  has already been computed exactly, whereas the norm bound certifies it from any nearby PPT state whose gap is known.

Balls of separable states are classical in a different form. Gurvits and Barnum [10] determine the largest ball of separable states centred at the maximally mixed state, working in a spectral norm and in arbitrary dimension. The balls of Corollary 5.3 are of another kind: they are centred at each separable Werner state  $\rho_\lambda$  with  $\lambda \leq \frac{1}{2}$ , their radii are stated in the trace norm, and they follow from a single operator inequality rather than from a volume or positivity argument. Because  $\rho_{1/4} = \frac{1}{4}I_4$ , the two families meet at  $\lambda = \frac{1}{4}$ , where the certified radius here is  $\eta(\frac{1}{4})/K = \sqrt{2} - 1$  in trace norm, that is  $(\sqrt{2} - 1)/2 = 0.2071$  in trace distance. No claim is made that this is the largest ball at that centre, and the comparison with the Gurvits–Barnum radius is not meaningful without converting norms, which is not attempted here. What is new is the factor  $1/K = 4(\sqrt{2} - 1)$  by which the sharp constant improves the elementary Weyl radius  $\eta(\lambda)$  uniformly along the whole separable Werner line, and the explicit refutation of the larger radius  $2\eta(\lambda)$  that a trace-norm-isometry claim for  $\Gamma$  would have given (Remark 5.5).

A second, closer body of work builds distances out of the partial transpose itself. Ganardi, Miller, Paterek and Żukowski [11] construct a hierarchy of correlation quantifiers from the distance to the set of states with positive partial transpose and show that in dimension  $2 \times 2$  the relevant member reproduces the negativity. Their object is a state-to-set distance evaluated after applying  $\Gamma$ ; the object here is the opposite direction, an upper bound on  $\Gamma$  itself as a map between two norms on traceless operators. The two are complementary rather than competing: their equality with negativity is an identity about a minimisation, while (1) is a pointwise inequality with a constant proved optimal by exhibiting an extremiser. The nearest-NPT distance  $d(\lambda)$  of Section 6 is, to our knowledge, not computed in closed form elsewhere; the closest standard quantity is the negativity itself, which measures the size of the negative part of  $\rho^\Gamma$  rather than how far one must move  $\rho$  to create one.

On the continuity side of the entanglement measure, Donald and Horodecki [12] prove continuity of  $E_R$ , and Audenaert [7] supplies the sharp trace-norm modulus for the von Neumann entropy,  $T \ln(d - 1) + h(T)$  with  $T = \frac{1}{2}\|\rho - \sigma\|_1$ . Combining the two gives a continuity estimate for  $E_R$  that is uniform but one-sided and dimension generic. The sandwich of Theorem 8.1 is sharper in a specific way: its lower edge is the exact Werner value  $F(\lambda(\rho))$ , the correction above that edge is the entropy deficit  $S(\rho_\lambda) - S(\rho) = D(\rho\|\rho_\lambda)$  rather than a generic modulus, and the Audenaert bound enters only to make the deficit uniform in  $\rho$ . Since the deficit vanishes exactly on the Werner line, both edges coincide there and the estimate is tight at every  $\lambda$ , whereas a modulus-only bound would leave an  $O(T \ln \frac{1}{T})$  slack on the very states where the exact value is known.

Two standard tools are used exactly as the literature states them, and nothing here improves them. That an optimisation over separable states can be restricted to symmetric competitors is the

twirling argument for group-invariant states [13], resting on monotonicity of the relative entropy under stochastic maps and its equality conditions [14]; both facts enter Section 7 as quoted, and the reduction is a means to the sandwich rather than a result. The value  $E_R(\rho_\lambda) = \ln 2 - H_{\text{nat}}(\lambda)$  is contained in the Bell-diagonal computation of Vedral, Plenio, Rippin and Knight [5], and the non-uniqueness of the closest separable state at the Bell point is established by Miranowicz and Ishizaka [6], who also give the inverse map from a closest-separable-state candidate back to the states it optimises and remark that a closed form for the closest separable state exists only in special cases. The contribution on this side is therefore threefold, and none of it is the value: the complete classification of closest separable states, including uniqueness for every  $0 < \lambda < 1$  and the endpoint  $\lambda = 0$ ; an operator-level cross-check that pushes the Miranowicz–Ishizaka inverse map through the entire entangled Werner ray and returns the same value to  $2 \times 10^{-14}$ ; and the observation that the Werner value together with the entropy-deficit identity produces a two-sided bound for *all* two-qubit states. The last is the point of the application. The Werner line is used as a calibrated ruler, not as the object of study.

### 3 Preliminaries

#### 3.1 States, the partial transpose, and the two-qubit criterion

Throughout,  $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$  with computational basis  $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ , and the four Bell vectors are

$$|\psi^\pm\rangle = \frac{|01\rangle \pm |10\rangle}{\sqrt{2}}, \quad |\phi^\pm\rangle = \frac{|00\rangle \pm |11\rangle}{\sqrt{2}}. \quad (2)$$

Write  $P_- = |\psi^-\rangle\langle\psi^-|$  for the singlet projector and  $P_{\text{sym}} = I_4 - P_- = |\psi^+\rangle\langle\psi^+| + |\phi^-\rangle\langle\phi^-| + |\phi^+\rangle\langle\phi^+|$  for the projector onto the three-dimensional symmetric subspace. A *state* is a positive semidefinite operator on  $\mathcal{H}$  with unit trace. The *separable* states are the convex hull of the product projectors,

$$\mathcal{S} = \left\{ \sum_k p_k |a_k \otimes b_k\rangle\langle a_k \otimes b_k| : p_k \geq 0, \sum_k p_k = 1 \right\},$$

a compact convex set.

The partial transpose with respect to the first factor is the linear map  $\Gamma$  defined on matrix units by

$$(|ab\rangle\langle cd|)^\Gamma = |cb\rangle\langle ad|, \quad \text{equivalently} \quad X_{ab,cd}^\Gamma = X_{cb,ad}, \quad (3)$$

where a compound index  $ab$  labels  $|a\rangle \otimes |b\rangle$ . It is an involution,  $(X^\Gamma)^\Gamma = X$ , it is Hermiticity preserving and unital,  $I_4^\Gamma = I_4$ , and it preserves the trace,  $\text{Tr } X^\Gamma = \text{Tr } X$ . A state is called *PPT* when  $\rho^\Gamma \geq 0$  and *NPT* otherwise.

**Lemma 3.1** (local covariance of  $\Gamma$ ). *For all operators  $X$  and all  $U, V \in \mathcal{B}(\mathbb{C}^2)$ ,  $((U \otimes V)X(U \otimes V)^\dagger)^\Gamma = (U^* \otimes V)X^\Gamma(U^* \otimes V)^\dagger$ . Consequently  $\Gamma$  preserves the spectrum of every operator of the form  $(U \otimes V)X(U \otimes V)^\dagger$  after partial transposition, up to the conjugating local unitary.*

*Proof.* For a product operator, (3) reads  $(A \otimes B)^\Gamma = A^\top \otimes B$ , where  $^\top$  is the ordinary matrix transpose in the fixed single-qubit basis. Taking  $X = |ab\rangle\langle cd| = |a\rangle\langle c| \otimes |b\rangle\langle d|$ ,

$$((U \otimes V)X(U \otimes V)^\dagger)^\Gamma = (U|a\rangle\langle c|U^\dagger)^\top \otimes V|b\rangle\langle d|V^\dagger = U^*(|a\rangle\langle c|)^\top (U^*)^\dagger \otimes V|b\rangle\langle d|V^\dagger,$$

using  $(UAU^\dagger)^\top = U^*A^\top(U^*)^\dagger$ , and the last line is  $(U^* \otimes V)X^\Gamma(U^* \otimes V)^\dagger$ . Linearity extends this to all  $X$ .  $\square$

The decisive structural fact on two qubits is that the PPT test is also sufficient.

**Theorem 3.2** (Peres [1]; Horodecki [2]). *For a two-qubit state  $\rho$ ,  $\rho \in \mathcal{S}$  if and only if  $\rho^\Gamma \geq 0$ .*

Two consequences are used repeatedly. First, every estimate that produces positivity of  $\rho^\Gamma$  produces separability, with no approximation. Second, a two-qubit partial transpose can carry at most one negative eigenvalue, and all its eigenvalues lie in  $[-\frac{1}{2}, 1]$  [3, 4]; the negativity

$$N(\rho) = \frac{\|\rho^\Gamma\|_1 - 1}{2} = \sum_{\mu < 0} |\mu|$$

therefore collapses to the single quantity  $[-\lambda_{\min}(\rho^\Gamma)]_+$ .

### 3.2 Norms, relative entropy, and the two functionals

For a Hermitian operator  $X$  we write  $\|X\|_1 = \text{Tr}|X|$  for the trace (Schatten 1-) norm,  $\|X\|_{\text{op}} = \lambda_{\max}|X|$  for the operator norm, and  $\|X\|_2^2 = \text{Tr} X^2$  for the Hilbert–Schmidt norm. For two states,  $\frac{1}{2}\|\rho - \sigma\|_1$  is the trace distance; all distances in this paper are stated in the trace norm  $\|\cdot\|_1$  and are therefore twice the trace distance, which is the convention in which the certified radii of Section 5 are sharp.

The relative entropy of two positive semidefinite operators is

$$D(\rho\|\sigma) = \text{Tr}[\rho(\ln \rho - \ln \sigma)],$$

with the natural logarithm, with the convention  $0 \ln 0 = 0$ , and with value  $+\infty$  when  $\text{supp } \rho \not\subseteq \text{supp } \sigma$ . The von Neumann entropy is  $S(\rho) = -\text{Tr}[\rho \ln \rho]$ , so that  $D(\rho\|\sigma) = -S(\rho) - \text{Tr}[\rho \ln \sigma]$ . We use four standard facts:  $D(\rho\|\sigma) \geq 0$  with equality if and only if  $\rho = \sigma$ ;  $D$  is jointly convex and, for fixed full-rank  $\rho$ , strictly convex in its second argument;  $D$  is monotone under completely positive trace-preserving maps,  $D(\Phi(\rho)\|\Phi(\sigma)) \leq D(\rho\|\sigma)$ , with equality conditions as reviewed in [14]; and  $\sigma \mapsto \text{Tr}[\rho \ln \sigma]$  is concave because the logarithm is operator concave.

**Definition 3.3** (relative entropy of entanglement).

$$E_R(\rho) = \min_{\sigma \in \mathcal{S}} D(\rho\|\sigma).$$

The minimum is attained because  $\mathcal{S}$  is compact and  $D(\rho\|\cdot)$  is lower semicontinuous [12, 5]. Any minimiser is called a *closest separable state* of  $\rho$ . Logs are natural, so  $E_R$  is in nats; one ebit is  $\ln 2$  nats.

### 3.3 The Werner family

For  $\lambda \in [0, 1]$  let

$$\rho_\lambda = \lambda P_- + \frac{1-\lambda}{3} P_{\text{sym}}. \quad (4)$$

This is the family invariant under  $U \otimes U$ , and in the computational basis

$$\rho_\lambda = \begin{pmatrix} \frac{1-\lambda}{3} & 0 & 0 & 0 \\ 0 & \frac{1+2\lambda}{6} & \frac{1-4\lambda}{6} & 0 \\ 0 & \frac{1-4\lambda}{6} & \frac{1+2\lambda}{6} & 0 \\ 0 & 0 & 0 & \frac{1-\lambda}{3} \end{pmatrix}, \quad \rho_\lambda^\Gamma = \begin{pmatrix} \frac{1-\lambda}{3} & 0 & 0 & \frac{1-4\lambda}{6} \\ 0 & \frac{1+2\lambda}{6} & 0 & 0 \\ 0 & 0 & \frac{1+2\lambda}{6} & 0 \\ \frac{1-4\lambda}{6} & 0 & 0 & \frac{1-\lambda}{3} \end{pmatrix}. \quad (5)$$

The second matrix is block diagonal, with  $|01\rangle, |10\rangle$  carrying the double eigenvalue  $\frac{1+2\lambda}{6}$  and  $|00\rangle, |11\rangle$  carrying  $\frac{1-\lambda}{3} \pm \frac{1-4\lambda}{6}$ . Hence

$$\text{spec}(\rho_\lambda^\Gamma) = \left\{ \frac{1-2\lambda}{2}, \underbrace{\frac{1+2\lambda}{6}, \frac{1+2\lambda}{6}, \frac{1+2\lambda}{6}}_{\text{threefold}} \right\}, \quad \eta(\lambda) = \lambda_{\min}(\rho_\lambda^\Gamma) = \min\left\{ \frac{1-2\lambda}{2}, \frac{1+2\lambda}{6} \right\}, \quad (6)$$

with eigenvector  $|\phi^+\rangle$  for the first eigenvalue and  $|\phi^-\rangle, |01\rangle, |10\rangle$  for the other three. The two branches cross at  $\lambda = \frac{1}{4}$ , so

$$\eta(\lambda) = \begin{cases} \frac{1+2\lambda}{6}, & 0 \leq \lambda \leq \frac{1}{4}, \\ \frac{1-2\lambda}{2}, & \frac{1}{4} \leq \lambda \leq \frac{1}{2}, \end{cases} \quad \eta(0) = \frac{1}{6}, \quad \eta\left(\frac{1}{4}\right) = \frac{1}{4}, \quad \eta\left(\frac{1}{2}\right) = 0.$$

By Theorem 3.2,  $\rho_\lambda$  is separable exactly when  $\lambda \leq \frac{1}{2}$ ;  $\rho_{1/4} = \frac{1}{4}I_4$  is the maximally mixed state,  $\rho_{1/2} = \sigma_*$  lies on the boundary, and  $\rho_1 = P_-$  is the singlet. The *singlet fraction* of a state is  $\lambda(\rho) = \text{Tr}[P_- \rho]$ , so  $\lambda(\rho_\lambda) = \lambda$ .

**Lemma 3.4** (maximal singlet fraction of a product vector). *For every unit product vector  $a \otimes b$ ,  $|\langle \psi^- | a \otimes b \rangle|^2 \leq \frac{1}{2}$ , with equality if and only if  $a \otimes b$  is, up to phase,  $|01\rangle - |10\rangle$  normalised or its local unitary images. Consequently  $\lambda(\sigma) \leq \frac{1}{2}$  for every  $\sigma \in \mathcal{S}$ .*

*Proof.* With  $a = (a_0, a_1)$  and  $b = (b_0, b_1)$ ,  $\langle \psi^- | a \otimes b \rangle = (a_1 b_0 - a_0 b_1)/\sqrt{2}$ , and the Cauchy–Schwarz–Binet identity gives  $|a_1 b_0 - a_0 b_1|^2 \leq (|a_0|^2 + |a_1|^2)(|b_0|^2 + |b_1|^2) = 1$ . Equality holds exactly when  $(a_0, a_1)$  and  $(b_0, b_1)$  are orthogonal in  $\mathbb{C}^2$ . The statement for  $\sigma \in \mathcal{S}$  follows because  $\lambda(\sigma)$  is a convex combination of product values.  $\square$

**Lemma 3.5** (the twirl). *Let  $\mathcal{T}(X) = \int (U \otimes U) X (U \otimes U)^\dagger dU$  be the Haar average over the diagonal action. Then  $\mathcal{T}$  is completely positive and trace preserving, it fixes every  $\rho_\lambda$ , it maps  $\mathcal{S}$  into  $\mathcal{S}$  and the PPT states into the PPT states, and*

$$\mathcal{T}(X) = \text{Tr}[P_- X] P_- + \frac{1 - \text{Tr}[P_- X]}{3} P_{\text{sym}} = \rho_G, \quad G = \text{Tr}[P_- X]. \quad (7)$$

*Proof.* The commutant of the diagonal action on  $\mathbb{C}^2 \otimes \mathbb{C}^2$  is spanned by  $P_-$  and  $P_{\text{sym}}$ , so  $\mathcal{T}(X) = \alpha P_- + \beta P_{\text{sym}}$  for scalars determined by  $\text{Tr}[P_- \mathcal{T}(X)] = \text{Tr}[P_- X]$  and  $\text{Tr}[\mathcal{T}(X)] = \text{Tr} X$ ; taking  $X$  a state gives  $\alpha = G$  and  $\beta = (1 - G)/3$ . Complete positivity and trace preservation are properties of a Haar average of unitary conjugations. Invariance of  $\mathcal{S}$  is because each  $U \otimes U$  conjugation preserves product projectors, and invariance of the PPT set is because Lemma 3.1 gives  $((U \otimes U) X (U \otimes U)^\dagger)^\Gamma = (U^* \otimes U) X^\Gamma (U^* \otimes U)^\dagger$ , which is positive whenever  $X^\Gamma$  is. Finally  $\rho_\lambda$  has the form (7) with  $G = \lambda$ , so it is fixed.  $\square$

### 3.4 The distance to the NPT set

**Definition 3.6** (nearest-NPT distance). For  $\lambda \in [0, \frac{1}{2}]$  set

$$d(\lambda) = \inf\{\|\sigma - \rho_\lambda\|_1 : \sigma \text{ a two-qubit state with } \lambda_{\min}(\sigma^\Gamma) < 0\}.$$

The set over which the infimum is taken is open in the state space, so the infimum need not be attained by an NPT state; it is attained in the closure, and in every case treated below we exhibit

an explicit NPT state or a limit of NPT states at the stated distance. A useful trivial competitor is the boundary state  $\sigma_* = \rho_{1/2}$ : since  $\sigma_* - \rho_\lambda = (\frac{1}{2} - \lambda)(P_- - \frac{1}{3}P_{\text{sym}})$  and  $\text{spec}(P_- - \frac{1}{3}P_{\text{sym}}) = \{1, -\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}\}$ ,

$$\|\sigma_* - \rho_\lambda\|_1 = 1 - 2\lambda, \quad (8)$$

and states arbitrarily close to  $\sigma_*$  on its NPT side give  $d(\lambda) \leq 1 - 2\lambda$  for every  $\lambda < \frac{1}{2}$ . The point of Section 6 is that this competitor is never optimal on  $[\lambda_0, \frac{1}{2})$  and that the true value is a different piecewise function.

## 4 A sharp operator-norm inequality for the partial transpose

This section proves (1). It is the only place in the paper where a genuinely new estimate is established from first principles; Sections 5–8 spend it.

### 4.1 The shape of the problem

The quantity to control is

$$\sup \left\{ \|X^\Gamma\|_{\text{op}} : X = X^\dagger, \text{Tr } X = 0, \|X\|_1 \leq 1 \right\} = \sup_{\|\psi\|=1} \sup_X |\text{Tr}[X(|\psi\rangle\langle\psi|)^\Gamma]|,$$

where the inner supremum runs over traceless Hermitian  $X$  with  $\|X\|_1 \leq 1$  and the outer one over unit vectors  $\psi$ . Two observations organise the rest.

First, the inner supremum is a duality pairing whose value is the *half-width* of the spectrum of  $(|\psi\rangle\langle\psi|)^\Gamma$ , not its operator norm. This is exactly where tracelessness enters: an unconstrained  $X$  would pair with  $(|\psi\rangle\langle\psi|)^\Gamma$  through its largest eigenvalue in absolute value, whereas a traceless  $X$  pairs with it only through the difference between its largest and smallest eigenvalues, because any multiple of the identity can be subtracted from  $(|\psi\rangle\langle\psi|)^\Gamma$  at no cost.

Second, by Lemma 3.1 the spectrum of  $(|\psi\rangle\langle\psi|)^\Gamma$  depends on  $\psi$  only through its two Schmidt weights, since local unitaries act on the partial transpose by conjugation. The outer supremum therefore reduces to a one-variable calculus problem on  $t \in [\frac{1}{2}, 1]$ .

### 4.2 Two lemmas

**Definition 4.1** (half-width). For  $A$  Hermitian set  $\text{hw}(A) = \frac{1}{2}(\lambda_{\max}(A) - \lambda_{\min}(A))$ .

**Lemma 4.2** (duality for traceless operators). *For every Hermitian  $A$  on a finite-dimensional space,*

$$\text{hw}(A) = \max \{ |\text{Tr}[XA]| : X = X^\dagger, \text{Tr } X = 0, \|X\|_1 \leq 1 \},$$

*and the maximum is attained at  $X = \frac{1}{2}(P_+ - P_-)$ , where  $P_\pm$  are eigenprojectors of  $A$  for  $\lambda_{\max}(A)$  and  $\lambda_{\min}(A)$ .*

*Proof.* Write  $X = \sum_i \mu_i |v_i\rangle\langle v_i|$  with  $\sum_i |\mu_i| = \|X\|_1 \leq 1$  and  $\sum_i \mu_i = \text{Tr } X = 0$ . Since  $\lambda_{\min}(A) \leq \langle v_i | A v_i \rangle \leq \lambda_{\max}(A)$ ,

$$\text{Tr}[XA] = \sum_i \mu_i \langle v_i | A v_i \rangle \leq \left( \sum_{\mu_i > 0} \mu_i \right) \lambda_{\max}(A) + \left( \sum_{\mu_i < 0} \mu_i \right) \lambda_{\min}(A) = \frac{s}{2} (\lambda_{\max}(A) - \lambda_{\min}(A)) \leq \text{hw}(A),$$

where  $s = \sum_i |\mu_i| \leq 1$  and  $\sum_{\mu_i > 0} \mu_i = -\sum_{\mu_i < 0} \mu_i = s/2$ . The same estimate applied to  $-X$  gives the absolute value. For the reverse inequality,  $X = \frac{1}{2}(P_+ - P_-)$  is traceless, has  $\|X\|_1 = 1$ , and gives  $\text{Tr}[XA] = \text{hw}(A)$ .  $\square$

**Lemma 4.3** (spectrum of a partially transposed pure state). *Let  $\psi$  be a unit vector in  $\mathbb{C}^2 \otimes \mathbb{C}^2$  with Schmidt weights  $t$  and  $1 - t$ ,  $t \in [\frac{1}{2}, 1]$ . Then*

$$\text{spec}((|\psi\rangle\langle\psi|)^\Gamma) = \{t, 1 - t, +\sqrt{t(1 - t)}, -\sqrt{t(1 - t)}\},$$

and consequently

$$\text{hw}((|\psi\rangle\langle\psi|)^\Gamma) = \frac{t + \sqrt{t(1 - t)}}{2}.$$

*Proof.* Local unitaries bring  $\psi$  to  $\sqrt{t}|00\rangle + \sqrt{1 - t}|11\rangle$ , and by Lemma 3.1 this conjugates  $(|\psi\rangle\langle\psi|)^\Gamma$  by a local unitary, which does not change the spectrum. For the standard form,

$$(|\psi\rangle\langle\psi|)^\Gamma = t|00\rangle\langle 00| + (1 - t)|11\rangle\langle 11| + \sqrt{t(1 - t)}(|01, 10\rangle\langle 01, 10| + |10, 01\rangle\langle 10, 01|),$$

which is block diagonal:  $|01\rangle, |10\rangle$  carry  $\pm\sqrt{t(1 - t)}$  and  $|00\rangle, |11\rangle$  carry  $t$  and  $1 - t$ . For the half-width, note that  $t \geq 1 - t$  and  $\sqrt{t(1 - t)} \leq t$  because  $t(1 - t) \leq t^2$  is equivalent to  $1 - t \leq t$ ; hence  $\lambda_{\max} = t$  and  $\lambda_{\min} = -\sqrt{t(1 - t)}$ .  $\square$

**Lemma 4.4** (the one-variable maximum). *The function  $W(t) = \frac{1}{2}(t + \sqrt{t(1 - t)})$  on  $[\frac{1}{2}, 1]$  attains its maximum at  $t_* = \frac{2+\sqrt{2}}{4} = 0.85355339\dots$ , with*

$$W(t_*) = K = \frac{1 + \sqrt{2}}{4} = 0.6035533905932737\dots, \quad W(\tfrac{1}{2}) = W(1) = \tfrac{1}{2} < K. \quad (9)$$

*The other critical point of  $W$  on  $(\frac{1}{2}, 1)$  is  $\frac{2-\sqrt{2}}{4} < \frac{1}{2}$ , which lies outside the interval.*

*Proof.*  $W'(t) = \frac{1}{2}(1 + \frac{1-2t}{2\sqrt{t(1-t)}})$ , so  $W'(t) = 0$  requires  $2\sqrt{t(1 - t)} = 2t - 1$ , which is consistent only for  $t \geq \frac{1}{2}$  and, on squaring, gives  $4t(1 - t) = (2t - 1)^2$ , that is  $8t^2 - 8t + 1 = 0$ , with roots  $\frac{2 \pm \sqrt{2}}{4}$ . Only  $t_* = \frac{2+\sqrt{2}}{4}$  lies in  $[\frac{1}{2}, 1]$ . At that point  $\sqrt{t_*(1 - t_*)} = (2t_* - 1)/2 = \frac{\sqrt{2}}{4}$ , so  $W(t_*) = \frac{1}{2}(\frac{2+\sqrt{2}}{4} + \frac{\sqrt{2}}{4}) = \frac{1+\sqrt{2}}{4}$ . The endpoint values are  $W(\frac{1}{2}) = \frac{1}{2}(\frac{1}{2} + \frac{1}{2}) = \frac{1}{2}$  and  $W(1) = \frac{1}{2}$ , both strictly smaller, so the interior critical point is the global maximum.  $\square$

### 4.3 The theorem

**Theorem 4.5** (sharp traceless partial-transpose bound). *Let  $X$  be a traceless Hermitian operator on  $\mathbb{C}^2 \otimes \mathbb{C}^2$ . Then*

$$\|X^\Gamma\|_{\text{op}} \leq K \|X\|_1, \quad K = \frac{1 + \sqrt{2}}{4}, \quad (10)$$

and equivalently

$$\lambda_{\min}(X^\Gamma) \geq -K \|X\|_1. \quad (11)$$

*Both inequalities are sharp, and the constant cannot be lowered.*

*Proof.* Let  $\psi$  be a unit vector with Schmidt weight  $t \in [\frac{1}{2}, 1]$ . Since  $\text{Tr } X = 0$ ,

$$|\langle\psi|X^\Gamma\psi\rangle| = |\text{Tr}[X(|\psi\rangle\langle\psi|)^\Gamma]| \leq \text{hw}((|\psi\rangle\langle\psi|)^\Gamma) \|X\|_1 = \frac{t + \sqrt{t(1 - t)}}{2} \|X\|_1 \leq K \|X\|_1,$$

where the first inequality is Lemma 4.2, the identity is Lemma 4.3 and the last step is Lemma 4.4. Taking the supremum over  $\psi$  gives  $\|X^\Gamma\|_{\text{op}} \leq K \|X\|_1$ , which is (10). For (11), note that  $\lambda_{\min}(X^\Gamma) = \min_{\|\psi\|=1} \langle\psi|X^\Gamma\psi\rangle \geq -\|X^\Gamma\|_{\text{op}}$ .

Sharpness is Proposition 4.6 below.  $\square$

**Proposition 4.6** (explicit extremisers). *Let*

$$X_* = \frac{1}{2}(|\phi^-\rangle\langle\phi^-| - |01\rangle\langle 01|), \quad X_2 = \frac{1}{2}(|00\rangle\langle 00| - |\psi^-\rangle\langle\psi^-|), \quad (12)$$

with  $|\phi^-\rangle$  and  $|\psi^-\rangle$  as in (2). Both are Hermitian and traceless with trace norm 1. Their partial transposes have spectra

$$\text{spec}(X_*^\Gamma) = \left\{-K, \frac{\sqrt{2}-1}{4}, \frac{1}{4}, \frac{1}{4}\right\}, \quad \lambda_{\min}(X_*^\Gamma) = -K, \quad (13)$$

$$\text{spec}(X_2^\Gamma) = \left\{-\frac{1}{4}, -\frac{1}{4}, -\frac{\sqrt{2}-1}{4}, K\right\}, \quad \lambda_{\max}(X_2^\Gamma) = +K, \quad (14)$$

so that  $\|X_*^\Gamma\|_{\text{op}} = \|X_2^\Gamma\|_{\text{op}} = K = K\|X_*\|_1 = K\|X_2\|_1$ : (10) is an equality for both, attained at the negative end for  $X_*$  and at the positive end for  $X_2$ . Moreover (11) is an equality for  $X_*$ . The eigenvector of  $X_*^\Gamma$  for  $-K$  is, up to phase,

$$\psi_0 = \sqrt{t_*}|01\rangle + \sqrt{1-t_*}|10\rangle = \frac{|01\rangle + (\sqrt{2}-1)|10\rangle}{\sqrt{4-2\sqrt{2}}}, \quad (15)$$

whose Schmidt weights are exactly  $t_*$  and  $1-t_*$ , the maximiser of Lemma 4.4;  $\psi_0$  lies in the span of  $|01\rangle, |10\rangle$ , which is the  $\eta(\lambda)$ -eigenspace of  $\rho_\lambda^\Gamma$  for every  $\lambda \leq \frac{1}{4}$  (compare (6)).

*Proof.*  $|\phi^-\rangle\langle\phi^-|$  and  $|01\rangle\langle 01|$  are orthogonal projectors, so  $X_*$  has eigenvalues  $\frac{1}{2}, -\frac{1}{2}, 0, 0$ , whence  $\|X_*\|_1 = 1$  and  $\text{Tr } X_* = 0$ ; likewise  $|00\rangle\langle 00|$  and  $|\psi^-\rangle\langle\psi^-|$  are orthogonal, so  $\|X_2\|_1 = 1$  and  $\text{Tr } X_2 = 0$ . In the computational basis,

$$X_* = \begin{pmatrix} \frac{1}{4} & 0 & 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{4} & 0 & 0 & \frac{1}{4} \end{pmatrix}, \quad X_*^\Gamma = \begin{pmatrix} \frac{1}{4} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{pmatrix},$$

because  $\Gamma$  fixes  $|00\rangle\langle 00|$ ,  $|11\rangle\langle 11|$  and  $|01\rangle\langle 01|$  and interchanges  $|00, 11\rangle\langle 00, 11|$  with  $|10, 01\rangle\langle 10, 01|$ . The  $|00\rangle$  and  $|11\rangle$  entries give the double eigenvalue  $\frac{1}{4}$ , and the  $|01\rangle, |10\rangle$  block  $\begin{pmatrix} -1/2 & -1/4 \\ -1/4 & 0 \end{pmatrix}$  has characteristic polynomial  $\mu^2 + \frac{1}{2}\mu - \frac{1}{16}$ , hence  $\mu = \frac{1}{4}(-1 \pm \sqrt{2})$ , that is  $-K$  and  $(\sqrt{2}-1)/4$ . This is (13). For the eigenvector,  $-\frac{1}{4}v_{01} = \mu v_{10}$  with  $\mu = -K$  gives  $v_{10} = v_{01}/(4K) = (\sqrt{2}-1)v_{01}$ , so  $\psi_0 \propto |01\rangle + (\sqrt{2}-1)|10\rangle$ ; normalising,  $1 + (\sqrt{2}-1)^2 = 4 - 2\sqrt{2}$  and  $1/(4 - 2\sqrt{2}) = (2 + \sqrt{2})/4 = t_*$ ,  $(\sqrt{2}-1)^2/(4 - 2\sqrt{2}) = (2 - \sqrt{2})/4 = 1 - t_*$ , which are the squared coefficients and also the Schmidt weights because  $\psi_0$  is supported on  $|01\rangle, |10\rangle$ .

For  $X_2$ , write  $|\psi^-\rangle\langle\psi^-| = \frac{1}{2}(|01\rangle\langle 01| - |01\rangle\langle 10| - |10\rangle\langle 01| + |10\rangle\langle 10|)$ ; then  $\Gamma$  fixes  $|01\rangle\langle 01|$  and  $|10\rangle\langle 10|$  and sends  $|01\rangle\langle 10|$  to  $|11\rangle\langle 00|$  and  $|10\rangle\langle 01|$  to  $|00\rangle\langle 11|$ , so

$$X_2 = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{4} & 0 \\ 0 & \frac{1}{4} & -\frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad X_2^\Gamma = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{4} \\ 0 & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & -\frac{1}{4} & 0 \\ \frac{1}{4} & 0 & 0 & 0 \end{pmatrix}.$$

The  $|01\rangle, |10\rangle$  entries give the double eigenvalue  $-\frac{1}{4}$ , and the  $|00\rangle, |11\rangle$  block  $\begin{pmatrix} 1/2 & 1/4 \\ 1/4 & 0 \end{pmatrix}$  has the same characteristic polynomial as before with the opposite sign of the off-diagonal product, hence roots  $\frac{1}{4}(1 \pm \sqrt{2}) = K, -(\sqrt{2}-1)/4$ . This is (14), and  $\|X_2^\Gamma\|_{\text{op}} = K$ .  $\square$

*Remark 4.7* (where equality can occur). Tracing the proof, equality in (10) requires both lemmas to be tight simultaneously: the vector  $\psi$  at which  $\|X^\Gamma\|_{\text{op}}$  is attained must have Schmidt weight  $t_*$  (or, by the symmetry  $t \leftrightarrow 1-t$  of  $W$ , must realise the same half-width), and  $X$  must be of the form  $\frac{1}{2}(P_+ - P_-)$  with  $P_\pm$  supported in the extremal eigenspaces of  $(|\psi\rangle\langle\psi|)^\Gamma$ . Both explicit extremisers satisfy this. A complete classification of the equality case is not attempted here and is not needed anywhere below.

*Remark 4.8* (tracelessness is essential). The restriction  $\text{Tr } X = 0$  is not cosmetic. For  $X = |00\rangle\langle 00|$  one has  $\|X\|_1 = 1$  and  $\|X^\Gamma\|_{\text{op}} = 1 > K$ , so (10) fails without it: the centring step in Lemma 4.2 is what produces  $K < 1$ , and it is legal only because  $\text{Tr } X = 0$ . Every application in this paper uses  $X = \rho - \sigma$  for two states, which is traceless by construction. For traceless  $X$  the constant 1 is also available directly, since  $\text{hw}((|\psi\rangle\langle\psi|)^\Gamma) \leq 1$  for every  $\psi$ ; Lemma 4.4 is the statement that the true value is  $K$ , a factor  $1/K = 1.6568542\dots$  better.

*Remark 4.9* (the constant is dimension free). Nothing in the proof of Theorem 4.5 uses  $d = 2$  except Lemma 4.3. Appendix A carries out the  $d \otimes d$  version: for Schmidt weights  $t_1 \geq \dots \geq t_d$  the spectrum of  $(|\psi\rangle\langle\psi|)^\Gamma$  is  $\{t_i\} \cup \{\pm\sqrt{t_i t_j} : i < j\}$ , the half-width is  $\frac{1}{2}(t_1 + \sqrt{t_1 t_2})$ , and the same constant  $K$  results, attained in every dimension by embedding  $X_*$ . What does not extend is the use of the inequality; see Remark A.3.

## 5 A quantitative Peres–Horodecki criterion

The inequality (1) turns the sign test of Theorem 3.2 into a stability statement. The input is any PPT reference state and the size of its partial-transpose gap; the output is a guarantee about every state within a controlled trace-norm distance of it.

**Theorem 5.1** (continuity law). *Let  $\sigma$  be a two-qubit PPT state and let  $g = \lambda_{\min}(\sigma^\Gamma) \geq 0$ . Then for every two-qubit state  $\rho$ ,*

$$\lambda_{\min}(\rho^\Gamma) \geq g - K \|\rho - \sigma\|_1, \quad K = \frac{1 + \sqrt{2}}{4}, \quad (16)$$

*and  $\rho$  is separable whenever  $\|\rho - \sigma\|_1 < g/K$ . The constant  $K$  cannot be replaced by any smaller number, not even if  $\sigma$  is restricted to the separable Werner states and  $\rho$  to the ray  $\sigma + sX_*$ ,  $s \geq 0$ .*

*Proof.*  $X = \rho - \sigma$  is Hermitian and traceless, so Weyl’s inequality and (11) give

$$\lambda_{\min}(\rho^\Gamma) = \lambda_{\min}(\sigma^\Gamma + X^\Gamma) \geq \lambda_{\min}(\sigma^\Gamma) + \lambda_{\min}(X^\Gamma) \geq g - K\|X\|_1.$$

If the right-hand side is positive then  $\rho^\Gamma > 0$ , so  $\rho$  is PPT and, by Theorem 3.2, separable. For the optimality statement, take  $\sigma = \rho_\lambda$  with  $\lambda_c \leq \lambda \leq \frac{1}{4}$  and  $\rho = \rho_\lambda + sX_*$  with  $s = \eta(\lambda)/K$ . The vector  $\psi_0$  of (15) lies in  $\text{span}\{|01\rangle, |10\rangle\}$ , on which  $\rho_\lambda^\Gamma$  acts as  $\eta(\lambda)$  times the identity by (6), and it is an eigenvector of  $X_*^\Gamma$  for  $-K$  by Proposition 4.6. Weyl’s inequality is therefore an equality on  $\psi_0$  and  $\lambda_{\min}(\rho^\Gamma) = g - Ks$  exactly; the same computation, including the positivity of  $\rho$ , is Proposition 6.5.  $\square$

**Corollary 5.2** (negativity bound). *Under the hypotheses of Theorem 5.1,*

$$N(\rho) \leq [K\|\rho - \sigma\|_1 - g]_+. \quad (17)$$

*Proof.* A two-qubit partial transpose has at most one negative eigenvalue and all its eigenvalues lie in  $[-\frac{1}{2}, 1]$  [3, 4]; hence  $N(\rho) = \sum_{\mu \in \text{spec}(\rho^\Gamma), \mu < 0} |\mu| = [-\lambda_{\min}(\rho^\Gamma)]_+$ , and (16) gives  $[-\lambda_{\min}(\rho^\Gamma)]_+ \leq [-g + K\|\rho - \sigma\|_1]_+$ .  $\square$

**Corollary 5.3** (certified separable balls along the Werner line). *Let  $0 \leq \lambda \leq \frac{1}{2}$  and  $\eta(\lambda) = \min\{\frac{1-2\lambda}{2}, \frac{1+2\lambda}{6}\}$ . Then every two-qubit state  $\rho$  with*

$$\|\rho - \rho_\lambda\|_1 < \frac{\eta(\lambda)}{K} = 4(\sqrt{2}-1)\eta(\lambda) = 1.6568542494923801\dots \eta(\lambda) \quad (18)$$

*is separable. The same conclusion holds under the weaker hypothesis  $\|\rho - \rho_\lambda\|_1 < \eta(\lambda)$ . At  $\lambda = \frac{1}{4}$ , where  $\rho_{1/4} = \frac{1}{4}I_4$  and  $\eta = \frac{1}{4}$ , the certified radius is  $\sqrt{2}-1 = 0.4142135623730951\dots$  in trace norm, that is  $0.2071067811865476\dots$  in trace distance.*

*Proof.* Take  $\sigma = \rho_\lambda$ , which is PPT for  $\lambda \leq \frac{1}{2}$  with  $g = \eta(\lambda)$  by (6), and apply Theorem 5.1. For the weaker statement, note that  $\text{hw}(|\psi\rangle\langle\psi|^\Gamma) \leq 1$  for every unit  $\psi$  by Lemma 4.3, so Lemma 4.2 gives  $\|X^\Gamma\|_{\text{op}} \leq \|X\|_1$  for every traceless Hermitian  $X$ ; the same Weyl argument then yields  $\lambda_{\min}(\rho^\Gamma) \geq \eta(\lambda) - \|\rho - \rho_\lambda\|_1$ . The numerical values are  $\eta(\frac{1}{4})/K = \frac{1}{4} \cdot 4(\sqrt{2}-1) = \sqrt{2}-1$  and  $1/K = 4(\sqrt{2}-1)$ .  $\square$

*Remark 5.4* (the two radii and where each is sharp). The gain of the sharp constant over the elementary Weyl argument is the factor  $1/K = 1.6568542494923801$ , uniformly in  $\lambda$ . The larger radius is attained, not merely certified: on  $[\lambda_c, \frac{1}{4}]$  the state  $\rho_\lambda + (\eta(\lambda)/K)X_*$  is a genuine state on the PPT boundary at distance exactly  $\eta(\lambda)/K$  (Proposition 6.6), so no larger radius can be proved by this route at those  $\lambda$ . We make no optimality claim at  $\lambda = \frac{1}{4}$  in the absolute sense: the largest ball of separable states centred at the maximally mixed state is a different question, treated in other norms and in arbitrary dimension by Gurvits and Barnum [10], and converting between their radius and ours is not part of this paper.

*Remark 5.5* (a retraction: the partial transpose is not a trace-norm isometry). A larger ball of radius  $2\eta(\lambda)$  would follow from the estimate  $\|X^\Gamma\|_{\text{op}} \leq \frac{1}{2}\|X\|_1$ , which in turn would be a consequence of  $\Gamma$  preserving the trace norm. It does not. The traceless Hermitian matrix

$$X = |00\rangle\langle 00| \cdot (-1) + |11\rangle\langle 11| - |01\rangle\langle 10| - |10\rangle\langle 01| = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (19)$$

has  $\text{spec}(X) = \{-1, -1, 1, 1\}$ , hence  $\|X\|_1 = 4$ , while  $X^\Gamma = -|00\rangle\langle 00| + |11\rangle\langle 11| - |00\rangle\langle 11| - |11\rangle\langle 00|$  has  $\text{spec}(X^\Gamma) = \{-\sqrt{2}, 0, 0, \sqrt{2}\}$  and  $\|X^\Gamma\|_1 = 2\sqrt{2}$ . The ratio is  $1/\sqrt{2} = 0.70710678\dots$ , so  $\Gamma$  contracts the trace norm on this operator.

The radius- $2\eta(\lambda)$  ball is correspondingly false. At  $\lambda = 0.24$ , where  $\eta = 0.2466666667$  and  $2\eta = 0.4933333333$ , the explicit state  $\sigma$  written out in Appendix B satisfies

$$\|\sigma - \rho_{0.24}\|_1 = 0.444034691994310 < 2\eta(0.24), \quad \lambda_{\min}(\sigma^\Gamma) = -9.706096 \times 10^{-4} < 0,$$

with  $\lambda_{\min}(\sigma) = 0.0276367338 > 0$  and  $\text{Tr } \sigma = 1$ ; it is therefore a genuinely NPT state inside the claimed ball. The corrected ball of Corollary 5.3 is not threatened: its radius at  $\lambda = 0.24$  is  $\eta/K = 0.4086907148$ , and the counterexample sits outside it by  $0.0353439772$ . The refutation is nonetheless delicate—the negative partial-transpose eigenvalue is of order  $10^{-3}$ —which is precisely why the radius has to be certified by an operator inequality rather than by sampling.

## 6 The distance from a separable Werner state to the nearest NPT state

### 6.1 Two reductions

The set of NPT states is the complement of a convex set, so it has no supporting hyperplane of its own and no dual norm formula. Two elementary reductions organise everything that follows: one turns boundary states into upper bounds, the other turns entanglement witnesses into lower bounds.

**Lemma 6.1** (boundary states give upper bounds). *Let  $\rho$  be a two-qubit state and let  $\sigma$  be a two-qubit state with  $\lambda_{\min}(\sigma^\Gamma) = 0$ . Then*

$$\inf\{\|\tau - \rho\|_1 : \tau \text{ a state, } \tau \text{ NPT}\} \leq \|\sigma - \rho\|_1.$$

*Proof.* Fix any NPT state  $\tau_0$  and set  $\sigma_t = (1-t)\sigma + t\tau_0$  for  $t \in (0,1)$ . Then  $\sigma_t$  is a state and  $\lambda_{\min}(\sigma_t^\Gamma) \leq (1-t)\lambda_{\min}(\sigma^\Gamma) + t\lambda_{\min}(\tau_0^\Gamma) < 0$ , so  $\sigma_t$  is NPT, while  $\|\sigma_t - \rho\|_1 \leq (1-t)\|\sigma - \rho\|_1 + t\|\tau_0 - \rho\|_1$  tends to  $\|\sigma - \rho\|_1$  as  $t \downarrow 0$ .  $\square$

**Lemma 6.2** (witnesses give lower bounds). *Let  $w$  be a unit vector and write  $A_w = (|w\rangle\langle w|)^\Gamma$  and*

$$D_w(\lambda) = \frac{\langle w | \rho_\lambda^\Gamma | w \rangle}{\text{hw}(A_w)}. \quad (20)$$

*Every two-qubit state  $\sigma$  with  $\langle w | \sigma^\Gamma | w \rangle \leq 0$  satisfies  $\|\sigma - \rho_\lambda\|_1 \geq D_w(\lambda)$ , and therefore*

$$d(\lambda) \geq \inf_{\|w\|=1} D_w(\lambda). \quad (21)$$

*Proof.* Put  $Z = \sigma - \rho_\lambda$ , which is Hermitian and traceless. Then  $\langle w | \sigma^\Gamma | w \rangle = \text{Tr}[ZA_w] + \langle w | \rho_\lambda^\Gamma | w \rangle$ , so the hypothesis gives  $\text{Tr}[ZA_w] \leq -\langle w | \rho_\lambda^\Gamma | w \rangle \leq 0$ , and Lemma 4.2 bounds  $\text{Tr}[ZA_w] \geq -\text{hw}(A_w)\|Z\|_1$ . Combining,  $\text{hw}(A_w)\|Z\|_1 \geq \langle w | \rho_\lambda^\Gamma | w \rangle$ . Every NPT state has  $\langle w | \sigma^\Gamma | w \rangle < 0$  for a unit eigenvector  $w$  of  $\sigma^\Gamma$  belonging to a negative eigenvalue, so (21) follows by taking the infimum over such  $w$ .  $\square$

The relaxation (21) is strict whenever the closest point of the halfspace  $\{\text{Tr}[YA_w] \leq 0\}$  to  $\rho_\lambda$  is not a positive operator; that is exactly what happens below  $\lambda_c$  (Table 1).

### 6.2 The Weyl lower bound

**Proposition 6.3** (lower bound). *For  $0 \leq \lambda < \frac{1}{2}$ ,*

$$d(\lambda) \geq \frac{\eta(\lambda)}{K} = \begin{cases} \frac{2(\sqrt{2}-1)}{3} (1+2\lambda), & 0 \leq \lambda \leq \frac{1}{4}, \\ 2(\sqrt{2}-1) (1-2\lambda), & \frac{1}{4} \leq \lambda < \frac{1}{2}. \end{cases} \quad (22)$$

*Proof.* Let  $\sigma$  be any NPT state. Applying Theorem 5.1 with the PPT reference  $\rho_\lambda$ , whose gap is  $g = \eta(\lambda)$  by (6), gives  $0 > \lambda_{\min}(\sigma^\Gamma) \geq \eta(\lambda) - K\|\sigma - \rho_\lambda\|_1$ , hence  $\|\sigma - \rho_\lambda\|_1 > \eta(\lambda)/K$ . Taking the infimum over  $\sigma$  gives (22); the two branches are  $\eta(\lambda) = (1+2\lambda)/6$  and  $\eta(\lambda) = (1-2\lambda)/2$  divided by  $K = (1+\sqrt{2})/4$ , using  $1/K = 4(\sqrt{2}-1)$ .  $\square$

*Remark 6.4* (the same bound from one witness). For  $\lambda \leq \frac{1}{4}$  the bound (22) is also the value of (20) at the single witness  $w = \psi_0$  of (15):  $\psi_0$  lies in the span of  $|01\rangle, |10\rangle$ , on which  $\rho_\lambda^\Gamma$  acts as  $\eta(\lambda)$  times the identity, so  $\langle \psi_0 | \rho_\lambda^\Gamma | \psi_0 \rangle = \eta(\lambda)$ , while  $\text{hw}((|\psi_0\rangle\langle\psi_0|)^\Gamma) = K$  because  $\psi_0$  has Schmidt weight  $t_*$ . The relaxation (21) therefore contains Proposition 6.3 as a special case, and it is exactly this witness that the numerical minimisation of (20) returns for every  $\lambda \leq \frac{1}{4}$  (Section 9).

### 6.3 Exactness on $[\lambda_c, \frac{1}{4}]$

The extremiser  $X_*$  of Proposition 4.6 generates a ray out of  $\rho_\lambda$  along which the Weyl bound is saturated.

**Proposition 6.5** (the saturating ray). *For  $0 \leq \lambda \leq \frac{1}{4}$  set  $s(\lambda) = \eta(\lambda)/K$  and*

$$\sigma(\lambda) = \rho_\lambda + s(\lambda)X_*. \quad (23)$$

*Then  $\text{Tr } \sigma(\lambda) = 1$ ,  $\lambda_{\min}(\sigma(\lambda)^\Gamma) = 0$ , and  $\|\sigma(\lambda) - \rho_\lambda\|_1 = s(\lambda)$ . Moreover  $\sigma(\lambda) \geq 0$  if and only if  $\lambda \geq \lambda_c$ , where*

$$\lambda_c = \frac{2 - \sqrt{2}}{2(1 + \sqrt{2})} = \frac{3\sqrt{2} - 4}{2} = 0.1213203435596426\dots \quad (24)$$

*Proof.* Write  $a = \frac{1-\lambda}{3}$ ,  $b = \frac{1+2\lambda}{6}$  and  $e = \frac{1-4\lambda}{6}$ , so that  $\rho_\lambda = a(|00\rangle\langle 00| + |11\rangle\langle 11|) + b(|01\rangle\langle 01| + |10\rangle\langle 10|) + e(|01\rangle\langle 10| + |10\rangle\langle 01|)$  and  $\rho_\lambda^\Gamma = a(|00\rangle\langle 00| + |11\rangle\langle 11|) + b(|01\rangle\langle 01| + |10\rangle\langle 10|) + e(|00\rangle\langle 11| + |11\rangle\langle 00|)$  by (3). Using (12) and the matrices computed in Proposition 4.6,

$$\sigma(\lambda) = \begin{pmatrix} a + \frac{s}{4} & 0 & 0 & -\frac{s}{4} \\ 0 & b - \frac{s}{2} & e & 0 \\ 0 & e & b & 0 \\ -\frac{s}{4} & 0 & 0 & a + \frac{s}{4} \end{pmatrix}, \quad \sigma(\lambda)^\Gamma = \begin{pmatrix} a + \frac{s}{4} & 0 & 0 & e \\ 0 & b - \frac{s}{2} & -\frac{s}{4} & 0 \\ 0 & -\frac{s}{4} & b & 0 \\ e & 0 & 0 & a + \frac{s}{4} \end{pmatrix}.$$

The trace is  $2a + 2b = 1$ . For  $\sigma(\lambda)^\Gamma$ , the  $|00\rangle, |11\rangle$  block has eigenvalues  $a + \frac{s}{4} \pm e$ , both non-negative because  $a - e = b > 0$ , and the  $|01\rangle, |10\rangle$  block has eigenvalues  $b - \frac{s}{4} \pm \frac{s}{4}\sqrt{2} = b - sK$ , whose smaller member is  $b - \eta(\lambda) = 0$  for  $\lambda \leq \frac{1}{4}$  because  $s = b/K$  and  $\eta(\lambda) = b$ ; this is the claimed boundary eigenvalue, and it shows that Weyl's inequality is an equality on this ray. For  $\sigma(\lambda)$  itself, the  $|00\rangle, |11\rangle$  block has eigenvalues  $a$  and  $a + \frac{s}{2}$ , both positive, and the  $|01\rangle, |10\rangle$  block  $\begin{pmatrix} b-s/2 & e \\ e & b \end{pmatrix}$  is positive semidefinite exactly when  $b - \frac{s}{2} \geq 0$  and its determinant  $b(b - \frac{s}{2}) - e^2$  is non-negative. The first condition is  $b(1 - \frac{1}{2K}) = b(3 - 2\sqrt{2}) > 0$ ; the second, using  $s = b/K$ , is  $b^2(3 - 2\sqrt{2}) \geq e^2$ , that is  $e \leq b(\sqrt{2} - 1)$  because  $\sqrt{3 - 2\sqrt{2}} = \sqrt{2} - 1$  and  $e, b \geq 0$  on  $[0, \frac{1}{4}]$ . Substituting the values of  $b$  and  $e$  gives  $1 - 4\lambda \leq (\sqrt{2} - 1)(1 + 2\lambda)$ , equivalently  $(2 - \sqrt{2}) \leq 2(1 + \sqrt{2})\lambda$ , which is  $\lambda \geq \lambda_c$  as in (24). Finally  $\|\sigma(\lambda) - \rho_\lambda\|_1 = s\|X_*\|_1 = s$  because  $\sigma(\lambda) - \rho_\lambda = sX_*$  and  $\|X_*\|_1 = 1$ .  $\square$

**Theorem 6.6** (exact value on  $[\lambda_c, \frac{1}{4}]$ ). *For every  $\lambda \in [\lambda_c, \frac{1}{4}]$ ,*

$$d(\lambda) = \frac{\eta(\lambda)}{K} = \frac{2(\sqrt{2} - 1)}{3} (1 + 2\lambda), \quad (25)$$

*with the endpoint values*

$$d(\lambda_c) = 6 - 4\sqrt{2} = 0.3431457505076198\dots, \quad d\left(\frac{1}{4}\right) = \sqrt{2} - 1 = 0.4142135623730951\dots$$

*The infimum is attained for  $\lambda_c < \lambda \leq \frac{1}{4}$  and is not attained at  $\lambda = \lambda_c$ .*

*Proof.* The lower bound is (22). For the upper bound, Proposition 6.5 supplies a state  $\sigma(\lambda)$  with  $\lambda_{\min}(\sigma(\lambda)^\Gamma) = 0$  at distance exactly  $\eta(\lambda)/K$ , so Lemma 6.1 gives  $d(\lambda) \leq \eta(\lambda)/K$ . At  $\lambda = \lambda_c$  the ray is the only feasible point of its own family, and the infimum is approached by states  $\rho_\lambda + sX_* + t\tau_0$  with  $t \downarrow 0$  rather than attained, because  $\sigma(\lambda_c)$  itself is singular. The endpoint values are  $\eta(\frac{1}{4})/K = \frac{1}{4} \cdot 4(\sqrt{2} - 1) = \sqrt{2} - 1$  and, since  $1 + 2\lambda_c = 3\sqrt{2} - 3 = 3(\sqrt{2} - 1)$  gives  $\eta(\lambda_c) = (\sqrt{2} - 1)/2$ ,  $\eta(\lambda_c)/K = \frac{\sqrt{2}-1}{2} \cdot \frac{4}{1+\sqrt{2}} = \frac{2(\sqrt{2}-1)}{1+\sqrt{2}} = 2(\sqrt{2} - 1)^2 = 6 - 4\sqrt{2}$ .  $\square$

The last identity deserves a line of its own: at  $\lambda = \lambda_c$  the two branches of the upper-bound construction meet, since  $d_{\text{iso}}(\lambda_c) = 6 - 4\sqrt{2}$  as well (Proposition 6.12 below), and the ray, which is infeasible on one side of  $\lambda_c$ , is exactly matched by the corner family on the other.

**Proposition 6.7** (the Weyl bound is strict above  $\frac{1}{4}$ ). *For  $\frac{1}{4} < \lambda < \frac{1}{2}$ ,  $d(\lambda) > \eta(\lambda)/K$ .*

*Proof.* Suppose  $d(\lambda) = \eta(\lambda)/K$ . The set of states is compact, so some state  $\sigma$  attains  $\|\sigma - \rho_\lambda\|_1 = \eta(\lambda)/K$  with  $\lambda_{\min}(\sigma^\Gamma) \leq 0$ ; but (16) with  $g = \eta(\lambda)$  gives  $\lambda_{\min}(\sigma^\Gamma) \geq \eta(\lambda) - K \cdot \eta(\lambda)/K = 0$ , so  $\lambda_{\min}(\sigma^\Gamma) = 0$  and both inequalities used are equalities. Equality in Weyl's inequality means that some unit  $\psi$  is simultaneously an eigenvector of  $\rho_\lambda^\Gamma$  for  $\eta(\lambda)$ , of  $X^\Gamma$  for  $\lambda_{\min}(X^\Gamma)$  and of  $\sigma^\Gamma$  for 0, where  $X = \sigma - \rho_\lambda$ . Equality in (11) means, by Remark 4.7, that the vector at which  $\lambda_{\min}(X^\Gamma)$  is attained has Schmidt weights  $(t_*, 1 - t_*)$ ; that vector is  $\psi$ . For  $\lambda > \frac{1}{4}$  the  $\eta(\lambda)$ -eigenspace of  $\rho_\lambda^\Gamma$  is  $\mathbb{C}|\phi^+\rangle$  by (6), because  $\frac{1-2\lambda}{2} < \frac{1+2\lambda}{6}$ , and every vector in it has Schmidt weights  $(\frac{1}{2}, \frac{1}{2}) \neq (t_*, 1 - t_*)$ . Contradiction.  $\square$

*Remark 6.8* (why the ray stops at  $\frac{1}{4}$ ). The obstruction in that proof is the same one that makes the ray of Proposition 6.5 useless above  $\frac{1}{4}$ : Weyl tightness forces the witness to live in the  $\eta(\lambda)$ -eigenspace of  $\rho_\lambda^\Gamma$ , while sharp-bound tightness forces it to have Schmidt weight  $t_* = (2 + \sqrt{2})/4 = 0.8535533906$ . The two eigenspaces meet for  $\lambda \leq \frac{1}{4}$ , where  $\eta(\lambda)$  is the triple eigenvalue  $b$  with eigenspace  $\text{span}\{|\phi^-\rangle, |01\rangle, |10\rangle\}$  containing  $\psi_0$ , and they are disjoint for  $\lambda > \frac{1}{4}$ , where that eigenspace collapses to  $\mathbb{C}|\phi^+\rangle$ . Below  $\lambda_c$  the eigenspaces still meet, but the resulting state  $\sigma(\lambda)$  stops being positive semidefinite; whether some other equality candidate survives there is the open question discussed at the end of this section.

#### 6.4 Exactness on $[\frac{1}{4}, \frac{1}{2})$

Above  $\frac{1}{4}$  the saturating direction leaves the Werner-symmetric subspace.

**Proposition 6.9** (the attaining family above  $\frac{1}{4}$ ). *For  $\frac{1}{4} \leq \lambda < \frac{1}{2}$  set  $R(\lambda) = \sqrt{(1-\lambda)(1+2\lambda)}$  and*

$$c(\lambda) = \frac{2R(\lambda) - (2\lambda + 1)}{3}, \quad \sigma(\lambda) = \rho_\lambda + c(\lambda)(P_- - |11\rangle\langle 11|). \quad (26)$$

*Then  $c(\lambda) > 0$  for  $\lambda > \frac{1}{4}$ ,  $c(\frac{1}{4}) = \frac{\sqrt{2}-1}{2}$ ,  $c(\lambda) \rightarrow 0$  as  $\lambda \rightarrow \frac{1}{2}$ , and  $\sigma(\lambda)$  is a state with  $\lambda_{\min}(\sigma(\lambda)^\Gamma) = 0$  whose spectra are*

$$\text{spec } \sigma(\lambda) = \left\{ a, a, \lambda + c, a - c \right\}, \quad \text{spec } \sigma(\lambda)^\Gamma = \left\{ 0, \frac{R}{3}, \frac{R}{3}, \frac{3-2R}{3} \right\}, \quad a = \frac{1-\lambda}{3}, \quad (27)$$

*and  $\|\sigma(\lambda) - \rho_\lambda\|_1 = 2c(\lambda)$ . Consequently*

$$d(\lambda) \leq d_{\text{up}}(\lambda) := 2c(\lambda) = \frac{2}{3} \left( 2\sqrt{(1-\lambda)(1+2\lambda)} - (2\lambda + 1) \right). \quad (28)$$

*Proof.*  $P_- - |11\rangle\langle 11|$  is traceless, so  $\text{Tr } \sigma(\lambda) = 1$ . Write  $a = \frac{1-\lambda}{3}$ ,  $b = \frac{1+2\lambda}{6}$ ,  $e = \frac{1-4\lambda}{6}$ , so that  $\rho_\lambda = a(|00\rangle\langle 00| + |11\rangle\langle 11|) + b(|01\rangle\langle 01| + |10\rangle\langle 10|) + e(|01\rangle\langle 10| + |10\rangle\langle 01|)$ , and note  $P_- = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10| - |01\rangle\langle 10| - |10\rangle\langle 01|)$ , whence  $\Gamma(P_-) = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10| - |00\rangle\langle 11| - |11\rangle\langle 00|)$ . Therefore

$$\sigma(\lambda) = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b + \frac{c}{2} & e - \frac{c}{2} & 0 \\ 0 & e - \frac{c}{2} & b + \frac{c}{2} & 0 \\ 0 & 0 & 0 & a - c \end{pmatrix}, \quad \sigma(\lambda)^\Gamma = \begin{pmatrix} a & 0 & 0 & e - \frac{c}{2} \\ 0 & b + \frac{c}{2} & 0 & 0 \\ 0 & 0 & b + \frac{c}{2} & 0 \\ e - \frac{c}{2} & 0 & 0 & a - c \end{pmatrix}.$$

The  $|01\rangle, |10\rangle$  block of  $\sigma(\lambda)$  has eigenvalues  $b + \frac{c}{2} \pm (e - \frac{c}{2})$ , namely  $b + e = a$  and  $b - e + c = \lambda + c$  because  $b - e = \lambda$ ; together with the  $|00\rangle$  entry  $a$  and the  $|11\rangle$  entry  $a - c$  this is the first line of (27). For the partial transpose,  $|01\rangle$  and  $|10\rangle$  give the double eigenvalue  $b + \frac{c}{2} = \frac{1+2\lambda}{6} + \frac{2R-(1+2\lambda)}{6} = \frac{R}{3}$ , and the  $|00\rangle, |11\rangle$  block is  $\begin{pmatrix} a & e-c/2 \\ e-c/2 & a-c \end{pmatrix}$  with determinant

$$a(a-c) - \left(e - \frac{c}{2}\right)^2 = a^2 - ac - e^2 + ec - \frac{c^2}{4} = (a-e)(a+e) - c(a-e) - \frac{c^2}{4} = b \frac{1-2\lambda}{2} - bc - \frac{c^2}{4},$$

using  $a-e = b$  and  $a+e = \frac{1-2\lambda}{2}$ . The block is therefore singular exactly when  $c^2 + 4bc - 2b(1-2\lambda) = 0$ , whose positive root is

$$c = -2b + \sqrt{4b^2 + 2b(1-2\lambda)} = -\frac{1+2\lambda}{3} + \frac{2}{3}\sqrt{(1-\lambda)(1+2\lambda)} = \frac{2R-(2\lambda+1)}{3},$$

because  $4b^2 + 2b(1-2\lambda) = \frac{4}{9}(1+2\lambda)(1-\lambda) = \frac{4}{9}R^2$ . That is (26), and it is the only value of  $c$  for which  $\lambda_{\min}(\sigma^\Gamma) = 0$ . The other root of the block is its trace  $2a - c = \frac{3-2R}{3}$ , which completes (27).

For positivity,  $\lambda + c > 0$  on  $(\frac{1}{4}, \frac{1}{2})$  because  $c > 0$  there, and  $c > 0$  is equivalent to  $4R^2 > (1+2\lambda)^2$ , that is  $4(1-\lambda)(1+2\lambda) > (1+2\lambda)^2$ , that is  $4(1-\lambda) > 1+2\lambda$ , that is  $\lambda < 1$ ; at  $\lambda = \frac{1}{4}$  this gives  $c = \frac{2 \cdot 3\sqrt{2}/4 - 3/2}{3} = \frac{\sqrt{2}-1}{2} > 0$ , and as  $\lambda \rightarrow \frac{1}{2}$ ,  $R \rightarrow 1$  and  $c \rightarrow 0$ . Next  $a - c \geq 0$  is equivalent to  $1 - \lambda \geq 2R - 1 - 2\lambda$ , i.e.  $(2+\lambda)^2 \geq 4R^2 = 4(1-\lambda)(1+2\lambda)$ , i.e.  $9\lambda^2 \geq 0$ ; and  $3 - 2R \geq 0$  holds because  $R^2 = (1-\lambda)(1+2\lambda) \leq \frac{9}{8} < \frac{9}{4}$  on  $[\frac{1}{4}, \frac{1}{2})$ , the maximum of  $(1-\lambda)(1+2\lambda)$  being attained at  $\lambda = \frac{1}{4}$ .

Finally  $\sigma(\lambda) - \rho_\lambda = c(P_- - |11\rangle\langle 11|)$ , and  $P_-$  and  $|11\rangle\langle 11|$  are orthogonal rank-one projectors, so  $\text{spec}(P_- - |11\rangle\langle 11|) = \{1, -1, 0, 0\}$  and  $\|\sigma(\lambda) - \rho_\lambda\|_1 = 2c(\lambda)$ . Since  $\lambda_{\min}(\sigma(\lambda)^\Gamma) = 0$ , Lemma 6.1 turns this into (28).  $\square$

*Remark 6.10.*  $d_{\text{up}}$  joins continuously onto the value of Theorem 6.6 at  $\lambda = \frac{1}{4}$ , where both equal  $\sqrt{2} - 1$ , and it tends to 0 as  $\lambda \rightarrow \frac{1}{2}$ , as it must because  $\rho_{1/2}$  is itself on the boundary. On  $(\frac{1}{4}, \frac{1}{2})$  it is strictly smaller than the Werner competitor  $1 - 2\lambda$  of (8): at  $\lambda = 0.3$  the two are 0.3444006992 and 0.4.

The matching lower bound comes from the witness relaxation. Restrict the witness to the real symmetric plane

$$w = a(|00\rangle + |11\rangle) + b(|01\rangle + |10\rangle), \quad a^2 + b^2 = \frac{1}{2}, \quad a, b \geq 0, \quad c := a^2 \in [0, \frac{1}{2}],$$

which contains  $|\phi^+\rangle$  ( $c = \frac{1}{2}$ ) and  $|\psi^+\rangle$  ( $c = 0$ ). For such a  $w$  a direct evaluation using (6) gives

$$\langle w | \rho_\lambda^\Gamma | w \rangle = 2a^2 \frac{1-2\lambda}{2} + 2b^2 \frac{1+2\lambda}{6} = \frac{1+2\lambda}{6} - \frac{2(4\lambda-1)}{3} c,$$

and the Schmidt weights of  $w$  are  $(a+b)^2$  and  $(a-b)^2$ , so Lemma 4.3 gives

$$\text{hw}((|w\rangle\langle w|)^\Gamma) = \frac{(a+b)^2 + |a^2 - b^2|}{2} = (a+b) \max\{a, b\} = \max\left\{c, \frac{1}{2} - c\right\} + \sqrt{c\left(\frac{1}{2} - c\right)} =: Q(c),$$

a function symmetric under  $c \leftrightarrow \frac{1}{2} - c$  whose maximum is  $K$ , attained at  $c = t_*/2 = (2 + \sqrt{2})/8$  and at  $\frac{1}{2} - t_*/2$ . Hence

$$D_w(\lambda) = \frac{\frac{1+2\lambda}{6} - \frac{2(4\lambda-1)}{3}c}{Q(c)} =: \frac{N_\lambda(c)}{Q(c)}, \quad (29)$$

and since  $N_\lambda$  is affine with slope  $-\frac{2(4\lambda-1)}{3} \leq 0$  for  $\lambda \geq \frac{1}{4}$  while  $Q$  is symmetric,  $D_w(\lambda)$  at  $c$  is at least its value at  $\frac{1}{2} - c$ ; the minimum over the plane is therefore attained at some  $c \in [\frac{1}{4}, \frac{1}{2}]$ , where  $Q(c) = c + \sqrt{c(\frac{1}{2} - c)}$ .

**Proposition 6.11** (the symmetric witness family is exactly sharp above  $\frac{1}{4}$ ). *For  $\frac{1}{4} \leq \lambda < \frac{1}{2}$  and every  $c \in [0, \frac{1}{2}]$ ,*

$$N_\lambda(c) \geq d_{\text{up}}(\lambda) Q(c), \quad (30)$$

*with equality if and only if  $c = c_0(\lambda)$ , where*

$$c_0(\lambda) = \frac{(1+2\lambda)(3+2R(\lambda))}{4(8\lambda^2-4\lambda+5)} = \frac{1+2u_*(\lambda)}{4}, \quad u_*(\lambda) = \frac{\sqrt{1-\lambda}(2\lambda+1)^{3/2} + (1-\lambda)(4\lambda-1)}{8\lambda^2-4\lambda+5}. \quad (31)$$

*Consequently  $\min_c D_w(\lambda) = d_{\text{up}}(\lambda)$  over this family, attained at  $c = c_0(\lambda)$ , and  $c_0(\frac{1}{4}) = \frac{1}{4} + \frac{\sqrt{2}}{8} = t_*/2$ .*

*Proof.* Write  $d = d_{\text{up}}(\lambda)$ ,  $A = \frac{1+2\lambda}{6}$  and  $B = \frac{2(4\lambda-1)}{3}$ , so that  $N_\lambda(c) = A - Bc$ . By the symmetry remark preceding the statement it is enough to prove (30) for  $c \in [\frac{1}{4}, \frac{1}{2}]$ , where  $Q(c) = c + \sqrt{c(\frac{1}{2} - c)}$ . Consider  $\varphi(c) = A - Bc - dQ(c)$  on  $[\frac{1}{4}, \frac{1}{2}]$ . Since  $Q$  is concave there—it is the sum of a linear term and  $\sqrt{c(\frac{1}{2} - c)}$ , whose second derivative is negative—and  $d > 0$ , the function  $\varphi$  is convex. Isolating the radical,  $\varphi(c) \geq 0$  is equivalent to  $A - (B+d)c \geq d\sqrt{c(\frac{1}{2} - c)}$ ; at the two endpoints this reads  $A \geq 0$  and  $A - \frac{B+d}{2} = \frac{1-2\lambda}{2} - \frac{d}{2} \geq 0$ , the latter being  $d \leq 1 - 2\lambda$ , which holds by the remark following Proposition 6.9, and since  $A - (B+d)c$  is affine it is non-negative on the whole interval. Squaring is therefore reversible, and  $\varphi(c) \geq 0$  is equivalent to

$$q(c) := (A - (B+d)c)^2 - d^2c\left(\frac{1}{2} - c\right) = \kappa c^2 + \beta c + A^2 \geq 0,$$

$$\kappa = (B+d)^2 + d^2, \quad \beta = -2A(B+d) - \frac{d^2}{2}.$$

The leading coefficient is positive because  $d > 0$  on  $[\frac{1}{4}, \frac{1}{2})$ , and the discriminant factors as

$$\begin{aligned} \Delta(\lambda) &= \beta^2 - 4\kappa A^2 = \left(2A(B+d) + \frac{d^2}{2}\right)^2 - 4A^2((B+d)^2 + d^2) \\ &= d^2 \left[ \frac{d^2}{4} + 2A(B+d) - 4A^2 \right] = \frac{4d^2}{9} \left[ R(\lambda)^2 - (1-\lambda)(1+2\lambda) \right] = 0, \end{aligned} \quad (32)$$

where the last equality is the defining identity  $R(\lambda)^2 = (1-\lambda)(1+2\lambda)$ . The intermediate step is the polynomial identity  $36\left[\frac{d^2}{4} + 2A(B+d) - 4A^2\right] = 16R^2 + 32\lambda^2 - 16\lambda - 16$ , which follows by substituting  $A = \frac{1+2\lambda}{6}$ ,  $B = \frac{2(4\lambda-1)}{3}$ ,  $d = \frac{2}{3}(2R - (2\lambda+1))$  and expanding; it is checked symbolically in Section 9. Hence  $q(c) = \kappa(c - c_0)^2$  with the double root

$$c_0 = -\frac{\beta}{2\kappa} = \frac{2A(B+d) + \frac{d^2}{2}}{2((B+d)^2 + d^2)},$$

so  $\varphi \geq 0$  with a unique zero at  $c_0$ . That zero is the unique minimiser of  $N_\lambda/Q$ , because  $\varphi(c_0) = 0$  says  $N_\lambda(c_0) = dQ(c_0)$  while  $\varphi(c) > 0$  for  $c \neq c_0$  says  $N_\lambda(c) > dQ(c)$ , and  $Q(c) > 0$  on  $(0, \frac{1}{2}]$ . Substituting the closed forms of  $A$ ,  $B$ ,  $d$  and  $R$  into  $-\beta/(2\kappa)$  and simplifying gives the first

expression in (31), and  $u = 2c - \frac{1}{2}$  gives the second. Because  $c_0(\lambda) \geq t_*/2 > \frac{1}{4}$  on  $[\frac{1}{4}, \frac{1}{2})$ , the minimum over the whole plane is attained there too. At  $\lambda = \frac{1}{4}$ , where  $B = 0$ ,  $A = \frac{1}{4}$ ,  $R = \frac{3\sqrt{2}}{4}$  and  $d = \sqrt{2} - 1$ , the same formula gives  $c_0 = \frac{2+\sqrt{2}}{8} = t_*/2$ , the maximiser of  $Q$ .  $\square$

**The value on  $[\frac{1}{4}, \frac{1}{2})$ .** Combining (28) and (30) through Lemma 6.2 gives

$$d(\lambda) = d_{\text{up}}(\lambda) = \frac{2}{3} \left( 2\sqrt{(1-\lambda)(1+2\lambda)} - (2\lambda + 1) \right), \quad \frac{1}{4} \leq \lambda < \frac{1}{2}, \quad (33)$$

provided the infimum in (21) is attained inside the symmetric plane just analysed. This paper establishes (33) on that reduction and does not re-derive the reduction itself: the argument that no witness outside  $\text{span}\{|\phi^+\rangle, |\psi^+\rangle\}$  beats  $c_0(\lambda)$  belongs to the source study [9], whose note is not part of the material reproduced here. Two independent facts support it. First, the reduction is a lower estimate rather than an optimisation heuristic: the numerator of (20) satisfies  $\langle w | \rho_\lambda^\Gamma w \rangle = \frac{1}{3} [(1-\lambda) - (4\lambda-1)u]$  with  $u = |\langle \phi^+ | w \rangle|^2 - \frac{1}{2}$  for *every* unit  $w$ , so the family (29) is the exact one-parameter image of a quantity defined on the whole sphere, and on the family the denominator attains its largest admissible value. Second, a multistart minimisation of (20) over the full complex unit sphere—24 starts per  $\lambda$ , including four structured ones—returned no witness below  $d_{\text{up}}(\lambda)$  at any of the thirteen sampled values of  $\lambda$  and reproduced (33) to a relative accuracy of  $6.5 \times 10^{-15}$  on  $[\lambda_c, \frac{1}{2})$ ; the minimisers it found for  $\lambda > \frac{1}{4}$  all lie in the symmetric plane, with Schmidt weights drifting from  $(t_*, 1 - t_*)$  at  $\lambda = \frac{1}{4}$  to  $(0.651, 0.349)$  at  $\lambda = 0.40$  (Section 9).

## 6.5 Certified upper bounds below $\lambda_c$

Below  $\lambda_c$  the ray of Proposition 6.5 is not a state, and the Weyl bound (22) is strict; what remains available are explicit boundary states of a different shape. Both families below are built from  $\rho_\lambda + c(P_- \text{-type direction} - |q\rangle\langle q|)$  and are *double corners*: both  $\sigma$  and  $\sigma^\Gamma$  are singular.

**Proposition 6.12** (the isospectral corner branch). *For  $0 \leq \lambda \leq \frac{1}{4}$  set*

$$\mathcal{R}(\lambda) = \frac{(1-\lambda)(1-4\lambda)}{10\lambda^2 - 2\lambda + 1}, \quad c_{\text{iso}}(\lambda) = \frac{10\lambda^2 - 2\lambda + 1}{6(1-\lambda)},$$

$$q(\lambda) = \sqrt{1-\mathcal{R}(\lambda)} |\psi^-\rangle - \sqrt{\mathcal{R}(\lambda)} |\psi^+\rangle, \quad \sigma_{\text{iso}}(\lambda) = \rho_\lambda + c_{\text{iso}}(\lambda) (|\phi^-\rangle\langle\phi^-| - |q(\lambda)\rangle\langle q(\lambda)|).$$

*Then  $\mathcal{R}(\lambda) \in [0, 1]$ ,  $\sigma_{\text{iso}}(\lambda)$  is a state with  $\lambda_{\min}(\sigma_{\text{iso}}^\Gamma) = 0$  and  $\lambda_{\min}(\sigma_{\text{iso}}) = 0$ ,  $\sigma_{\text{iso}}$  and  $\sigma_{\text{iso}}^\Gamma$  have the same spectrum*

$$\left\{ 0, \frac{1-\lambda}{3}, \frac{1-\lambda}{3} + c_{\text{iso}}, \frac{1+4\lambda-14\lambda^2}{6(1-\lambda)} \right\},$$

*and*

$$d(\lambda) \leq d_{\text{iso}}(\lambda) := 2c_{\text{iso}}(\lambda) = \frac{10\lambda^2 - 2\lambda + 1}{3(1-\lambda)} \quad (34)$$

*on  $[0, \frac{1}{4}]$ . Moreover  $d_{\text{iso}}(\lambda_c) = 6 - 4\sqrt{2} = \eta(\lambda)/K$  at  $\lambda = \lambda_c$ ,  $d_{\text{iso}}(0) = \frac{1}{3}$ , and  $d_{\text{iso}}(\lambda) < 1 - 2\lambda$  for  $\lambda < \frac{1}{4}$ .*

**Proposition 6.13** (the rank-two low branch). *For  $0 < \lambda \leq \lambda_0$  set  $s = \sqrt{3(2\lambda+1)}$  and*

$$x = \frac{2\lambda(s-3\lambda)}{(1-4\lambda)(3\lambda+1)}, \quad t = \frac{2\lambda s + 5\lambda + 1}{2(1-\lambda)(3\lambda+1)}, \quad c_{\text{low}} = \frac{(1-\lambda)(3\lambda+1)}{3\lambda+2s+3}, \quad (35)$$

$$\begin{aligned} q_{\text{low}}(\lambda) &= \sqrt{x} |\psi^-\rangle + \sqrt{1-x-t} |\phi^+\rangle + \sqrt{t} |\psi^+\rangle, \\ \sigma_{\text{low}}(\lambda) &= \rho_\lambda + c_{\text{low}}(|\phi^-\rangle\langle\phi^-| - |q_{\text{low}}(\lambda)\rangle\langle q_{\text{low}}(\lambda)|). \end{aligned}$$

Then  $x > 0$ ,  $t > 0$ ,  $x + t < 1$ ,  $\sigma_{\text{low}}(\lambda)$  is a state with  $\lambda_{\min}(\sigma_{\text{low}}^\Gamma) = \lambda_{\min}(\sigma_{\text{low}}) = 0$ , and

$$d(\lambda) \leq \delta_{\text{low}}(\lambda) := 2c_{\text{low}}(\lambda) = \frac{2(1-\lambda)(3\lambda+1)}{3\lambda+2\sqrt{3(2\lambda+1)}+3} \quad (36)$$

on  $(0, \lambda_0]$ , where

$$\lambda_0 = \frac{7-3\sqrt{5}}{4} = 0.0729490168751577\dots \quad (37)$$

is the exact crossover at which  $1-x-t=0$  and  $\delta_{\text{low}}(\lambda_0) = d_{\text{iso}}(\lambda_0) = 0.3262379212492639\dots$ . For  $\lambda > \lambda_0$  the parameters (35) violate  $x+t \leq 1$ , so the branch does not extend. Taking the limit  $\lambda \downarrow 0$  gives the state  $\sigma_0 = \rho_0 + \frac{1}{3+2\sqrt{3}}(|\phi^-\rangle\langle\phi^-| - |q_0\rangle\langle q_0|)$  with  $q_0 = (|\phi^+\rangle + |\psi^+\rangle)/\sqrt{2}$ , hence

$$d(0) \leq \frac{2}{3+2\sqrt{3}} = 0.3094010767585030\dots < d_{\text{iso}}(0) = \frac{1}{3}. \quad (38)$$

Both propositions are proved by direct computation in Appendix B, where the states are written out as matrices, their spectra are displayed in closed form, and the double-corner and tangency equations that determine (35) are derived. The ordering of the three upper bounds and their relation to (22) is

$$\frac{\eta(\lambda)}{K} \leq d(\lambda) \leq \begin{cases} \delta_{\text{low}}(\lambda), & 0 < \lambda \leq \lambda_0, \\ d_{\text{iso}}(\lambda), & \lambda_0 \leq \lambda \leq \lambda_c, \\ \frac{\eta(\lambda)}{K}, & \lambda_c \leq \lambda \leq \frac{1}{4}, \\ d_{\text{up}}(\lambda), & \frac{1}{4} \leq \lambda < \frac{1}{2}, \end{cases} \quad (39)$$

with equality in the third line by Theorem 6.6 and in the fourth by (33). Figure 1 plots (39); Table 1 lists the numbers.

## 6.6 What is proved and what remains open below $\lambda_c$

Two statements complete the picture. The first is a theorem; the second is the remaining gap.

### Strictness of the Weyl bound.

**Theorem 6.14** (the Weyl bound is strict below  $\lambda_c$ ). *For every  $\lambda$  with  $0 < \lambda < \lambda_c$ ,*

$$\frac{\eta(\lambda)}{K} < d(\lambda) \leq \delta_{\text{low}}(\lambda),$$

*so the certified interval of (39) has non-empty interior at every such  $\lambda$ .*

*Proof sketch; the full argument is Appendix A.4.* Assume  $d(\lambda) = \eta(\lambda)/K$ . Compactness gives a limit state  $\sigma$  at that distance with  $\lambda_{\min}(\sigma^\Gamma) = 0$ , and both inequalities used in Theorem 5.1 are equalities. Tracing the equality cases of Weyl's inequality and of (10) through Remark 4.7 shows (Lemma A.4) that  $\sigma$  must have the form  $\rho_\lambda + c(|u_-\rangle\langle u_-| - |u_+\rangle\langle u_+|)$  with  $c = \eta(\lambda)/(2K)$ , where  $\psi \in E_\eta = \{|\phi^+\rangle\}^\perp$  has Schmidt weight  $t_*$  and  $u_\pm$  are the extremal eigenvectors of  $(|\psi\rangle\langle\psi|)^\Gamma$ . Every such  $\psi$  obeys the two structural identities  $\langle u_-|\psi^-\rangle = 0$  and  $|\langle u_+|\psi^-\rangle|^2 = \frac{1}{2}$ , which follow from the moment identities  $\langle \psi^-|R^k|\psi^- \rangle = \frac{1}{2}(t_1^k + t_2^k)$ ,  $k \leq 3$ , together with the distinctness of  $\text{spec } R$

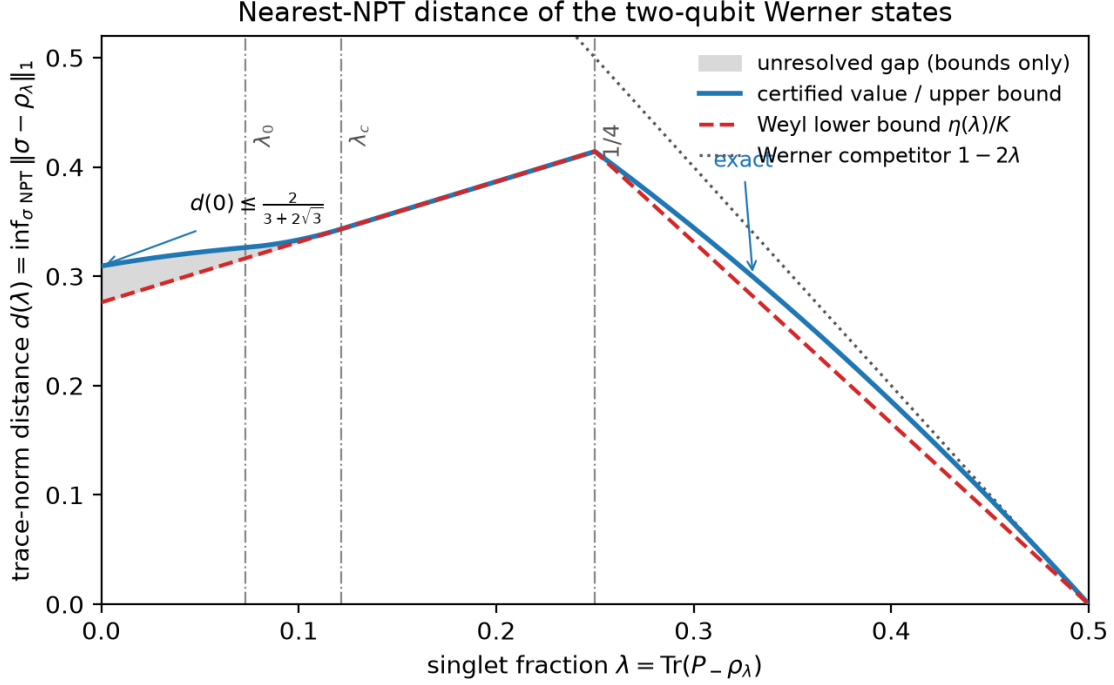


Figure 1: Nearest-NPT trace-norm distance  $d(\lambda)$  for the two-qubit Werner family. The solid blue curve is the certified value where the bounds coincide ( $\lambda \geq \lambda_c$ ) and the best certified upper bound elsewhere; the dashed red curve is the Weyl lower bound  $\eta(\lambda)/K$  of Proposition 6.3; the dotted grey curve is the trivial Werner competitor  $1 - 2\lambda$  of (8). The shaded band on  $[0, \lambda_c]$  is the region where only bounds are known: its width is  $\delta_{\text{low}}(0) - \eta(0)/K = \frac{2}{3+2\sqrt{3}} - \frac{2(\sqrt{2}-1)}{3} = 0.033259$  at  $\lambda = 0$  and closes at  $\lambda_c$ , where the corner branch and the ray meet with the common value  $6 - 4\sqrt{2}$ . The three vertical markers are  $\lambda_0$ ,  $\lambda_c$  and  $\frac{1}{4}$ , the two crossover points of (39) and the kink of  $\eta(\lambda)$ . The two arrows label the segment on which the value is exact and the limit bound  $d(0) \leq 2/(3 + 2\sqrt{3})$  of (38), which is the value the curve takes at the left edge.

(Lemma A.5). Hence  $\rho_\lambda^{-1} = \frac{3}{1-\lambda}I + (\frac{1}{\lambda} - \frac{3}{1-\lambda})P_-$  has the diagonal Gram matrix  $\text{diag}(\frac{1+2\lambda}{2\lambda(1-\lambda)}, \frac{3}{1-\lambda})$  on  $\{u_+, u_-\}$  (Lemma A.6), and the Schur complement criterion turns  $\sigma \geq 0$  into the single scalar condition  $F(\lambda) \leq 1$  for

$$F(\lambda) = \frac{(1+2\lambda)^2}{6(1+\sqrt{2})\lambda(1-\lambda)}.$$

The equation  $F(\lambda) = 1$  is  $(10 + 6\sqrt{2})\lambda^2 - (2 + 6\sqrt{2})\lambda + 1 = 0$ , whose exact roots are  $\lambda_c$  and  $\frac{1}{7} + \frac{3\sqrt{2}}{14} = 0.4459029062\dots$ ; since the leading coefficient is positive,  $F(\lambda) > 1$  for  $0 < \lambda < \lambda_c$ . No candidate is a state there, so equality is impossible. The right inequality is (36).  $\square$

The witness relaxation cannot certify this: by Remark 6.4 it returns exactly  $\eta(\lambda)/K$  for every  $\lambda \leq \frac{1}{4}$ , so the strictness is invisible to witnesses and is a statement about the geometry of the equality candidates instead. What makes it quantitative is the gap, which Section 9 measures at  $3.33 \times 10^{-2}$  at  $\lambda = 0$  and  $9.81 \times 10^{-3}$  at  $\lambda_0$ . The exact value of  $d(\lambda)$  on  $(0, \lambda_c)$  remains open.

**The endpoint  $\lambda = 0$ .** The bound (38) is a limit of feasible states, and  $\rho_0 = P_{\text{sym}}/3$  is rank three, so the rank-two branch does not apply at  $\lambda = 0$  itself. Whether  $d(0) = 2/(3 + 2\sqrt{3})$  is open; the

Table 1: Certified values of  $d(\lambda)$ . Column 2 is the Weyl lower bound  $\eta(\lambda)/K$  of Proposition 6.3; column 3 is the numerical value of the witness relaxation  $\inf_w D_w(\lambda)$  of Lemma 6.2, obtained by 24-start minimisation over the complex unit sphere; column 4 is the best certified upper bound from an explicit boundary state ( $\delta_{\text{low}}$ ,  $d_{\text{iso}}$ , the ray, or  $d_{\text{up}}$  as in (39)); column 5 is their difference. A zero in column 5 means the value is exact; a positive entry means only bounds are known. Values are printed to ten decimals and were computed in IEEE double precision.

|                         | $\lambda$     | $\eta(\lambda)/K$ | $\inf_w D_w(\lambda)$ | certified bound | status                           |
|-------------------------|---------------|-------------------|-----------------------|-----------------|----------------------------------|
|                         | 0             | 0.2761423749      | 0.2761423749          | 0.3094010768    | bound, gap $3.33 \times 10^{-2}$ |
|                         | 0.02          | 0.2871880699      | 0.2871880699          | 0.3151362310    | bound, gap $2.79 \times 10^{-2}$ |
|                         | 0.05          | 0.3037566124      | 0.3037566124          | 0.3221202833    | bound, gap $1.84 \times 10^{-2}$ |
| $\lambda_0 = 0.0729490$ | 0.3164310045  | 0.3164310045      | 0.3262379212          | 0.3262379212    | bound, gap $9.81 \times 10^{-3}$ |
|                         | 0.09          | 0.3258480024      | 0.3258480024          | 0.3300366300    | bound, gap $4.19 \times 10^{-3}$ |
| $\lambda_c = 0.1213203$ | 0.3431457505  | 0.3431457505      | 0.3431457505          | 0.3431457505    | exact, $6 - 4\sqrt{2}$           |
|                         | 0.16          | 0.3645079349      | 0.3645079349          | 0.3645079349    | exact                            |
|                         | 0.20          | 0.3865993249      | 0.3865993249          | 0.3865993249    | exact                            |
|                         | $\frac{1}{4}$ | 0.4142135624      | 0.4142135624          | 0.4142135624    | exact, $\sqrt{2} - 1$            |
|                         | 0.30          | 0.3313708499      | 0.3444006992          | 0.3444006992    | exact*                           |
|                         | 0.35          | 0.2485281374      | 0.2682530694          | 0.2682530694    | exact*                           |
|                         | 0.40          | 0.1656854249      | 0.1856406461          | 0.1856406461    | exact*                           |
|                         | 0.49          | 0.0165685425      | 0.0198507380          | 0.0198507380    | exact*                           |

\*Exact subject to the witness-family reduction described before (33); the upper bound itself is unconditional.

source study's own list of open items includes exactly this endpoint.

## 7 The relative entropy of entanglement of the Werner states

This section is an application. Its value,  $E_R(\rho_\lambda) = \ln 2 - H_{\text{nat}}(\lambda)$ , and the candidate minimiser  $\sigma_*$  are classical [5, 13, 6]; what is assembled here is the complete closest-state classification and the calibration that Section 8 turns into a bound for every two-qubit state.

### 7.1 Reduction to one variable

For  $G \in [0, 1]$  write  $\sigma_G = G P_- + \frac{1-G}{3} P_{\text{sym}} = \rho_G$  for the Werner state with singlet fraction  $G$ .

**Theorem 7.1** (twirling reduction). *For every  $\lambda \in [0, 1]$ ,*

$$E_R(\rho_\lambda) = \min_{0 \leq G \leq \frac{1}{2}} f_\lambda(G), \quad f_\lambda(G) = \lambda \ln \frac{\lambda}{G} + (1 - \lambda) \ln \frac{1 - \lambda}{1 - G}, \quad (40)$$

with the conventions  $0 \ln 0 = 0$  and  $f_\lambda(0) = +\infty$  when  $\lambda > 0$ . Moreover  $\sigma_G \in \mathcal{S}$  if and only if  $G \leq \frac{1}{2}$ .

*Proof.* Let  $\sigma \in \mathcal{S}$  and put  $G = \text{Tr}[P_- \sigma]$ . By Lemma 3.4,  $G \leq \frac{1}{2}$ , and by Lemma 3.5,  $\mathcal{T}(\sigma) = \sigma_G$  is again separable while  $\mathcal{T}(\rho_\lambda) = \rho_\lambda$ . Monotonicity of the relative entropy under completely positive trace-preserving maps [14] gives

$$D(\rho_\lambda \| \sigma) \geq D(\mathcal{T}(\rho_\lambda) \| \mathcal{T}(\sigma)) = D(\rho_\lambda \| \sigma_G),$$

so the infimum over  $\mathcal{S}$  equals the infimum over  $\{\sigma_G : 0 \leq G \leq \frac{1}{2}\}$ . Conversely every such  $\sigma_G$  is separable:  $\sigma_G$  is a Werner state, whose partial transpose has spectrum  $\{\frac{1-2G}{2}, \frac{1+2G}{6}, \frac{1+2G}{6}, \frac{1+2G}{6}\}$

by (6), so  $\sigma_G$  is PPT exactly when  $G \leq \frac{1}{2}$ , and PPT is separability on two qubits (Theorem 3.2). Finally,  $\rho_\lambda$  and  $\sigma_G$  commute and are both functions of  $P_-$  and  $P_{\text{sym}}$ , so

$$\begin{aligned} D(\rho_\lambda \| \sigma_G) &= \text{Tr} \left[ \rho_\lambda \left( \ln \lambda P_- + \ln \frac{1-\lambda}{3} P_{\text{sym}} - \ln G P_- - \ln \frac{1-G}{3} P_{\text{sym}} \right) \right] \\ &= \lambda \ln \frac{\lambda}{G} + (1-\lambda) \ln \frac{1-\lambda}{1-G} = f_\lambda(G), \end{aligned}$$

using  $\text{Tr}[P_- \rho_\lambda] = \lambda$  and  $\text{Tr}[P_{\text{sym}} \rho_\lambda] = 1 - \lambda$ .  $\square$

## 7.2 The minimum and the value

**Theorem 7.2** (Werner relative entropy of entanglement). *For  $\lambda \in [0, 1]$ ,*

$$E_R(\rho_\lambda) = F(\lambda) = \begin{cases} 0, & 0 \leq \lambda \leq \frac{1}{2}, \\ \lambda \ln(2\lambda) + (1-\lambda) \ln(2(1-\lambda)) = \ln 2 - H_{\text{nat}}(\lambda), & \frac{1}{2} < \lambda \leq 1, \end{cases} \quad (41)$$

where  $H_{\text{nat}}(\lambda) = -\lambda \ln \lambda - (1-\lambda) \ln(1-\lambda)$  is the binary Shannon entropy in nats. In base-two units the entangled branch reads  $1 - H_2(\lambda)$  ebits. The minimum in (40) is attained at  $G = \lambda$  when  $\lambda \leq \frac{1}{2}$  and at  $G = \frac{1}{2}$  when  $\lambda > \frac{1}{2}$ .

*Proof.* On  $(0, 1)$ ,

$$f'_\lambda(G) = -\frac{\lambda}{G} + \frac{1-\lambda}{1-G}, \quad f''_\lambda(G) = \frac{\lambda}{G^2} + \frac{1-\lambda}{(1-G)^2} > 0,$$

so  $f_\lambda$  is strictly convex with a single stationary point, at  $G = \lambda$ , where  $f_\lambda(\lambda) = 0$ . If  $\lambda \leq \frac{1}{2}$  that point lies in the feasible interval  $[0, \frac{1}{2}]$ , so the minimum is 0; this agrees with  $\rho_\lambda \in \mathcal{S}$  and with  $D(\rho_\lambda \| \rho_\lambda) = 0$ . If  $\lambda > \frac{1}{2}$  then  $f_\lambda$  is strictly decreasing on  $(0, \lambda) \supseteq (0, \frac{1}{2}]$ , so the minimum over  $[0, \frac{1}{2}]$  is at  $G = \frac{1}{2}$ , with value

$$f_\lambda\left(\frac{1}{2}\right) = \lambda \ln(2\lambda) + (1-\lambda) \ln(2(1-\lambda)) = \ln 2 - H_{\text{nat}}(\lambda),$$

where the identity is  $\lambda \ln \lambda + (1-\lambda) \ln(1-\lambda) = -H_{\text{nat}}(\lambda)$  and  $\lambda \ln 2 + (1-\lambda) \ln 2 = \ln 2$ . The endpoints are covered by the same formula:  $F(0) = 0$  and  $F(1) = \ln 2$ , the latter being one ebit.  $\square$

*Remark 7.3* (shape of the curve).  $F$  is continuous on  $[0, 1]$ , identically zero on the separable half, and on  $(\frac{1}{2}, 1)$  it satisfies

$$F'(\lambda) = \ln \frac{\lambda}{1-\lambda}, \quad F''(\lambda) = \frac{1}{\lambda(1-\lambda)} > 0,$$

so it is strictly increasing and strictly convex there, with  $F'(\lambda) \rightarrow 0$  as  $\lambda \downarrow \frac{1}{2}$  and  $F'(\lambda) \rightarrow +\infty$  as  $\lambda \uparrow 1$ . The value and the first derivative are therefore continuous across the PPT threshold, while the second derivative jumps from 0 on the separable side to 4 on the entangled side:  $F$  is  $C^1$  but not  $C^2$  at  $\lambda = \frac{1}{2}$ . Representative values are  $F(\frac{1}{2}) = 0$ ,  $F(\frac{3}{4}) = 0.1308120359$  nats = 0.1887218755 ebits and  $F(1) = \ln 2 = 0.6931471806$  nats = 1 ebit. Figure 2 shows the curve.

*Remark 7.4* (the isotropic parametrisation). The same family is often written as an isotropic state,  $\rho_p = p P_- + (1-p) \frac{1}{4} I_4$  with  $p \in [0, 1]$ . Its singlet fraction is  $\lambda = \text{Tr}[P_- \rho_p] = p + (1-p)/4 = (1+3p)/4$ , so  $p = (4\lambda - 1)/3$ , the PPT threshold  $\lambda = \frac{1}{2}$  is  $p = \frac{1}{3}$ , and the value of Theorem 7.2 reads  $E_R(\rho_p) = \ln 2 - H_{\text{nat}}(\frac{1+3p}{4})$  for  $p > \frac{1}{3}$  and 0 otherwise. Figure 2 is drawn against  $\lambda$  and marks  $p = \frac{1}{3}$  at the threshold for readers who use the other convention.

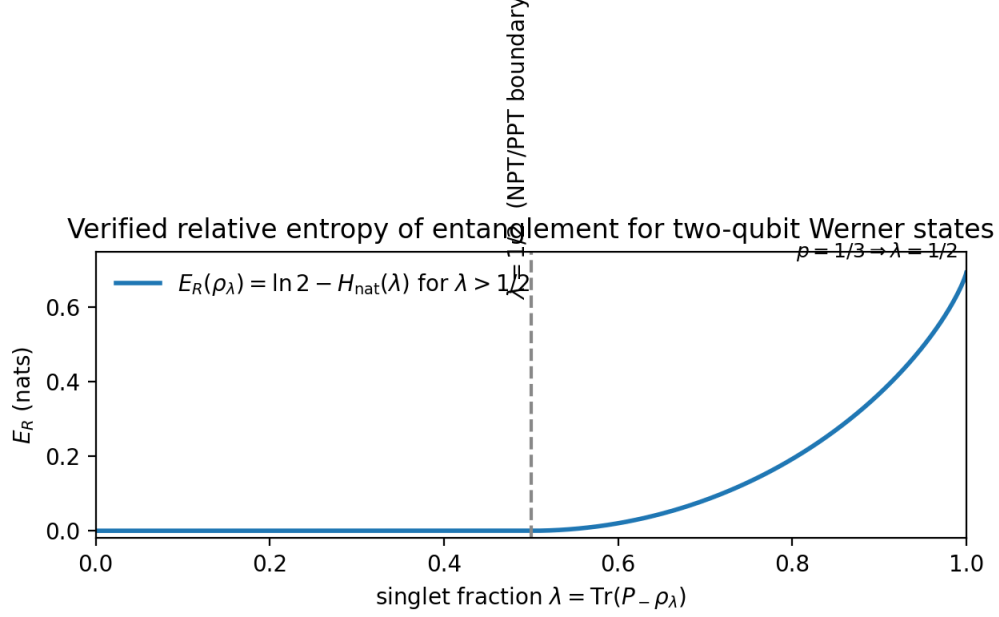


Figure 2: Relative entropy of entanglement of the two-qubit Werner state  $\rho_\lambda$ , in nats, against the singlet fraction  $\lambda$  (solid curve): flat and zero on the separable half  $[0, \frac{1}{2}]$ , then  $\ln 2 - H_{\text{nat}}(\lambda)$  on  $(\frac{1}{2}, 1]$ , rising with vertical tangent to one ebit at the Bell point. The dashed vertical line marks the PPT threshold  $\lambda = \frac{1}{2}$ , where the closest separable state jumps from  $\rho_\lambda$  to the fixed boundary state  $\sigma_* = \rho_{1/2}$  while the value and the slope stay continuous. The annotation at the upper right gives the same threshold in the isotropic parametrisation of Remark 7.4,  $p = \frac{1}{3}$ .

### 7.3 The closest separable states

**Theorem 7.5** (complete classification). *Let  $0 \leq \lambda \leq 1$ .*

- (i) *For  $0 \leq \lambda \leq \frac{1}{2}$  the state  $\rho_\lambda$  is separable and is its own unique closest separable state.*
- (ii) *For  $\frac{1}{2} < \lambda < 1$  the unique closest separable state is*

$$\sigma_* = \rho_{1/2} = \frac{1}{2}P_- + \frac{1}{6}P_{\text{sym}} = \frac{1}{6}(I_4 + 2P_-), \quad \text{spec}(\sigma_*) = \left\{\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{2}\right\}, \quad (42)$$

*which is independent of  $\lambda$ .*

- (iii) *For  $\lambda = 1$  the closest separable states are exactly the separable  $\sigma$  with  $\sigma|\psi^-\rangle = |\psi^-\rangle/2$ . This set contains  $\sigma_*$  and the continuous family*

$$\sigma_\theta = \frac{1}{2}P_- + \frac{1}{2}|u_\theta\rangle\langle u_\theta|, \quad u_\theta = \cos\theta|\psi^+\rangle + \sin\theta|\phi^-\rangle, \quad 0 \leq \theta < \pi, \quad (43)$$

*whose members are pairwise distinct, have  $\text{spec}(\sigma_\theta) = \{0, 0, \frac{1}{2}, \frac{1}{2}\}$  and  $\text{spec}(\sigma_\theta^\Gamma) = \{0, 0, \frac{1}{2}, \frac{1}{2}\}$ , and all satisfy  $D(P_- \| \sigma_\theta) = \ln 2$ .*

*Proof.* (i) For  $\lambda < \frac{1}{2}$ ,  $\rho_\lambda \in \mathcal{S}$  and  $D(\rho_\lambda \| \rho_\lambda) = 0$ , so the minimum is 0 and it is attained only where the relative entropy vanishes, which is exactly at  $\sigma = \rho_\lambda$ . At  $\lambda = 0$  the same argument applies:  $\rho_0 = P_{\text{sym}}/3$  is separable and  $D(\rho_0 \| \sigma) = 0$  forces  $\sigma = \rho_0$ .

(ii) For  $\frac{1}{2} < \lambda < 1$  the state  $\rho_\lambda$  is positive definite, since its eigenvalues are  $\lambda$  and  $(1 - \lambda)/3$  with multiplicity three. Suppose two distinct  $\sigma_1, \sigma_2 \in \mathcal{S}$  both attain the minimum. Then  $\bar{\sigma} = (\sigma_1 + \sigma_2)/2 \in \mathcal{S}$ , and because  $\rho_\lambda > 0$  both  $\sigma_1$  and  $\sigma_2$  are positive definite:  $D(\rho_\lambda \| \sigma_i) < \infty$  forces  $\text{supp } \rho_\lambda \subseteq \text{supp } \sigma_i$ . The logarithm is strictly operator concave, so  $\ln \bar{\sigma} \geq \frac{1}{2}(\ln \sigma_1 + \ln \sigma_2)$  with equality only if  $\sigma_1 = \sigma_2$  [14]; taking the expectation in the positive definite state  $\rho_\lambda$  preserves the strictness and gives  $D(\rho_\lambda \| \bar{\sigma}) < \frac{1}{2}D(\rho_\lambda \| \sigma_1) + \frac{1}{2}D(\rho_\lambda \| \sigma_2) = E_R(\rho_\lambda)$ , a contradiction. Hence the minimiser is unique, and by Theorem 7.2 it is  $\sigma_{1/2} = \sigma_*$ .

(iii) At  $\lambda = 1$ ,  $D(P_- \| \sigma) = -\langle \psi^- | (\ln \sigma) \psi^- \rangle$  whenever  $\text{supp } P_- \subseteq \text{supp } \sigma$ , and  $+\infty$  otherwise; a minimiser must therefore satisfy  $\langle \psi^- | \sigma \psi^- \rangle > 0$ . For such a  $\sigma$ , operator concavity of the logarithm gives  $\langle \psi^- | (\ln \sigma) \psi^- \rangle \leq \ln \langle \psi^- | \sigma \psi^- \rangle$ , with equality if and only if  $|\psi^- \rangle$  is an eigenvector of  $\sigma$  [14], and Lemma 3.4 gives  $\langle \psi^- | \sigma \psi^- \rangle \leq \frac{1}{2}$ . Hence

$$D(P_- \| \sigma) \geq -\ln \langle \psi^- | \sigma \psi^- \rangle \geq -\ln \frac{1}{2} = \ln 2 = F(1),$$

and equality throughout forces  $\langle \psi^- | \sigma \psi^- \rangle = \frac{1}{2}$  together with  $\sigma |\psi^- \rangle = \frac{1}{2} |\psi^- \rangle$ ; conversely that eigenvector equation implies both equalities, so it characterises the minimisers. For  $\sigma_\theta$ , the vector  $u_\theta$  lies in the span of  $|\psi^+ \rangle, |\phi^- \rangle$ , which is orthogonal to  $|\psi^- \rangle$ , so  $\sigma_\theta |\psi^- \rangle = \frac{1}{2} |\psi^- \rangle$ , and  $\sigma_\theta \geq 0$  with trace 1 and eigenvalues  $\{0, 0, \frac{1}{2}, \frac{1}{2}\}$ . Its partial transpose has the same spectrum, as is computed in Appendix B, so  $\sigma_\theta$  is PPT and hence separable by Theorem 3.2. Distinct  $\theta \in [0, \pi)$  give distinct projectors  $|u_\theta \rangle \langle u_\theta|$  and hence distinct states.  $\square$

*Remark 7.6.* Part (iii) is the only degeneracy in the family, and it is a rank phenomenon: the Bell state is the unique member with  $\rho_\lambda$  singular, and the strict convexity argument of part (ii) needs positive definiteness of the first argument. The degeneracy is classical [6]; what the classification adds is that it is exhaustive, and that no other  $\lambda$  shares it.

## 7.4 An operator-level cross-check

The value and the minimiser can be checked without performing any minimisation, by running the inverse construction of Miranowicz and Ishizaka [6] backwards. Let  $\sigma$  be a full-rank separable two-qubit state whose partial transpose has a one-dimensional kernel spanned by  $\varphi$ , let  $\{\lambda_i\}$  be the eigenvalues of  $\sigma$  with eigenvectors  $|i \rangle$ , and set

$$\mathcal{G}(\sigma) = \sum_{i,j} G_{ij} |i \rangle \langle i| (|\varphi \rangle \langle \varphi|)^\Gamma |j \rangle \langle j|, \quad G_{ij} = \begin{cases} \lambda_i, & \lambda_i = \lambda_j, \\ \frac{\lambda_i - \lambda_j}{\ln \lambda_i - \ln \lambda_j}, & \lambda_i \neq \lambda_j, \end{cases} \quad (44)$$

where the second line is the divided difference of  $\ln$ . Their result is that the entangled states having  $\sigma$  as a closest separable state are exactly  $\rho = \sigma - x \mathcal{G}(\sigma)$  for  $x \geq 0$  in the feasible range, with

$$E_R(\rho) = S(\sigma) - S(\rho) + x \text{Tr}[(|\varphi \rangle \langle \varphi|)^\Gamma \sigma \ln \sigma]. \quad (45)$$

Applying (44) to  $\sigma_*$ , whose partial transpose has kernel  $\mathbb{C}|\phi^+ \rangle$  by (6) at  $\lambda = \frac{1}{2}$ , produces exactly the entangled Werner ray:

$$\sigma_* - x \mathcal{G}(\sigma_*) = \rho_{\lambda(x)}, \quad \lambda(x) = \frac{1}{2} + \frac{x}{4}, \quad x \in [0, 2], \quad (46)$$

and (45) returns  $F(\lambda(x))$ . This is an independent route to the same closed form, using a formula derived for a different purpose, and it also confirms that  $\sigma_*$  is not merely a closest separable state

of the entangled Werner states but the one predicted by the general theory. Numerically, (46) holds with a maximum entrywise operator error of  $3.3 \times 10^{-16}$  over 41 values of  $x$ , and (45) matches  $F(\lambda)$  to  $2.0 \times 10^{-14}$  (Section 9).

*Remark 7.7* (the only eigenvalue pair that matters). Since  $\text{spec}(\sigma_*) = \{\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{2}\}$ , the divided difference in (44) is used only on the degenerate blocks, where it reduces to  $G_{ij} = \lambda_i$ , and on the single mixed pair  $(\frac{1}{2}, \frac{1}{6})$ , where it equals  $(\frac{1}{2} - \frac{1}{6})/(\ln \frac{1}{2} - \ln \frac{1}{6}) = \frac{1}{3}/\ln 3 = 0.3034128647\dots$ . The construction (44) therefore involves no limiting procedure.

## 8 A two-sided bound for every two-qubit state

The Werner value is a number attached to a one-parameter family. The point of this section is that the same twirling data turns it into a bound on  $E_R(\rho)$  for an arbitrary two-qubit state, with an explicit and uniformly controlled error.

### 8.1 The sandwich

**Theorem 8.1** (two-sided singlet-fraction bound). *Let  $\rho$  be any two-qubit state, let  $\lambda = \lambda(\rho) = \text{Tr}[P_- \rho]$  be its singlet fraction, and let  $\mu = \min\{\lambda, \frac{1}{2}\}$ . Then*

$$F(\lambda) \leq E_R(\rho) \leq D(\rho \| \rho_\mu), \quad (47)$$

where  $F$  is the Werner value (41). The competitor  $\rho_\mu$  on the right is separable, and both inequalities become equalities for every  $\rho = \rho_\lambda$ .

*Proof.* For the lower bound, let  $\sigma \in \mathcal{S}$  and  $G = \text{Tr}[P_- \sigma] \leq \frac{1}{2}$  by Lemma 3.4. Lemma 3.5 gives  $\mathcal{T}(\sigma) = \sigma_G$  and  $\mathcal{T}(\rho) = \rho_\lambda = \rho_\lambda$ , so monotonicity of the relative entropy yields  $D(\rho \| \sigma) \geq D(\rho_\lambda \| \sigma_G) \geq E_R(\rho_\lambda) = F(\lambda)$  by Theorem 7.2. Taking the infimum over  $\sigma$  gives the left inequality.

For the upper bound,  $\mu \leq \frac{1}{2}$  makes  $\rho_\mu$  PPT and hence separable, so  $E_R(\rho) \leq D(\rho \| \rho_\mu)$ .

On the Werner line,  $\rho = \rho_\lambda$  gives  $\lambda(\rho) = \lambda$  and  $\mu = \min\{\lambda, \frac{1}{2}\}$ , and the right-hand side is  $D(\rho_\lambda \| \rho_\mu) = F(\lambda)$  by Theorem 7.2: the two sides coincide.  $\square$

### 8.2 The correction is an entropy deficit

The right-hand side of (47) differs from the left by a quantity that has two useful exact expressions.

**Proposition 8.2** (deficit identity). *With  $\lambda, \mu$  as in Theorem 8.1,*

$$D(\rho \| \rho_\mu) = F(\lambda) + \Delta(\rho), \quad \Delta(\rho) := S(\rho_\lambda) - S(\rho) = D(\rho \| \rho_\lambda) \geq 0, \quad (48)$$

and  $\Delta(\rho) = 0$  if and only if  $\rho = \rho_\lambda$ .

*Proof.* Both  $\rho_\mu$  and  $\rho_\lambda$  are functions of  $P_-$  and  $P_{\text{sym}}$ , so  $\ln \rho_\mu = \ln \mu P_- + \ln \frac{1-\mu}{3} P_{\text{sym}}$  and  $\ln \rho_\lambda = \ln \lambda P_- + \ln \frac{1-\lambda}{3} P_{\text{sym}}$ . Since  $\text{Tr}[P_- \rho] = \lambda$  and  $\text{Tr}[P_{\text{sym}} \rho] = 1 - \lambda$ ,

$$D(\rho \| \rho_\mu) = -S(\rho) - \lambda \ln \mu - (1 - \lambda) \ln \frac{1 - \mu}{3}, \quad D(\rho_\lambda \| \rho_\mu) = -S(\rho_\lambda) - \lambda \ln \mu - (1 - \lambda) \ln \frac{1 - \mu}{3},$$

and their difference is  $S(\rho_\lambda) - S(\rho)$ . The second quantity equals  $F(\lambda)$  because  $\rho_\mu$  is the closest separable state of  $\rho_\lambda$  (Theorem 7.5), which gives the first identity in (48). For the second, the same computation with  $\mu = \lambda$  gives  $\text{Tr}[\rho \ln \rho_\lambda] = \text{Tr}[\rho_\lambda \ln \rho_\lambda] = -S(\rho_\lambda)$ , hence  $D(\rho \| \rho_\lambda) = S(\rho_\lambda) - S(\rho) = \Delta(\rho)$ , which is non-negative and vanishes only at  $\rho = \rho_\lambda$ .  $\square$

*Remark 8.3.* The identity  $\Delta(\rho) = D(\rho\|\rho_\lambda)$  says that the gap between the two sides of (47) is exactly the information lost by replacing  $\rho$  with the Werner state of the same singlet fraction. It is therefore a measure of non-Werner-ness, not of entanglement, and it vanishes on the whole Werner line whether or not the state is entangled.

**Corollary 8.4** (uniform trace-norm modulus). *Let  $T_F = \frac{1}{2}\|\rho - \rho_\lambda\|_1 \in [0, 1]$  and  $h(T) = -T \ln T - (1 - T) \ln(1 - T)$ . Then*

$$F(\lambda) \leq E_R(\rho) \leq F(\lambda) + T_F \ln 3 + h(T_F), \quad (49)$$

*so the remainder is  $O(T_F \ln \frac{1}{T_F})$  uniformly as  $\rho$  approaches the Werner line.*

*Proof.*  $\Delta(\rho) = S(\rho_\lambda) - S(\rho)$  is bounded by the sharp entropy continuity estimate in dimension four,  $|S(\rho) - S(\rho_\lambda)| \leq T_F \ln(4 - 1) + h(T_F)$  [7].  $\square$

*Remark 8.5* (where the two controls differ). The modulus (49) is state independent and dimension generic; the deficit form (48) is exact and is the one to use when  $\rho$  is known, because  $T_F \ln 3 + h(T_F)$  is an upper bound for it and is loose away from  $T_F = 0$ . Conversely, near  $\lambda = \frac{1}{2}$  on the separable side the deficit can be replaced by the stronger statement that  $E_R$  vanishes identically: by Corollary 5.3, every state within trace-norm distance  $\eta(\lambda)/K$  of  $\rho_\lambda$  is separable, so  $E_R \equiv 0$  on that ball for  $\lambda < \frac{1}{2}$  and all its directional derivatives there are zero.

### 8.3 First-order variation along the entangled Werner line

On the entangled half of the Werner line the sandwich pins  $E_R$  to first order.

**Lemma 8.6** (vanishing first variation of the relative entropy). *Let  $\sigma$  be a positive definite state and let  $Z(\varepsilon)$  be a family of Hermitian operators with  $\text{Tr } Z(\varepsilon) = 0$ ,  $\|Z(\varepsilon)\| = O(\varepsilon)$  and  $\sigma + Z(\varepsilon) \geq 0$ . Then  $D(\sigma + Z(\varepsilon) \|\sigma) = O(\varepsilon^2)$ .*

*Proof.* The map  $Z \mapsto \ln(\sigma + Z)$  is analytic near  $Z = 0$  with derivative  $L(Z) = \int_0^\infty (\sigma + t)^{-1} Z (\sigma + t)^{-1} dt$ , so  $\ln(\sigma + Z) - \ln \sigma = L(Z) + O(\|Z\|^2)$  and

$$D(\sigma + Z \|\sigma) = \text{Tr}[\sigma L(Z)] + \text{Tr}[Z L(Z)] + O(\|Z\|^3).$$

Because  $\sigma$  commutes with  $(\sigma + t)^{-1}$ ,  $\int_0^\infty \sigma(\sigma + t)^{-2} dt = I$ , hence  $\text{Tr}[\sigma L(Z)] = \text{Tr } Z = 0$ . The remaining terms are  $O(\|Z\|^2)$ .  $\square$

**Theorem 8.7** (Fréchet derivative on the entangled Werner line). *Let  $\frac{1}{2} < \lambda < 1$ , let  $H$  be Hermitian and traceless, and let  $\rho_\varepsilon = \rho_\lambda + \varepsilon H \geq 0$  for  $|\varepsilon|$  small. Then*

$$E_R(\rho_\varepsilon) = F(\lambda) + \varepsilon \text{Tr}[P_- H] \ln \frac{\lambda}{1 - \lambda} + O(\varepsilon^2), \quad (50)$$

*that is,  $E_R$  is Fréchet differentiable at  $\rho_\lambda$  with derivative*

$$H \longmapsto \text{Tr}[H(\ln \rho_\lambda - \ln \sigma_*)] = \text{Tr}[P_- H] \ln \frac{\lambda}{1 - \lambda}. \quad (51)$$

*Proof.* Write  $\lambda_\varepsilon = \lambda(\rho_\varepsilon) = \lambda + \varepsilon \text{Tr}[P_- H]$  and  $\sigma_\varepsilon = \rho_{\lambda_\varepsilon}$ ; for small  $\varepsilon$  one has  $\lambda_\varepsilon > \frac{1}{2}$ , so  $\sigma_\varepsilon$  is positive definite. Then  $Z_\varepsilon = \rho_\varepsilon - \sigma_\varepsilon = \varepsilon(H - \text{Tr}[P_- H](P_- - \frac{1}{3}P_{\text{sym}}))$  is traceless with  $\|Z_\varepsilon\| = O(\varepsilon)$ , and Lemma 8.6 gives  $\Delta(\rho_\varepsilon) = D(\rho_\varepsilon \|\sigma_\varepsilon) = O(\varepsilon^2)$ . Theorem 8.1 with Proposition 8.2 then sandwiches  $E_R(\rho_\varepsilon)$  between  $F(\lambda_\varepsilon)$  and  $F(\lambda_\varepsilon) + O(\varepsilon^2)$ . Since  $F$  is analytic on  $(\frac{1}{2}, 1)$  with  $F'(\lambda) = \ln \frac{\lambda}{1 - \lambda}$ , expanding gives (50). For (51),  $\ln \rho_\lambda - \ln \sigma_* = \ln(2\lambda)P_- + \ln(2(1 - \lambda))P_{\text{sym}}$ , and  $\text{Tr } H = 0$  gives  $\text{Tr}[HP_{\text{sym}}] = -\text{Tr}[HP_-]$ , so the trace collapses to  $[\ln(2\lambda) - \ln(2(1 - \lambda))] \text{Tr}[P_- H]$ .  $\square$

*Remark 8.8* (what the derivative says). The derivative depends on  $H$  only through  $\text{Tr}[P_-H]$ , the displacement of the singlet fraction. Directions that preserve the singlet fraction leave  $E_R$  stationary to first order, and the second-order term is controlled by  $\Delta(\rho_\varepsilon) = O(\varepsilon^2)$ ; in particular the closest separable state of  $\rho_\varepsilon$  stays within  $O(\varepsilon)$  of  $\sigma_*$ , so the minimiser is pinned to first order even though it moves. On the separable side,  $\lambda < \frac{1}{2}$ , the same argument is unnecessary: Remark 8.5 shows that  $E_R$  is identically zero in a neighbourhood of  $\rho_\lambda$ , so every directional derivative vanishes there and  $F'(\lambda) = 0$  matches.

## 9 Numerical results

Every quantitative statement in this paper is proved, and the proofs use only explicit spectra and one-variable calculus. The computations summarised here have a different job: they locate arithmetic slips in the closed forms, they test the inequalities against random inputs that no hand example would produce, and in one place they supply the only available evidence for a step whose proof is not reproduced in this paper. All of it is ordinary floating-point arithmetic—IEEE double precision for eigenvalues, norms and optimisation, exact rational and radical arithmetic for the symbolic identities—with fixed seeds and no adaptive tolerances.

### 9.1 What was computed

Tables 2 and 3 list the checks by the statement they support, the first covering the geometry and the second the entanglement results. Three points about the design are worth making explicit.

First, the random tests are adversarial rather than confirmatory. For the sharp bound (10), 4000 traceless Hermitian matrices were drawn from a Gaussian ensemble on the 15-dimensional traceless subspace; the largest ratio  $\|X^\Gamma\|_{\text{op}}/\|X\|_1$  observed was 0.5544, which is  $0.9186K$ , and no draw violated the bound. Random draws do not come close to saturating it, which is why the extremisers of Proposition 4.6 are exhibited exactly rather than searched for.

Second, the closed forms are compared against the objects they describe, not against each other. The branch  $\delta_{\text{low}}$  is tested by building the state  $\sigma_{\text{low}}(\lambda)$  from (35) and checking positive semidefiniteness, the double-corner condition, and  $\|\sigma_{\text{low}} - \rho_\lambda\|_1 = 2c$  numerically, rather than by checking that two formulas for  $2c$  agree. The same applies to the isospectral branch, to the upper-interval family, and to the ray.

Third, the symbolic checks are exact. The discriminant identity (32) was reduced to zero as a polynomial identity in  $\lambda$  and  $R$ ; the twirling objective, its stationary point, its strict convexity, the identity  $f_\lambda(\frac{1}{2}) = \ln 2 - H_{\text{nat}}(\lambda)$ , and the partial-transpose spectrum of  $\sigma_G$  were all manipulated in exact arithmetic. Where a numerical tolerance appears in either table it is the residual of a floating-point evaluation of an exactly verified identity, or the outcome of a random search.

Two rows of Table 3 deserve naming because they test a formula rather than an inequality. The inverse-map row is the operator-level cross-check of Section 7: the Miranowicz–Ishizaka construction (46) is pushed through 41 values of the ray parameter, each returning a state whose relative entropy (45) reproduces, with operator error  $3.3 \times 10^{-16}$  and value error  $2.0 \times 10^{-14}$ . The Fréchet row scans  $\varepsilon$  over a 33-fold range in both directions  $H$  and  $-H$  at twelve values of  $\lambda$ ; the residual divided by  $\varepsilon^2$  stays within 2.1% of a constant, which is the signature of a genuine second-order remainder rather than of a fitted slope. The ensembles, seeds and tolerances behind every row of both tables are specified in Appendix C.

Table 2: Arithmetic checks behind the geometry of Sections 3–6. “Sample” is the number of independent inputs, “outcome” the worst residual or closest approach to a violation over that sample.

| statement                                             | method                                                  | sample                                         | outcome                                                                                                                                                                                                                                        |
|-------------------------------------------------------|---------------------------------------------------------|------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Lemma 4.3, pure-state PT spectrum                     | eigendecomposition vs. closed form                      | 201 values                                     | Schmidt max error $1.1 \times 10^{-16}$                                                                                                                                                                                                        |
| Lemma 4.4, max $W = K$ at $t_*$                       | dense grid, then exact evaluation                       | $2 \times 10^5$ points                         | grid max 0.60355183; exact value $K$ at $t_*$                                                                                                                                                                                                  |
| Theorem 4.5, $\ X^\Gamma\ _{\text{op}} \leq K\ X\ _1$ | random traceless Hermitian draws                        | 4000                                           | 0 violations; sup ratio 0.5544 = 0.9186 $K$                                                                                                                                                                                                    |
| Proposition 4.6, both extremisers                     | exact matrices, spectra                                 | 2                                              | $\ X_*^\Gamma\ _{\text{op}} = \ X_2^\Gamma\ _{\text{op}} = K$ to $10^{-16}$                                                                                                                                                                    |
| Thm. 5.1, Cor. 5.2                                    | random PPT/state pairs                                  | 1200                                           | 0 violations; min slack $2.74 \times 10^{-1}$                                                                                                                                                                                                  |
| one negative PT eigenvalue in $2 \otimes 2$           | same pairs                                              | 1200                                           | maximum number of negative eigenvalues seen: 1                                                                                                                                                                                                 |
| Theorem 5.1, Weyl-aligned saturation                  | ray $\rho_\lambda + sX_*$                               | 5 values of $\lambda$                          | residual $1.1 \times 10^{-16}$                                                                                                                                                                                                                 |
| Cor. 5.3, both radii                                  | random states in each ball                              | 1200 each                                      | min PT eigenvalue $9.2 \times 10^{-3}$ ( $\eta$ ), $8.0 \times 10^{-3}$ ( $\eta/K$ )                                                                                                                                                           |
| Remark 5.5, non-isometry and the false $2\eta$ ball   | explicit matrices                                       | 2                                              | ratio $1/\sqrt{2}$ ; witness at distance 0.4440346920 with $\lambda_{\min} = -9.7 \times 10^{-4}$                                                                                                                                              |
| Prop. 6.5, Thm. 6.6                                   | the ray at $\lambda \geq \lambda_c$                     | $9 + 1$                                        | PSD, boundary, distance = $\eta(\lambda)/K$ to $1.1 \times 10^{-16}$ ; below $\lambda_c$ , $\lambda_{\min} = -3.13 \times 10^{-2}$                                                                                                             |
| Proposition 6.9, spectra and distance                 | explicit family                                         | 41                                             | spectra to $1.7 \times 10^{-16}$ , distance = $2c$ to $8.3 \times 10^{-17}$                                                                                                                                                                    |
| Proposition 6.11, identity (32)                       | exact radical arithmetic (SymPy)                        | symbolic                                       | $\Delta(\lambda) \equiv 0$ ; $c_0$ closed form to $2.2 \times 10^{-16}$                                                                                                                                                                        |
| Proposition 6.11, $N_\lambda \geq d_{\text{up}}Q$     | grid over $(\lambda, c)$                                | $40 \times 40001$                              | minimum $1.3 \times 10^{-14}$ , attained at $c_0$                                                                                                                                                                                              |
| numerator identity for all witnesses                  | random unit vectors                                     | 2000                                           | max error $1.7 \times 10^{-16}$                                                                                                                                                                                                                |
| Lemma 6.2, relaxation $\inf_w D_w$                    | multistart minimisation                                 | $13\lambda \times 24$ starts                   | certified value on $[\lambda_c, \frac{1}{2})$ to $6.5 \times 10^{-15}$ ; $\eta(\lambda)/K$ below                                                                                                                                               |
| Thm. 6.14, App. A.4                                   | Gröbner reduction of (57); Schur criterion; exact roots | $550+2000+500$ vectors, 12 values of $\lambda$ | (58) to $2.8 \times 10^{-14}$ ; weights $(\frac{1}{2}, \frac{1}{2}, 0, 0)$ to $2.2 \times 10^{-13}$ ; the displayed $R \psi^-\rangle$ and its norm to $5.6 \times 10^{-16}$ ; sign of $F - 1$ matches $\lambda_{\min}(\sigma)$ in all 12 cases |
| Props. 6.12, 6.13, corner branches                    | explicit double corners                                 | $26 + 26$                                      | PSD, both corners, distance to $1.7 \times 10^{-16}$ ; $B^2 = 5.6 \times 10^{-16}$ at $\lambda_0$                                                                                                                                              |

## 9.2 The two checks that carry argumentative weight

Most rows of the two tables confirm a proof. Two of them are load bearing in a weaker sense, and the distinction matters.

**The witness relaxation on  $[\frac{1}{4}, \frac{1}{2})$ .** Lemma 6.2 is a theorem, but evaluating its right-hand side requires minimising  $D_w(\lambda)$  over the whole unit sphere of  $\mathbb{C}^4$ , which is a seven-real-parameter problem. Proposition 6.11 solves the restriction to the symmetric plane exactly. The multistart minimisation over the full sphere—24 starting points per  $\lambda$ , four of them structured ( $\psi_0$ ,  $|01\rangle$ ,  $|\phi^+\rangle$ ,  $|\psi^-\rangle$ ) and 20 drawn from a complex Gaussian—returned the symmetric-plane value at every  $\lambda$  sampled in  $[\lambda_c, \frac{1}{2})$ , to a relative accuracy of  $6.5 \times 10^{-15}$ , and returned no smaller value anywhere. The minimisers found for  $\lambda > \frac{1}{4}$  have Schmidt weights (0.792, 0.208) at  $\lambda = 0.30$ , (0.724, 0.276) at  $\lambda = 0.35$  and (0.651, 0.349) at  $\lambda = 0.40$ , drifting away from  $(t_*, 1 - t_*)$  as  $\lambda$  increases; all lie in the symmetric plane. This is evidence, not proof, and (33) is stated accordingly.

**The relaxation below  $\lambda_c$ .** The same minimisation returns exactly  $\eta(\lambda)/K$  for every  $\lambda \leq \frac{1}{4}$ , which is the content of Remark 6.4: on that half of the family the best witness is the fixed vector  $\psi_0$  and the relaxation adds nothing to Proposition 6.3. The certified upper bounds (36)–(34) lie strictly above it, by  $3.33 \times 10^{-2}$  at  $\lambda = 0$  and by  $9.81 \times 10^{-3}$  at  $\lambda_0$ , which is what Theorem 6.14 predicts and what no witness can detect: strictness there is a statement about equality candidates, and its proof in Appendix A.4 uses the Schur criterion rather than the relaxation. The band plotted in Figure 1 is therefore not an artefact of a weak numerical search: it is the gap between the best bound any single witness can give and the best boundary state that has been constructed.

## 9.3 Values used in the text

For convenience, the constants that appear in statements are collected here with the precision to which they were evaluated:

$$\begin{aligned} K &= 0.6035533905932737, & 1/K &= 1.6568542494923801, \\ t_* &= 0.8535533905932737, & \lambda_c &= 0.1213203435596426, \\ \lambda_0 &= 0.0729490168751577, & d(\lambda_c) &= 0.3431457505076198, \\ d(\tfrac{1}{4}) &= 0.4142135623730951, & \delta_{\text{low}}(0) &= 0.3094010767585030, \\ d_{\text{iso}}(0) &= 0.3333333333333333, & d_{\text{iso}}(\lambda_0) &= 0.3262379212492639. \end{aligned}$$

The first six are  $\frac{1+\sqrt{2}}{4}$ ,  $4(\sqrt{2}-1)$ ,  $\frac{2+\sqrt{2}}{4}$ ,  $\frac{3\sqrt{2}-4}{2}$ ,  $\frac{7-3\sqrt{5}}{4}$  and  $6-4\sqrt{2}$ ; the two distances at  $\lambda = 0$  are the limit value (38) and  $d_{\text{iso}}(0) = \frac{1}{3}$ . In nats the Werner value is  $F(\frac{3}{4}) = 0.1308120359411370 = 0.1887218755408672$  ebits and  $F(1) = \ln 2 = 0.6931471805599453$  nats = 1 ebit.

# 10 Discussion and open questions

## 10.1 What the four results say together

The four results of this paper are not independent. One number,  $K = (1+\sqrt{2})/4$ , is the whole input: it is the norm of the partial transpose as a map from the traceless trace-norm ball to Hermitian operators with the operator norm, and everything else is a consequence of that norm being finite, sharp and attained.

The continuity law (16) is the statement that the PPT test degrades gracefully: knowing a nearby PPT state and its gap  $g$  certifies separability within a radius  $g/K$ . The certified separable balls of Corollary 5.3 are the same statement applied to the Werner line, where the gap is known

Table 3: Arithmetic checks behind the entanglement results of Sections 7–8, in the same conventions as Table 2.

| statement                                                      | method                              | sample                  | outcome                                                                                              |
|----------------------------------------------------------------|-------------------------------------|-------------------------|------------------------------------------------------------------------------------------------------|
| Theorem 7.2, closed form                                       | exact arithmetic, then evaluation   | 13 values of $\lambda$  | identities exact; max error $1.4 \times 10^{-14}$                                                    |
| Theorem 7.1, minimiser over $G$                                | derivative-free global minimisation | 6                       | max error $9.7 \times 10^{-17}$ ; separable case returns $G^* = \lambda$                             |
| Lemma 3.4, $\lambda(\sigma) \leq \frac{1}{2}$ on $\mathcal{S}$ | random separable states             | 200                     | maximum singlet fraction 0.4815209                                                                   |
| Theorem 7.5, no challenger beats $F$                           | random separable and PPT states     | 1200 + 400              | minimal excess 0.2894 and 0.2879                                                                     |
| Theorem 7.5(ii), uniqueness                                    | second differences along directions | 160                     | minimum 0.7367 > 0                                                                                   |
| Theorem 7.5(iii), Bell family                                  | angles $\theta \in [0, \pi]$        | 37                      | value $\ln 2$ to $1.2 \times 10^{-13}$ ; 36 pairwise distinct                                        |
| (46), (45), inverse map                                        | operator construction               | 41 values of $x$        | operator error $3.3 \times 10^{-16}$ , value error $2.0 \times 10^{-14}$                             |
| Thm. 8.1, Prop. 8.2                                            | random two-qubit states             | 400                     | min. lower gap 0.3877; deficit identity to $2.3 \times 10^{-15}$ ; min. $\Delta$ off the line 0.1187 |
| Corollary 8.4                                                  | same states                         | 400                     | minimal slack 0.5162                                                                                 |
| Theorem 8.7                                                    | $\varepsilon$ -scan over directions | 12 pairs $(\lambda, H)$ | max  bound – linear / $\varepsilon^2 = 3.66$ ; width/ $\varepsilon^2$ drifts 2.1% over $33 \times$   |

in closed form, and the factor  $1/K = 1.6569$  is the entire content of the improvement over the elementary Weyl radius. The nearest-NPT distance  $d(\lambda)$  is the same quantity read in the opposite direction—how far one must move before the gap closes—and it is exactly  $\eta(\lambda)/K$  wherever the extremiser direction  $X_*$  stays inside the state space, which is precisely the interval  $[\lambda_c, \frac{1}{4}]$ . The coincidence of the separable-ball radius and the boundary distance on that interval is not an accident: both are the value of the same support function, and Proposition 6.5 exhibits the state that attains it.

The application to the relative entropy of entanglement uses none of this. It needs only the twirl and the PPT criterion, and it is classical. Its role in the paper is to provide a calibrated reference value: because  $F(\lambda)$  is known exactly and because the twirl cannot increase relative entropy, every two-qubit state inherits the lower bound  $F(\lambda(\rho))$ , and because the Werner state with the same singlet fraction is always a legal competitor, every state inherits the upper bound  $D(\rho||\rho_\mu)$ . The interesting content of Section 8 is that the difference between these two bounds is exactly the entropy deficit  $\Delta(\rho) = D(\rho||\rho_\lambda)$ : the sandwich is not a pair of unrelated estimates but a single identity plus a non-negative remainder, and that remainder measures the distance of  $\rho$  from the Werner line rather than its entanglement.

## 10.2 Where the geometry changes character

Three thresholds organise the boundary-distance picture, and each has a distinct cause.

At  $\lambda = \frac{1}{4}$  the  $\eta(\lambda)$ -eigenspace of  $\rho_\lambda^\Gamma$  collapses from the three-dimensional space spanned by  $|\phi^-\rangle$ ,  $|01\rangle$  and  $|10\rangle$  to the line  $\mathbb{C}|\phi^+\rangle$ . Since that line consists of maximally entangled vectors and the extremisers of (10) have Schmidt weight  $t_* \neq \frac{1}{2}$ , Weyl tightness and sharp-bound tightness become mutually exclusive above  $\frac{1}{4}$  (Proposition 6.7), and the optimal displacement leaves the Werner-symmetric subspace: the attaining direction changes from  $X_*$  to  $P_- - |11\rangle\langle 11|$ .

At  $\lambda = \lambda_c$  the ray  $\rho_\lambda + sX_*$  stops being a state. The obstruction is the  $|01\rangle, |10\rangle$  block, whose determinant is  $b^2(3 - 2\sqrt{2}) - e^2$ ; it changes sign exactly when  $e = b(\sqrt{2} - 1)$ , which is  $\lambda = \lambda_c$ . Below

that point the direction that saturates the operator inequality is still available—the  $\eta(\lambda)$ -eigenspace still contains vectors of Schmidt weight  $t_*$ —but moving the full distance  $\eta(\lambda)/K$  along it leaves the state space, and the certified upper bounds come from double-corner states of a different shape.

At  $\lambda = \lambda_0$  the two corner families exchange roles. The isospectral branch requires  $\mathcal{R}(\lambda) \geq 0$ , that is  $\lambda \leq \frac{1}{4}$ , and the rank-two branch requires  $B^2 = 1 - x - t \geq 0$ ; the second condition fails above  $\lambda_0 = \frac{7-3\sqrt{5}}{4}$ , where  $B^2 = 0$  and the two values agree to 0.3262379212. The exchange is a change of ansatz, not a change in  $d$ : whether  $d$  itself has a kink at  $\lambda_0$  is not known.

### 10.3 Limitations

The following are limits of what is established here, not of what is true.

*The all-state problem is untouched.* The question that motivated this line of work asks for a finite closed-form rule returning the closest separable state of an arbitrary two-qubit density matrix [8]. Nothing in this paper answers it. Theorem 7.5 classifies the closest separable states of the Werner family, which is a one-parameter set of measure zero in the 15-dimensional two-qubit state space, and Theorem 8.1 bounds  $E_R(\rho)$  for all  $\rho$  without identifying a minimiser. The inverse map (44) runs in the other direction: it produces states from a candidate minimiser, and using it as a direct rule would require inverting a map whose fibres are not classified.

*Below  $\lambda_c$  there are bounds, not values.* The exact value of  $d(\lambda)$  on  $(0, \lambda_c)$  is open, including the endpoint  $\lambda = 0$ , where (38) gives  $d(0) \leq 2/(3 + 2\sqrt{3})$  and nothing more. The gap between the Weyl lower bound and the best certified upper bound is  $3.33 \times 10^{-2}$  at  $\lambda = 0$ ,  $1.84 \times 10^{-2}$  at  $\lambda = 0.05$  and  $9.81 \times 10^{-3}$  at  $\lambda_0$ ; it closes at  $\lambda_c$ . That the lower edge is not attained on  $(0, \lambda_c)$  is Theorem 6.14, proved in Appendix A.4 by classifying the equality candidates and reducing their positivity to  $F(\lambda) \leq 1$ ; the analogous strictness on  $(\frac{1}{4}, \frac{1}{2})$  is Proposition 6.7, where the eigenspace argument is shorter because  $E_\eta$  collapses to  $\mathbb{C}|\phi^+\rangle$ .

*One reduction above  $\frac{1}{4}$  is taken on trust.* The exactness of  $d_{\text{up}}(\lambda)$  rests on the claim that the witness minimum in Lemma 6.2 is attained in the symmetric plane. The plane is analysed exactly, the numerator identity holds for every witness, and a full-sphere multistart search finds no counterexample, but the general reduction is not proved here.

*The applications, not the inequality, are two-dimensional.* Theorem A.2 shows that (10) holds in every dimension with the same constant  $K$ , so the operator estimate itself is not the limitation; its two uses are. PPT is not separability from  $3 \otimes 3$  onwards [2], and a partial transpose can have several negative eigenvalues there, so  $\mathcal{N}(\rho)$  is no longer  $[-\lambda_{\min}(\rho^\Gamma)]_+$  and the continuity law no longer certifies separability from a nearby PPT state. The distance computations of Section 6 use the explicit  $4 \times 4$  corner structure, the single-negative-eigenvalue property and the Werner twirl, none of which transfers. The sandwich of Section 8 uses the twirl and the PPT criterion as an equivalence, so it too is a two-qubit statement, although its lower edge  $F(\lambda(\rho))$  is meaningful wherever a Werner-type symmetry reduction is available.

*Floating point, not formal.* The certificates are double-precision evaluations of exact formulas, with residuals of order  $10^{-16}$  and margins of order  $10^{-2}$  to  $10^{-1}$  between the tested quantities and the thresholds they had to respect. No statement here has been checked by a proof assistant, and the two claims identified above rest on arguments this paper reports rather than re-derives.

### 10.4 Open questions

1. **The exact value of  $d(\lambda)$  on  $(0, \lambda_c)$ , and  $d(0)$ .** The candidates are explicit:  $\delta_{\text{low}}$  on  $(0, \lambda_0]$  and  $d_{\text{iso}}$  on  $[\lambda_0, \lambda_c]$ , and both are attained by double-corner states whose optimality inside

their own witness halfspace is certified by the source study's dual functionals at machine-precision Karush–Kuhn–Tucker residuals. What is missing is the global statement that no witness halfspace forces a larger distance. The endpoint  $\lambda = 0$  is separately interesting because  $\rho_0 = P_{\text{sym}}/3$  is rank three and the rank-two branch degenerates there, so (38) is a limit rather than an attained value.

2. **A singlet-fraction-only upper envelope for  $E_R$ .** The lower side of (47) depends on  $\lambda(\rho)$  alone. The upper side does not, and  $\sup\{E_R(\rho) : \text{Tr}[P_- \rho] = \lambda\}$  is not computed here, nor is the set of states attaining it, nor the set attaining only the lower bound. A closed envelope would make the sandwich a two-sided function of a single measurable quantity.
3. **The witness minimum as a function of  $\lambda$ .**  $\inf_w D_w(\lambda)$  is a well-defined, computable function that agrees with  $d(\lambda)$  on  $[\lambda_c, \frac{1}{2})$  and equals  $\eta(\lambda)/K$  below. Whether the two agree on  $(0, \lambda_c)$  is the same question as item 1 in dual language, and whether the minimiser is unique—it is  $\psi_0$  for  $\lambda \leq \frac{1}{4}$  by Remark 6.4, and lies in the symmetric plane above—bears on the structure of the closest boundary states.
4. **The equality case of the sharp inequality.** Remark 4.7 gives necessary conditions and two extremisers but no classification. A complete description of the traceless operators  $X$  with  $\|X^\Gamma\|_{\text{op}} = K\|X\|_1$  would settle, in particular, whether the ray of Proposition 6.5 is the only way to saturate the continuity law, and hence would decide item 1.

## 10.5 Conclusion

The partial transpose of a two-qubit operator, restricted to traceless operators and measured from the trace norm to the operator norm, has norm exactly  $(1 + \sqrt{2})/4$ . That single sharp constant makes the Peres–Horodecki criterion quantitative, upgrades every certified separable ball around a separable Werner state by the factor  $4(\sqrt{2} - 1)$ , replaces a doubled radius that is false, and gives the distance from  $\rho_\lambda$  to the nearest NPT state exactly on  $[\lambda_c, \frac{1}{2})$  with explicit certified bounds and a quantified gap below. Used in the other direction, the same Werner data yields a two-sided bound on the relative entropy of entanglement of every two-qubit state whose correction is an entropy deficit with a uniform modulus and a first-order expansion. The closed-form closest separable state of a general two-qubit state remains unknown, and this paper does not claim it.

## A Derivations

### A.1 The sharp bound is dimension free

The constant  $K$  was derived for two qubits. It is not a two-qubit constant.

**Lemma A.1** (spectrum of a partially transposed pure state in  $d \otimes d$ ). *Let  $d \geq 2$  and let  $\psi \in \mathbb{C}^d \otimes \mathbb{C}^d$  be a unit vector with Schmidt weights  $t_1 \geq t_2 \geq \dots \geq t_d \geq 0$ ,  $\sum_i t_i = 1$ . Then*

$$\text{spec}((|\psi\rangle\langle\psi|)^\Gamma) = \{t_1, \dots, t_d\} \cup \{+\sqrt{t_i t_j}, -\sqrt{t_i t_j} : 1 \leq i < j \leq d\},$$

and

$$\text{hw}((|\psi\rangle\langle\psi|)^\Gamma) = \frac{t_1 + \sqrt{t_1 t_2}}{2}. \quad (52)$$

*Proof.* Local unitaries reduce to  $\psi = \sum_i \sqrt{t_i} |ii\rangle$  and, by Lemma 3.1, conjugate  $(|\psi\rangle\langle\psi|)^\Gamma$  without changing its spectrum. In that form

$$(|\psi\rangle\langle\psi|)^\Gamma = \sum_{i,j} \sqrt{t_i t_j} |ij\rangle\langle ji|,$$

which is block diagonal: the  $d$  one-dimensional blocks  $|ii\rangle$  carry  $t_i$ , and for each pair  $i < j$  the two-dimensional block on  $\text{span}\{|ij\rangle, |ji\rangle\}$  is  $\sqrt{t_i t_j}$  times  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , with eigenvalues  $\pm\sqrt{t_i t_j}$ . The blocks account for  $d + 2\binom{d}{2} = d^2$  dimensions. For the half-width,  $\sqrt{t_i t_j} \leq t_1$  for every pair, so  $\lambda_{\max} = t_1$ , and the most negative eigenvalue is  $-\sqrt{t_1 t_2}$  because  $t_1 t_2 \geq t_i t_j$  for  $i < j$  other than  $(1, 2)$ .  $\square$

**Theorem A.2** (dimension-free sharp bound). *For every  $d \geq 2$  and every traceless Hermitian  $X$  on  $\mathbb{C}^d \otimes \mathbb{C}^d$ ,*

$$\|X^\Gamma\|_{\text{op}} \leq K \|X\|_1, \quad K = \frac{1 + \sqrt{2}}{4}, \quad (53)$$

*and the constant is attained for every  $d$ .*

*Proof.* The proof of Theorem 4.5 uses only Lemma 4.2, which is dimension free, and the half-width of  $(|\psi\rangle\langle\psi|)^\Gamma$ . By Lemma A.1 the latter is  $\frac{1}{2}(t_1 + \sqrt{t_1 t_2})$ , and

$$t_1 + \sqrt{t_1 t_2} \leq t_1 + \sqrt{t_1(1 - t_1)} \leq \frac{1 + \sqrt{2}}{2},$$

the first inequality because  $t_2 \leq 1 - t_1$  and the second being Lemma 4.4. Hence  $\|X^\Gamma\|_{\text{op}} \leq K \|X\|_1$ . For sharpness, embed the two-qubit extremiser  $X_*$  of Proposition 4.6 into  $\mathbb{C}^d \otimes \mathbb{C}^d$  on the subspace spanned by  $|0\rangle, |1\rangle$  of each factor: the embedded operator is traceless with trace norm 1, and its partial transpose is the embedding of  $X_*^\Gamma$  together with zeros, so its operator norm is  $\|X_*^\Gamma\|_{\text{op}} = K$ .  $\square$

The bound (53) was checked in  $d = 3$  and  $d = 4$  by comparing (52) with the numerically computed half-width of  $(|\psi\rangle\langle\psi|)^\Gamma$  for 2000 random unit vectors each (maximum deviation  $1.0 \times 10^{-15}$ ), by verifying the displayed spectrum exactly for  $(t_1, t_2, t_3) = (\frac{3}{5}, \frac{3}{10}, \frac{1}{10})$ , and by confirming that the embedded  $X_*$  has trace norm 1 and partial-transpose operator norm  $\frac{1}{4} + \frac{\sqrt{2}}{4} = K$  in  $d = 3, 4, 5$ .

*Remark A.3* (what does not extend). The inequality extends; its use does not. Section 5 converts (53) into a separability statement through  $\text{PPT} = \text{SEP}$ , which fails from  $3 \otimes 3$  onwards [2], and through  $\mathcal{N}(\rho) = [-\lambda_{\min}(\rho^\Gamma)]_+$ , which fails as soon as a partial transpose can have two negative eigenvalues—possible already in  $3 \otimes 3$  by Lemma A.1, whose spectrum contains  $-\sqrt{t_1 t_2}$  and  $-\sqrt{t_1 t_3}$  simultaneously. The distance computations of Section 6 use the explicit  $4 \times 4$  corner structure and do not transfer either. Only the operator inequality itself is dimension free.

## A.2 The discriminant identity behind Proposition 6.11

The proof of Proposition 6.11 reduces to the vanishing of

$$\Delta(\lambda) = \left(2A(B + d) + \frac{d^2}{2}\right)^2 - 4A^2((B + d)^2 + d^2),$$

where

$$A = \frac{1 + 2\lambda}{6}, \quad B = \frac{2(4\lambda - 1)}{3}, \quad d = \frac{2}{3}(2R - (2\lambda + 1)),$$

with  $R = R(\lambda) = \sqrt{(1-\lambda)(1+2\lambda)}$ . Expanding the square cancels the  $4A^2(B+d)^2$  terms and leaves

$$\Delta = d^2 \left[ \frac{d^2}{4} + 2A(B+d) - 4A^2 \right], \quad (54)$$

so it is enough to show that the bracket vanishes. Put  $u = 1 + 2\lambda$ , so that  $\lambda = \frac{u-1}{2}$ ,  $1 - \lambda = \frac{3-u}{2}$ ,  $4\lambda - 1 = 2u - 3$  and  $R^2 = (1-\lambda)(1+2\lambda) = \frac{u(3-u)}{2}$ ; then  $A = \frac{u}{6}$ ,  $B = \frac{2(2u-3)}{3}$ ,  $d = \frac{2}{3}(2R-u)$ , and multiplying the bracket by 9,

$$\begin{aligned} 9 \left[ \frac{d^2}{4} + 2A(B+d) - 4A^2 \right] &= (2R-u)^2 + 2u[(2u-3) + (2R-u)] - u^2 \\ &= 4R^2 - 4Ru + u^2 + 2u^2 - 6u + 4Ru - u^2 \\ &= 4R^2 + 2u^2 - 6u = 4 \cdot \frac{u(3-u)}{2} + 2u^2 - 6u = 0. \end{aligned}$$

The two  $\mp 4Ru$  terms cancel and the defining relation of  $R$  finishes the computation, so  $\Delta(\lambda) \equiv 0$  identically on  $[\frac{1}{4}, \frac{1}{2}]$ . Consequently  $q(c) = \kappa c^2 + \beta c + A^2$  with  $\kappa = (B+d)^2 + d^2 > 0$  and  $\beta = -2A(B+d) - \frac{d^2}{2}$  is a perfect square,  $q(c) = \kappa(c - c_0)^2$  with  $c_0 = -\beta/(2\kappa)$ , and substituting the closed forms of  $A$ ,  $B$ ,  $d$  and  $R$  gives (31). Three identities were verified in exact radical arithmetic, writing  $R$  as  $\sqrt{1-\lambda}\sqrt{2\lambda+1}$ :  $\Delta \equiv 0$ ,  $c_0 + \beta/(2\kappa) \equiv 0$ , and  $q'(c_0) \equiv 0$ ; together they give  $q(c_0) = -\Delta/(4\kappa) = 0$ . In floating point the residuals over 40 values of  $\lambda$  in  $[\frac{1}{4}, \frac{1}{2}]$  are  $8.8 \times 10^{-17}$  for  $q(c_0)$  and  $2.8 \times 10^{-16}$  for  $c_0 + \beta/(2\kappa)$ .

### A.3 The feasibility function of the rank-two branch

Proposition 6.13 is usable exactly where  $B^2 = 1 - x - t \geq 0$ , since  $B$  is the  $|\phi^+\rangle$  coefficient of the corner vector. With  $s = \sqrt{3(2\lambda+1)}$  and the parameters (35), clearing denominators gives

$$B^2(\lambda) = \frac{N(\lambda)}{D(\lambda)}, \quad \begin{aligned} N(\lambda) &= 12\lambda^3 + 12\sqrt{3}\lambda^2\sqrt{2\lambda+1} + 10\lambda^2 - 6\sqrt{3}\lambda\sqrt{2\lambda+1} - 5\lambda + 1, \\ D(\lambda) &= 2(1-4\lambda)(1-\lambda)(3\lambda+1), \end{aligned} \quad (55)$$

an identity verified in exact arithmetic. On  $(0, \frac{1}{4})$  the denominator is positive, so the sign of  $B^2$  is the sign of  $N$ . The number  $\lambda_0 = \frac{7-3\sqrt{5}}{4}$  is a root of  $N$ : substituting it clears the radicals and returns exactly zero. The remaining positive root of  $N$  is  $\frac{1+\sqrt{13}}{12} = 0.3837959396$ , outside  $(0, \frac{1}{4}]$ . Since  $N(0) = 1 > 0$ ,  $N$  is positive on  $(0, \lambda_0)$  and negative on  $(\lambda_0, \frac{1}{4})$ , so (36) is available exactly on  $(0, \lambda_0]$  and the branch terminates there with  $B = 0$ , meaning the corner vector has lost its  $|\phi^+\rangle$  component. At that endpoint  $x + t = 1$ , and  $\delta_{\text{low}}(\lambda_0) = d_{\text{iso}}(\lambda_0)$  exactly, which is the crossover claimed in Proposition 6.13: the rank-two corner vector becomes  $\sqrt{x}|\psi^-\rangle + \sqrt{t}|\psi^+\rangle$ , and the isospectral family at  $\lambda_0$  has  $\mathcal{R}(\lambda_0) = t(\lambda_0)$ , so the two constructions agree. The limits at  $\lambda \downarrow 0$  are  $x \rightarrow 0$ ,  $t \rightarrow \frac{1}{2}$ ,  $B^2 \rightarrow \frac{1}{2}$  and  $c_{\text{low}} \rightarrow \frac{1}{3+2\sqrt{3}}$ , all exact, which gives the state  $\sigma_0$  of (38).

### A.4 No equality candidate exists below $\lambda_c$

This appendix proves Theorem 6.14, quoted in Section 6: the Weyl lower bound is *strict* on  $(0, \lambda_c)$ , so  $d(\lambda) > \eta(\lambda)/K$  there. The argument is a chain of four exact steps—classify the candidates, prove two structural identities for the vectors they are built from, evaluate a  $2 \times 2$  Gram matrix, and apply a Schur complement—and it ends in a quadratic equation with radical roots.

**Lemma A.4** (the equality candidates). *Fix  $0 < \lambda \leq \frac{1}{4}$  and suppose that  $d(\lambda) = \eta(\lambda)/K$ . In this range the eigenspace  $E_\eta = \ker(\rho_\lambda^\Gamma - \eta(\lambda)I)$  is  $\{|\phi^+\rangle\}^\perp = \text{span}\{|\phi^-\rangle, |01\rangle, |10\rangle\}$  by (6). Then*

there are a unit vector  $\psi \in E_\eta$  of Schmidt weight  $t_*$ , unit vectors  $u_+, u_-$  spanning the  $\lambda_{\max}$ - and  $\lambda_{\min}$ -eigenspaces of  $R := (|\psi\rangle\langle\psi|)^\Gamma$ , and the number

$$c = \frac{\eta(\lambda)}{2K} = \frac{1 + 2\lambda}{3(1 + \sqrt{2})}, \quad (56)$$

such that  $\sigma = \rho_\lambda + c(|u_-\rangle\langle u_-| - |u_+\rangle\langle u_+|)$  is a state with  $\lambda_{\min}(\sigma^\Gamma) = 0$  and  $\|\sigma - \rho_\lambda\|_1 = \eta(\lambda)/K$ .

*Proof.* The state space is compact and  $\{\tau : \lambda_{\min}(\tau^\Gamma) \leq 0\}$  is closed, so a sequence of NPT states whose distances to  $\rho_\lambda$  tend to  $d(\lambda)$  has a limit point  $\sigma$  with  $\|\sigma - \rho_\lambda\|_1 = \eta(\lambda)/K$  and  $\lambda_{\min}(\sigma^\Gamma) \leq 0$ . Theorem 5.1 with the reference  $\rho_\lambda$  and  $g = \eta(\lambda)$  gives  $\lambda_{\min}(\sigma^\Gamma) \geq \eta(\lambda) - K \cdot \eta(\lambda)/K = 0$ , so  $\lambda_{\min}(\sigma^\Gamma) = 0$  and both inequalities used are equalities. Put  $X = \sigma - \rho_\lambda$  and  $s = \|X\|_1 = \eta(\lambda)/K$ .

Equality in Weyl's inequality  $\lambda_{\min}(\sigma^\Gamma) \geq \lambda_{\min}(\rho_\lambda^\Gamma) + \lambda_{\min}(X^\Gamma)$  means that one unit vector  $\psi$  is simultaneously an eigenvector of  $\rho_\lambda^\Gamma$  for  $\eta(\lambda)$ , of  $X^\Gamma$  for  $\lambda_{\min}(X^\Gamma)$ , and of  $\sigma^\Gamma$  for 0. The first of these says  $\psi \in E_\eta$ . Together they give  $\lambda_{\min}(X^\Gamma) = \langle\psi|X^\Gamma|\psi\rangle - \eta(\lambda) = -Ks$ , and since  $\text{Tr}[X^\Gamma|\psi\rangle\langle\psi|] = \text{Tr}[X(|\psi\rangle\langle\psi|)^\Gamma] = \text{Tr}[XR]$ , the operator  $X$  attains  $-K\|X\|_1$  in the pairing of Lemma 4.2 against  $R$ . Equality there forces  $\text{hw}(R) = K$  and  $X = \frac{s}{2}(P_- - P_+)$ , where  $P_\pm$  are the eigenprojectors of  $R$  for  $\lambda_{\max}(R)$  and  $\lambda_{\min}(R)$ ; this is the equality case recorded in Remark 4.7. Because  $\text{hw}(R) = K$ , Lemma 4.4 gives the Schmidt weights  $(t_*, 1 - t_*)$ , so  $\text{spec } R = \{t_*, 1 - t_*, \pm \frac{1}{2\sqrt{2}}\}$  consists of four distinct numbers by Lemma 4.3 and  $P_\pm = |u_\pm\rangle\langle u_\pm|$  are rank one. Substituting  $s = \eta(\lambda)/K$  gives  $c = s/2$  as in (56), and  $\sigma = \rho_\lambda + X$ .  $\square$

**Lemma A.5** (two structural identities). *Let  $\psi$  be a unit vector with  $\langle\phi^+|\psi\rangle = 0$ , let  $R = (|\psi\rangle\langle\psi|)^\Gamma$ , and let  $t_1 \geq t_2$  be the Schmidt weights of  $\psi$  with  $\Delta = t_1 t_2$ . Then*

$$\langle\psi^-|R^k|\psi^-\rangle = \frac{1}{2}(t_1^k + t_2^k), \quad k = 0, 1, 2, 3, \quad (57)$$

and if the four eigenvalues  $t_1, t_2, \pm\sqrt{\Delta}$  are distinct then  $\psi^-$  has spectral weight  $\frac{1}{2}$  on each of the  $t_1$ - and  $t_2$ -eigenspaces and weight 0 on both  $\pm\sqrt{\Delta}$ -eigenspaces. In particular, for  $\psi \in E_\eta$  of Schmidt weight  $t_*$ ,

$$\langle u_-|\psi^-\rangle = 0, \quad |\langle u_+|\psi^-\rangle|^2 = \frac{1}{2}. \quad (58)$$

*Proof.* Write  $\psi = p|00\rangle + q|01\rangle + s|10\rangle - p|11\rangle$  with  $p \geq 0$  real and  $q, s \in \mathbb{C}$ : the condition  $\langle\phi^+|\psi\rangle = 0$  forces the  $|11\rangle$  coefficient to be minus the  $|00\rangle$  coefficient, and a global phase, which leaves  $R$  unchanged, makes  $p$  real. Put  $Q = |q|^2$ ,  $S = |s|^2$ , so  $2p^2 + Q + S = 1$ , and let  $\Psi = \begin{pmatrix} p & q \\ s & -p \end{pmatrix}$  be the coefficient matrix of  $\psi$ , so that  $\det \Psi = -p^2 - qs$  and hence  $\Delta = |\det \Psi|^2 = p^4 + 2p^2 \text{Re}(qs) + QS$  by Lemma 4.3, which also gives  $t_1 + t_2 = 1$ .

For  $k = 0$  both sides are 1. For  $k = 1$ , the matrix entries of  $R$  give  $\langle\psi^-|R|\psi^-\rangle = \frac{1}{2}(R_{01,01} + R_{10,10} - R_{01,10} - R_{10,01}) = p^2 + \frac{1}{2}(Q + S) = p^2 + \frac{1}{2}(1 - 2p^2) = \frac{1}{2}$ , while  $\frac{1}{2}(t_1 + t_2) = \frac{1}{2}$ . For  $k = 2$ ,  $R|\psi^-\rangle = \frac{1}{\sqrt{2}}(p(\bar{q} - s), p^2 + Q, -(p^2 + S), p(\bar{s} - q))$ , so

$$\begin{aligned} \langle\psi^-|R^2|\psi^-\rangle &= \|R\psi^-\|_1^2 = p^2(Q + S - 2\text{Re}(qs)) + \frac{1}{2}[(p^2 + Q)^2 + (p^2 + S)^2] \\ &= p^4 + 2p^2(Q + S) - 2p^2 \text{Re}(qs) + \frac{1}{2}(Q + S)^2 = \frac{1}{2} - \Delta, \end{aligned}$$

where the last step substitutes  $Q + S = 1 - 2p^2$  and cancels; and  $\frac{1}{2}(t_1^2 + t_2^2) = \frac{1}{2}((t_1 + t_2)^2 - 2\Delta) = \frac{1}{2} - \Delta$ . For  $k = 3$  the same expansion produces a degree-six polynomial in  $(p, q_1, q_2, s_1, s_2)$ ; reduced modulo  $2p^2 + Q + S - 1$  it equals  $\frac{1}{2}(1 - 3\Delta) = \frac{1}{2}(t_1^3 + t_2^3)$ , since  $t_1^3 + t_2^3 = (t_1 + t_2)^3 - 3\Delta(t_1 + t_2)$ . The reduction is an exact Gröbner-basis computation in five real variables with rational coefficients, and it is recorded among the identities of Section 9.

Let  $w_1, w_2, w_+, w_- \geq 0$  be the spectral weights of  $\psi^-$  on the four eigenspaces, and write  $u = w_+ + w_-$ ,  $v = w_+ - w_-$ . The identities (57) for  $k = 0, 1, 2$  are the three equations

$$w_1 + w_2 + u = 1, \quad w_1 t_1 + w_2 t_2 + v\sqrt{\Delta} = \frac{1}{2}, \quad w_1 t_1^2 + w_2 t_2^2 + u\Delta = \frac{1}{2}(t_1^2 + t_2^2),$$

whose solutions are  $w_1 = \frac{1}{2} - ut_2$ ,  $w_2 = \frac{1}{2} - ut_1$  and  $v = 2u\sqrt{\Delta}$ . The  $k = 3$  equation then reads  $\frac{1}{2}(t_1^3 + t_2^3) + u\Delta(4\Delta - 1) = \frac{1}{2}(t_1^3 + t_2^3)$ , so  $u\Delta(4\Delta - 1) = 0$ . For  $\psi$  of Schmidt weight  $t_*$  one has  $\Delta = t_*(1 - t_*) = \frac{1}{8}$ , whence  $4\Delta - 1 = -\frac{1}{2} \neq 0$  and  $u = 0$ ; then  $v = 0$ ,  $w_1 = w_2 = \frac{1}{2}$ , and  $w_+ = w_- = 0$ . The claims (58) are  $w_- = |\langle u_- | \psi^- \rangle|^2 = 0$  and  $w_1 = |\langle u_+ | \psi^- \rangle|^2 = \frac{1}{2}$ .  $\square$

**Lemma A.6** (the Gram matrix of  $\rho_\lambda^{-1}$ ). *For  $0 < \lambda < 1$ ,  $\rho_\lambda^{-1} = \frac{3}{1-\lambda}I + (\frac{1}{\lambda} - \frac{3}{1-\lambda})P_-$ . If  $u_+, u_-$  satisfy (58), then*

$$A := \langle u_+ | \rho_\lambda^{-1} | u_+ \rangle = \frac{1+2\lambda}{2\lambda(1-\lambda)}, \quad B := \langle u_- | \rho_\lambda^{-1} | u_- \rangle = \frac{3}{1-\lambda}, \quad \langle u_+ | \rho_\lambda^{-1} | u_- \rangle = 0. \quad (59)$$

*Proof.*  $\rho_\lambda = \frac{1-\lambda}{3}(I - P_-) + \lambda P_-$  is diagonal in the decomposition  $\mathbb{C}^4 = \text{ran } P_- \oplus \text{ran}(I - P_-)$ , with eigenvalues  $\lambda$  and  $\frac{1-\lambda}{3}$ , which gives the inverse. Substituting  $P_- = |\psi^- \rangle \langle \psi^-|$  and using (58),  $A = \frac{3}{1-\lambda} + (\frac{1}{\lambda} - \frac{3}{1-\lambda})\frac{1}{2} = \frac{1+2\lambda}{2\lambda(1-\lambda)}$ ,  $B = \frac{3}{1-\lambda} + 0$ , and  $\langle u_+ | \rho_\lambda^{-1} | u_- \rangle = \frac{3}{1-\lambda} \langle u_+ | u_- \rangle + (\frac{1}{\lambda} - \frac{3}{1-\lambda}) \langle u_+ | \psi^- \rangle \langle \psi^- | u_- \rangle = 0$ .  $\square$

*Proof of Theorem 6.14.* Suppose equality holds. Lemma A.4 produces a candidate  $\sigma = \rho_\lambda + UDU^\dagger$  with  $U = [u_+ \ u_-]$  having orthonormal columns and  $D = \text{diag}(-c, c)$ ,  $c > 0$ . Since  $\rho_\lambda > 0$  for  $\lambda > 0$ , the Schur complement criterion gives

$$\sigma \geq 0 \iff I + GDG^\dagger \geq 0 \text{ with } G = \rho_\lambda^{-1/2}U \iff D^{-1} + U^\dagger \rho_\lambda^{-1}U \geq 0,$$

the second equivalence holding because  $GDG^\dagger$  and  $D^{1/2}G^\dagger GD^{1/2}$  have the same non-zero spectrum and congruence by  $D^{-1/2}$  preserves signs. By Lemma A.6 the last matrix is  $\text{diag}(-\frac{1}{c} + A, \frac{1}{c} + B)$ , and  $\frac{1}{c} + B > 0$ , so

$$\sigma \geq 0 \iff A \leq \frac{1}{c} \iff \frac{1+2\lambda}{2\lambda(1-\lambda)} \leq \frac{3(1+\sqrt{2})}{1+2\lambda} \iff F(\lambda) \leq 1,$$

with

$$F(\lambda) = \frac{(1+2\lambda)^2}{6(1+\sqrt{2})\lambda(1-\lambda)}. \quad (60)$$

On  $(0, \frac{1}{2})$  the denominator is positive, so  $F(\lambda) \leq 1$  is the quadratic inequality  $(10 + 6\sqrt{2})\lambda^2 - (2 + 6\sqrt{2})\lambda + 1 \leq 0$ , whose two roots are

$$\lambda_c = \frac{3\sqrt{2}-4}{2} = 0.1213203435596426\dots, \quad \frac{2+3\sqrt{2}}{14} = \frac{1}{7} + \frac{3\sqrt{2}}{14} = 0.4459029062228061\dots,$$

both in  $(0, \frac{1}{2})$ , and both exact: substituting either clears the radicals. The leading coefficient is positive, so  $F(\lambda) > 1$  for  $0 < \lambda < \lambda_c$ . Every equality candidate is therefore indefinite there, contradicting Lemma A.4, and  $d(\lambda) > \eta(\lambda)/K$ .  $\square$

*Remark A.7.* Three points of scope. First, the criterion  $F(\lambda) \leq 1$  is an *equivalence*, so it also explains why the ray of Proposition 6.5 becomes feasible exactly at  $\lambda_c$ . Its second root,  $\frac{1}{7} + \frac{3\sqrt{2}}{14} = 0.4459\dots$ , lies above  $\frac{1}{4}$ , where  $E_\eta$  has collapsed to  $\mathbb{C}|\phi^+\rangle$  and vectors of  $\{|\phi^+\rangle\}^\perp$  stop being equality

candidates altogether; Proposition 6.7 rules out equality there by that collapse, so the upper root has no bearing on  $d(\lambda)$ . Second, nothing is claimed at  $\lambda = 0$ , where  $\rho_\lambda$  is not invertible and Lemma A.6 does not apply; the value  $d(0)$  is left open in Section 6. Third, the identities (57) hold on all of  $\{|\phi^+\rangle\}^\perp$ , not merely on the Schmidt- $t_*$  part, so the same computation applies verbatim to every equality candidate in the whole range  $\lambda \leq \frac{1}{4}$ .

## B The explicit boundary states

This appendix supplies the computations that Section 6 quotes: the two double-corner families below  $\lambda_c$ , the limit state at  $\lambda = 0$ , the Bell family at  $\lambda = 1$ , and the state that refutes the doubled separable ball. In every case the object is an explicit  $4 \times 4$  matrix and the claims about it are identities in  $\lambda$  that were checked in exact rational and radical arithmetic, not only in floating point.

Throughout,  $\Gamma$  denotes the first-factor partial transpose, taken in the computational basis, so that  $X_{(a,b),(c,e)}^\Gamma = X_{(c,b),(a,e)}$ , and

$$\begin{aligned}\rho_\lambda &= \lambda P_- + \frac{1-\lambda}{3} P_{\text{sym}} \\ &= a(|00\rangle\langle 00| + |11\rangle\langle 11|) + b(|01\rangle\langle 01| + |10\rangle\langle 10|) + e(|01\rangle\langle 10| + |10\rangle\langle 01|),\end{aligned}$$

with  $a = \frac{1-\lambda}{3}$ ,  $b = \frac{1+2\lambda}{6}$ ,  $e = \frac{1-4\lambda}{6}$  as in Section 3.

### B.1 The isospectral corner family

The ansatz is a rank-one displacement inside the three-dimensional subspace  $\text{span}\{|\phi^-\rangle, |\psi^+\rangle, |\psi^-\rangle\}$ :

$$\sigma = \rho_\lambda + c(|\phi^-\rangle\langle\phi^-| - |q\rangle\langle q|), \quad q = \sqrt{1-\mathcal{R}}|\psi^-\rangle - \sqrt{\mathcal{R}}|\psi^+\rangle, \quad 0 \leq \mathcal{R} \leq 1, \quad c > 0, \quad (61)$$

so that  $\text{Tr } \sigma = 1$  automatically. Writing  $\mathcal{R}$  and  $c$  as independent unknowns and expanding, the two determinants factor completely:

$$\det \sigma = \frac{1}{27} (\lambda - 1)(3c - \lambda + 1) [c((1 - \lambda) - \mathcal{R}(1 - 4\lambda)) - \lambda(1 - \lambda)], \quad (62)$$

$$\begin{aligned}\det \sigma^\Gamma &= -\frac{1}{432} (6\mathcal{R}c + 2\lambda + 1)(2\mathcal{R}c - 2c + 2\lambda - 1) \\ &\quad \times [36\mathcal{R}^2c^2 - 36\mathcal{R}c^2 - 12c\lambda - 6c + 4\lambda^2 + 4\lambda + 1].\end{aligned} \quad (63)$$

For  $\lambda \in (0, \frac{1}{4})$  the factors  $\lambda - 1$  and  $3c - \lambda + 1$  cannot vanish at a feasible point, since the second would require  $c = \frac{\lambda-1}{3} < 0$ ; likewise  $6\mathcal{R}c + 2\lambda + 1 > 0$  and  $2\mathcal{R}c - 2c + 2\lambda - 1 = 0$  would give  $c = \frac{1-2\lambda}{2(\mathcal{R}-1)} \leq 0$ . The double-corner condition therefore reduces to the two bracketed factors,

$$c((1 - \lambda) - \mathcal{R}(1 - 4\lambda)) = \lambda(1 - \lambda), \quad 36\mathcal{R}^2c^2 - 36\mathcal{R}c^2 - 12c\lambda - 6c + 4\lambda^2 + 4\lambda + 1 = 0, \quad (64)$$

a linear and a quadratic equation in the two unknowns. Eliminating  $c$  gives a quadratic whose two roots produce exactly two solutions of (64):

$$(\mathcal{R}, c) = \left( \frac{(1+2\lambda)(1-\lambda)}{10\lambda^2 - 2\lambda + 1}, \frac{10\lambda^2 - 2\lambda + 1}{18\lambda} \right) \quad \text{and} \quad (\mathcal{R}, c) = \left( \frac{(1-\lambda)(1-4\lambda)}{10\lambda^2 - 2\lambda + 1}, \frac{10\lambda^2 - 2\lambda + 1}{6(1-\lambda)} \right).$$

The first is infeasible: its  $\mathcal{R}$  exceeds 1 on  $(0, \frac{1}{4})$  because

$$\mathcal{R} - 1 = \frac{3\lambda(1-4\lambda)}{10\lambda^2 - 2\lambda + 1} > 0 \quad (10\lambda^2 - 2\lambda + 1 > 0 \text{ for all } \lambda),$$

and  $\mathcal{R} > 1$  makes  $q$  in (61) undefined. The second is exactly the pair  $(\mathcal{R}(\lambda), c_{\text{iso}}(\lambda))$  of Proposition 6.12, so the feasible double-corner state in this ansatz is unique.

Substituting it back,  $\sigma_{\text{iso}}$  has the four eigenvalues

$$0, \quad \frac{1-\lambda}{3}, \quad \frac{1-\lambda}{3} + c_{\text{iso}} = \frac{4\lambda^2 - 2\lambda + 1}{2(1-\lambda)}, \quad \frac{1+4\lambda-14\lambda^2}{6(1-\lambda)}, \quad (65)$$

and  $\sigma_{\text{iso}}^\Gamma$  has the same four: each of them is a root of both  $\det(\sigma_{\text{iso}} - rI)$  and  $\det(\sigma_{\text{iso}}^\Gamma - rI)$ , an identity in  $\lambda$  verified exactly, and they are distinct on  $(0, \frac{1}{4})$ , so both characteristic polynomials split completely over them. That is the *isospectral* property. All four are non-negative on  $[0, \frac{1}{4}]$ :  $\mathcal{R} \in [0, 1]$  there,  $c_{\text{iso}} > 0$ , and the last eigenvalue is positive because  $1+4\lambda-14\lambda^2 > 0$  for  $\lambda \leq \frac{1}{4}$ . At  $\lambda = \frac{1}{4}$  the spectrum degenerates to  $\{0, \frac{1}{4}, \frac{1}{4}, \frac{1}{2}\}$ , which is  $\text{spec}(\sigma_*)$ : the isospectral corner state at  $\lambda = \frac{1}{4}$  is  $\sigma_*$  itself, and  $c_{\text{iso}}(\frac{1}{4}) = \frac{1}{4}$  gives  $d_{\text{iso}}(\frac{1}{4}) = \frac{1}{2} = 1 - 2\lambda$ , the Werner competitor. That coincidence is why the band in Figure 1 closes smoothly.

Since  $\sigma_{\text{iso}} - \rho_\lambda = c_{\text{iso}}(|\phi^-\rangle\langle\phi^-| - |q\rangle\langle q|)$  is a difference of two orthogonal rank-one projectors times  $c_{\text{iso}}$ —orthogonal because  $\langle\phi^-|q\rangle = 0$ —its spectrum is  $\{c_{\text{iso}}, -c_{\text{iso}}, 0, 0\}$  and  $\|\sigma_{\text{iso}} - \rho_\lambda\|_1 = 2c_{\text{iso}} = d_{\text{iso}}(\lambda)$ , verified exactly at  $\lambda = \frac{1}{20}, \frac{1}{10}, \frac{3}{20}, \frac{1}{4}$ , where the values are  $\frac{37}{114}, \frac{1}{3}, \frac{37}{102}, \frac{1}{2}$ .

## B.2 The rank-two corner family

The second ansatz keeps three Bell directions:

$$\sigma = \rho_\lambda + c(|\phi^-\rangle\langle\phi^-| - |q\rangle\langle q|), \quad q = \sqrt{x}|\psi^-\rangle + \sqrt{1-x-t}|\phi^+\rangle + \sqrt{t}|\psi^+\rangle, \quad (66)$$

with  $x, t \geq 0$ ,  $x+t \leq 1$ ,  $c > 0$ . Set  $A = \sqrt{x}$ ,  $B = \sqrt{1-x-t}$ ,  $D = \sqrt{t}$ , so that  $A^2 + B^2 + D^2 = 1$ . Expanding  $\det \sigma$  as a polynomial in  $(\lambda, A, B, D, c)$  gives the complete factorisation

$$\det \sigma = -\frac{1}{27}(\lambda-1)(3c-\lambda+1)[A^2c\lambda - A^2c - 3B^2c\lambda - 3D^2c\lambda - \lambda^2 + \lambda], \quad (67)$$

and on the sphere  $B^2 = 1 - A^2 - D^2$  the bracket becomes  $c[3\lambda + (1-4\lambda)x] - \lambda(1-\lambda)$ . Hence the first corner condition is exactly

$$c = c_{\text{psd}}(\lambda, x) = \frac{\lambda(1-\lambda)}{3\lambda + (1-4\lambda)x}, \quad (68)$$

the same relation the source study uses, and it depends on  $x$  but not on  $t$ .

The second corner condition is more interesting:  $\det \sigma^\Gamma$ , which a priori is a polynomial in  $A, B, D$  as well, depends on the corner vector only through  $x = A^2$  and  $t = D^2$ , and equals a quartic in  $c$ ,

$$\det \sigma^\Gamma = P(c) = a_4c^4 + a_3c^3 + a_2c^2 + a_1c + a_0, \quad (69)$$

with

$$\begin{aligned} a_4 &= -t^2(t-1)^2, \\ a_3 &= \frac{1}{3}t(t-1)(4\lambda x - 2\lambda - x + 2), \\ a_2 &= \frac{1}{36}(2\lambda+1)(8\lambda t^2 + 16\lambda tx - 8\lambda t - 8\lambda x + 6\lambda - 2t^2 - 4tx + 2t + 2x - 3), \\ a_1 &= \frac{1}{108}x(2\lambda+1)^2(4\lambda-1), \\ a_0 &= \frac{1}{432}(1-2\lambda)(2\lambda+1)^3. \end{aligned} \quad (70)$$

The identity (69) was verified in exact rational arithmetic at six rational parameter points  $(\lambda, x, t, c)$  with  $x + t < 1$ , and in floating point at random points, where the ratio  $\det \sigma^\Gamma / P(c)$  equals 1 to all printed digits. The branch of Proposition 6.13 is the solution of the three equations

$$c = c_{\text{psd}}(\lambda, x), \quad P(c) = 0, \quad \frac{\partial P}{\partial t}(\lambda, x, t, c) = 0, \quad (71)$$

the last being the tangency condition that makes the corner double in the  $t$ -direction. Substituting the closed forms (35) satisfies all three identically in  $\lambda$ —verified exactly, with the radicals  $s = \sqrt{3(2\lambda + 1)}$  left unevaluated—and gives  $\lambda_{\min} \sigma = \lambda_{\min} \sigma^\Gamma = 0$  with the remaining eigenvalues positive: at  $\lambda = \frac{1}{100}, \frac{1}{50}, \frac{1}{20}$  the partial-transpose spectra are

$$\begin{aligned} \lambda = \frac{1}{100} : & \{0, 0.239146, 0.253095, 0.507759\}, \\ \lambda = \frac{1}{50} : & \{0, 0.234755, 0.262117, 0.503127\}, \\ \lambda = \frac{1}{20} : & \{0, 0.224277, 0.288863, 0.486860\}, \end{aligned}$$

with the zero eigenvalue exact rather than numerical, and the floating-point margins of the other eigenvalues are of order  $10^{-17}$ . The distance is again  $2c$ :  $\sigma - \rho_\lambda = c(|\phi^-\rangle\langle\phi^-| - |q\rangle\langle q|)$  with  $\langle\phi^-|q\rangle = 0$ , so  $\|\sigma - \rho_\lambda\|_1 = 2c$ , which was verified exactly at the same three values of  $\lambda$ , where  $2c_{\text{low}}$  equals  $\frac{1}{50}\sqrt{23801 - 4040\sqrt{34}}, \frac{1}{75}\sqrt{54609 - 6120\sqrt{78}}$  and  $\frac{1}{30}\sqrt{9249 - 504\sqrt{330}}$  respectively.

### B.3 The limit state at $\lambda = 0$

Taking  $\lambda \downarrow 0$  in (35) gives, exactly,  $x \rightarrow 0$ ,  $t \rightarrow \frac{1}{2}$ ,  $B^2 \rightarrow \frac{1}{2}$  and  $c_{\text{low}} \rightarrow \frac{1}{3+2\sqrt{3}}$ , so  $q \rightarrow q_0 = (|\phi^+\rangle + |\psi^+\rangle)/\sqrt{2}$  and

$$\sigma_0 = \rho_0 + \frac{1}{3+2\sqrt{3}}(|\phi^-\rangle\langle\phi^-| - |q_0\rangle\langle q_0|), \quad \rho_0 = \frac{1}{3}P_{\text{sym}}.$$

Then  $\det \sigma_0 = 0$  exactly,  $\det \sigma_0^\Gamma = 0$  to  $1.5 \times 10^{-16}$ , and

$$\begin{aligned} \text{spec}(\sigma_0) &= \{0, 0.1786327950, 0.3333333333, 0.4880338717\}, \\ \text{spec}(\sigma_0^\Gamma) &= \{0, 0.2440169359, 0.2440169359, 0.5119661283\}, \end{aligned}$$

so  $\sigma_0$  is a state on the PPT boundary. Its distance to  $\rho_0$  is  $2c = 2/(3+2\sqrt{3}) = 0.3094010767585031$ , which is (38). Because  $\rho_0$  is rank three the branch does not include  $\lambda = 0$  itself, and (38) is a bound obtained as a limit of feasible states; the value  $d(0)$  remains open.

### B.4 The Bell family at $\lambda = 1$

For  $\theta \in [0, \pi)$  let  $u_\theta = \cos \theta |\psi^+\rangle + \sin \theta |\phi^-\rangle$  and  $\sigma_\theta = \frac{1}{2}P_- + \frac{1}{2}|u_\theta\rangle\langle u_\theta|$ , as in (43). Since  $\langle\psi^-|u_\theta\rangle = 0$ , the two summands are orthogonal and  $\sigma_\theta^2 = \frac{1}{4}P_- + \frac{1}{4}|u_\theta\rangle\langle u_\theta| = \frac{1}{2}\sigma_\theta$ , an identity in  $\theta$  verified exactly; a positive semidefinite operator satisfying  $M^2 = \frac{1}{2}M$  with  $\text{Tr } M = 1$  has spectrum  $\{0, 0, \frac{1}{2}, \frac{1}{2}\}$ . The same identity holds after partial transposition,  $(\sigma_\theta^\Gamma)^2 = \frac{1}{2}\sigma_\theta^\Gamma$  with  $\text{Tr } \sigma_\theta^\Gamma = 1$ , again verified exactly in  $\theta$ , so  $\text{spec}(\sigma_\theta^\Gamma) = \{0, 0, \frac{1}{2}, \frac{1}{2}\} \geq 0$  and  $\sigma_\theta$  is PPT, hence separable by Theorem 3.2. Finally  $\sigma_\theta|\psi^-\rangle = \frac{1}{2}|\psi^-\rangle$  holds identically in  $\theta$ , which is the eigenvector condition characterising the minimisers at  $\lambda = 1$  in Theorem 7.5(iii), and  $D(P_- \parallel \sigma_\theta) = -\ln \frac{1}{2} = \ln 2$ . Different  $\theta$  give different states:  $\sigma_0 - \sigma_{\pi/4} \neq 0$  exactly.

## B.5 The state that refutes the doubled separable ball

Remark 5.5 uses an explicit NPT state inside the ball of radius  $2\eta(\lambda)$  at  $\lambda = 0.24$ . In the computational basis it is

$$\sigma = \begin{pmatrix} 0.3549364950431938 & 0.0686811431460248 & -0.0546630696778 & 0.0686106044832 \\ 0.0686811431460248 & 0.0589933160512 & -0.0654761739481 & 0.0556736143022 \\ -0.0546630696778 & -0.0654761739481 & 0.2748490367138 & -0.0375468736388 \\ 0.0686106044832 & 0.0556736143022 & -0.0375468736388 & 0.3112211521919 \end{pmatrix}, \quad (72)$$

a real symmetric matrix with rational-independent entries, found by a random search and retained in decimal form; every property quoted below is a direct eigenvalue computation on the printed matrix, not a property of the search that produced it. Its trace is 1 to  $10^{-16}$ ,

$$\text{spec}(\sigma) \ni \lambda_{\min}(\sigma) = 0.0276367338 > 0, \quad \lambda_{\min}(\sigma^\Gamma) = -9.7060961312 \times 10^{-4} < 0,$$

so  $\sigma$  is a state and is NPT, and

$$\|\sigma - \rho_{0.24}\|_1 = 0.444034691994310 < 2\eta(0.24) = 0.4933333333,$$

$$\|\sigma - \rho_{0.24}\|_1 - \frac{\eta(0.24)}{K} = 0.0353439771 > 0.$$

The first line refutes the radius- $2\eta(\lambda)$  ball; the second shows that the corrected radius  $\eta(\lambda)/K$  of Corollary 5.3 survives, with the counterexample outside it. The small size of  $\lambda_{\min}(\sigma^\Gamma)$ , of order  $10^{-3}$ , is the reason the refutation needs an explicit matrix rather than a sampling argument.

The companion matrix for the non-isometry of  $\Gamma$  is

$$X = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad X^\Gamma = \begin{pmatrix} -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \quad (73)$$

with  $\text{spec}(X) = \{-1, -1, 1, 1\}$  and  $\text{spec}(X^\Gamma) = \{-\sqrt{2}, 0, 0, \sqrt{2}\}$ , so  $\|X\|_1 = 4$  while  $\|X^\Gamma\|_1 = 2\sqrt{2}$  and the ratio is  $1/\sqrt{2} = 0.707106781186547$ , verified to  $10^{-15}$ . The partial transpose therefore *contracts* the trace norm on this traceless operator, and no argument that treats  $\Gamma$  as a trace-norm isometry can be repaired by a change of convention. The estimate  $\|X^\Gamma\|_{\text{op}} \leq \frac{1}{2}\|X\|_1$ , which would have produced the doubled radius, is separately false: the extremiser  $X_*$  of Proposition 4.6 has  $\|X_*\|_1 = 1$  and  $\|X_*^\Gamma\|_{\text{op}} = K = 0.6035533906 > \frac{1}{2}$ . The constant  $\frac{1}{2}$  is not attained by any traceless operator, and  $K$  is the correct one.

## C Numerical protocols

Section 9 reports what was computed. This appendix records how, in enough detail that each number can be reproduced and each check can be seen to test what it claims to test.

### C.1 Certifying a claimed boundary distance

A claim of the form  $d(\lambda) \leq v$  with an explicit state  $\sigma$  is certified by five tests, in this order, each of which can fail independently:

1. **Normalisation.**  $\text{Tr } \sigma = 1$  to within  $10^{-13}$ , and  $\sigma$  Hermitian by construction rather than by symmetrisation.
2. **Positivity.**  $\lambda_{\min}(\sigma) \geq -10^{-12}$ . A negative eigenvalue of order  $10^{-13}$  is treated as zero only when step 3 confirms that  $\sigma$  is exactly singular; otherwise the state is rejected.
3. **The corner.**  $\det \sigma^\Gamma = 0$  as an *exact* identity in  $\lambda$ , evaluated in rational and radical arithmetic. This is preferred to an eigenvalue test because the two corner families of Appendix B are singular by construction, and a floating-point determinant of order  $10^{-17}$  is otherwise indistinguishable from a small positive eigenvalue.
4. **Partial positivity.**  $\lambda_{\min}(\sigma^\Gamma) \geq -10^{-12}$ , with the full spectrum of  $\sigma^\Gamma$  recorded, because the double-corner families are positive semidefinite on both sides and a sign error in either would otherwise pass steps 1–3.
5. **The distance.**  $\|\sigma - \rho_\lambda\|_1$  is computed from the eigenvalues of the difference, never from the parameter  $2c$  that appears in the closed form. Agreement between the two is the content of the test, and it is what detects a normalisation error in the corner vector  $q$ .

Steps 1–5 were applied to the ray  $\sigma(\lambda)$  of Proposition 6.5 at 9 values of  $\lambda$  in  $[\lambda_c, \frac{1}{4}]$  and one below it, to the isospectral family at 26 values in  $[\lambda_0, \frac{1}{4}]$ , to the rank-two family at 26 values in  $(0, \lambda_0]$ , and to the upper-interval family of Proposition 6.9 at 41 values in  $[\frac{1}{4}, \frac{1}{2})$ . The worst residual over all of them was  $1.7 \times 10^{-16}$  for the spectra and  $8.3 \times 10^{-17}$  for the distances.

## C.2 The witness relaxation

The relaxation (21) requires minimising

$$w \longmapsto D_w(\lambda) = \frac{\langle w | \rho_\lambda^\Gamma w \rangle}{\text{hw}(|w\rangle\langle w|^\Gamma)}$$

over the unit sphere of  $\mathbb{C}^4$ , which is seven real parameters after fixing the phase. A vector  $w$  is written as  $x + iy$  with  $x, y \in \mathbb{R}^4$  and the constraint  $\|x\|^2 + \|y\|^2 = 1$  imposed by renormalisation after every step, so the optimisation is unconstrained in eight variables. The denominator is evaluated from the exact spectrum of  $(|w\rangle\langle w|)^\Gamma$  of Lemma 4.3, using the two Schmidt weights obtained from the singular values of the  $2 \times 2$  matrix of  $w$ ; this avoids a  $4 \times 4$  eigendecomposition per evaluation and keeps the objective smooth except where the two Schmidt weights coincide.

At each of 13 values of  $\lambda$  in  $[0, \frac{1}{2})$  the minimisation was started from 24 points: the four structured vectors  $\psi_0$ ,  $|01\rangle$ ,  $|\phi^+\rangle$  and  $|\psi^-\rangle$ , and 20 draws from a standard complex Gaussian in  $\mathbb{C}^4$ , renormalised. Each start was followed to a relative change below  $10^{-14}$  between successive iterates with gradient norm below  $10^{-10}$ , and the smallest value over the 24 starts was recorded together with the Schmidt weights of its minimiser. The outcome: on  $[\lambda_c, \frac{1}{2})$  the minimum agrees with the certified value of (39) to a relative accuracy of  $6.5 \times 10^{-15}$ ; on  $[0, \frac{1}{4}]$  it equals  $\eta(\lambda)/K$  exactly, and the minimiser is  $\psi_0$  from every start that converges to the global minimum, as Remark 6.4 predicts. The minimisers for  $\lambda > \frac{1}{4}$  all lie in the symmetric plane  $\text{span}\{|\phi^+\rangle, |\psi^+\rangle\}$  and their Schmidt weights are (0.792394, 0.207606) at  $\lambda = 0.30$ , (0.724137, 0.275863) at  $\lambda = 0.35$  and (0.651085, 0.348915) at  $\lambda = 0.40$ , matching the closed form  $c_0(\lambda)$  of (31) through  $t = (a + b)^2$  with  $a^2 = c_0$ .

### C.3 Sampling ensembles

Four ensembles were used, all with fixed seeds and all in double precision.

*Traceless Hermitian operators.* A Ginibre matrix  $G$  with independent complex standard normal entries is Hermitised as  $H = (G + G^\dagger)/2$  and projected onto the traceless subspace by  $X = H - \frac{1}{4}(\text{Tr } H)I_4$ . This is the ensemble behind the 4000-draw test of (10); its largest observed ratio  $\|X^\Gamma\|_{\text{op}}/\|X\|_1$  was 0.5544496, which is  $0.9186K$ .

*Separable states.* Four product vectors  $|a_i\rangle \otimes |b_i\rangle$  with independent complex Gaussian factors, combined with Dirichlet(1, 1, 1, 1) weights. This ensemble cannot produce entangled states and was used for the 1200-challenger test of Theorem 7.2, the 200-state test of Lemma 3.4, and the ball tests of Corollary 5.3.

*PPT states.* A Ginibre state  $\tau = GG^\dagger/\text{Tr}(GG^\dagger)$ , pulled back through the partial transpose as  $\sigma = \tau^\Gamma/\text{Tr } \tau^\Gamma$ ; since  $\Gamma$  is an involution and preserves positivity of the transpose,  $\sigma^\Gamma = \tau/\text{Tr } \tau \geq 0$ . This gives the 400 PPT challengers of Theorem 7.2 and the PPT reference states of the continuity test.

*Arbitrary states.* Ginibre states, used for the 400-state tests of Theorem 8.1, Proposition 8.2 and Corollary 8.4.

For the continuity test, pairs  $(\sigma, \rho)$  were drawn with  $\sigma$  PPT from the third ensemble and  $\rho = \sigma + \varepsilon Z$  with  $Z$  a traceless Hermitian draw from the first ensemble and  $\varepsilon$  uniform in  $[0, 0.6]$ , renormalised and projected back to positivity when necessary; the bound (16) was then tested with  $g = \lambda_{\min}(\sigma^\Gamma)$  and the negativity form of Corollary 5.2 with the same data. No pair violated either bound and the smallest slack was  $2.741 \times 10^{-1}$ .

### C.4 A worked certification

At  $\lambda = 0.35$  the full chain reads as follows. The upper-interval family gives  $R(0.35) = \sqrt{0.65 \times 1.7} = 1.0511898$ , hence  $c(0.35) = \frac{1}{3}(2R - (2\lambda + 1)) = \frac{1}{3}(2.1023796 - 1.7) = 0.1341265347$ , so  $\sigma(0.35)$  is a state with  $\lambda_{\min}(\sigma^\Gamma) = 0$  at distance  $2c = 0.2682530694$ . The symmetric-plane analysis gives  $c_0(0.35) = 0.4734740900$ , whose witness has Schmidt weights  $(0.724137, 0.275863)$ , and at that point  $N_{0.35}(c_0) = d_{\text{up}}(0.35)Q(c_0)$  to within  $10^{-16}$ , while  $N_{0.35}(c) - d_{\text{up}}(0.35)Q(c) \geq 0$  at all 40001 grid points of  $c \in [0, \frac{1}{2}]$ . The witness relaxation over the full sphere, from all 24 starts, returns 0.2682530694, agreeing with the family value to a relative accuracy below  $10^{-14}$ . The Weyl bound at the same  $\lambda$  is  $\eta(\lambda)/K = 0.2485281374$ , strictly smaller, as Proposition 6.7 requires. These are the three numerical columns of the  $\lambda = 0.35$  row of Table 1.

### C.5 Conventions that silently invalidate a check

Three sources of error in this subject are worth naming, because each of them leaves the most natural sanity checks intact.

*The partial-transpose index convention.* In a basis indexed by pairs  $(a, b)$  there are three operations that are commonly called a partial transpose: transpose in the first factor,  $X_{(a,b),(c,e)}^\Gamma = X_{(c,b),(a,e)}$ ; transpose in the second factor,  $X_{(a,b),(c,e)}^{\Gamma_2} = X_{(a,e),(c,b)}$ ; and conjugation by the swap operator,  $X \mapsto FXF$ , which in tensor index order is the permutation  $(a, b, c, e) \mapsto (b, a, e, c)$ . All three coincide on Bell-diagonal states, and the first two have identical spectra on every state, so any test built from the Werner family or from spectra alone cannot distinguish them. They differ on general operators: the third is an algebra automorphism and preserves the spectrum exactly, so with it in place  $\|X^\Gamma\|_{\text{op}} = \|X\|_{\text{op}}$ , the trace norm is preserved, and a constant  $K < 1$  in (10) becomes unreachable while every Werner-family computation still passes. Any check of the sharp

bound must therefore be applied to operators that are not Bell-diagonal, and must first confirm that  $\Gamma$  is not the swap conjugation, for instance by testing  $\|X^\Gamma\|_{\text{op}} < \|X\|_{\text{op}}$  on a single traceless example.

*Eigenvectors versus eigenvalues.* The extremal eigenvector of  $X_*^\Gamma$ , written in (15), is determined only up to an overall phase, and the sibling extremiser obtained by replacing  $|\phi^-\rangle$  with  $|\phi^+\rangle$  has an eigenvector of the same shape with the opposite relative sign between its two computational components. A printed eigenvector whose relative sign does not match the printed matrix is invisible to every spectral test, so it has to be checked by applying the matrix to it: the residual  $\|X_*^\Gamma\psi_0 + K\psi_0\|$  is  $1.1 \times 10^{-16}$  for the sign written in (15).

*Logarithm base.* Entropies and relative entropies are in natural logarithms throughout, and the value  $F(1) = \ln 2 = 0.6931471806$  is one ebit. In base two the same quantity is 1, and the closed form reads  $1 - H_2(\lambda)$ . Mixing the two conventions changes every numerical comparison by a factor  $\ln 2$  while leaving all inequalities structurally intact, so it is caught only by an absolute value such as  $F(1)$  or  $F(\frac{3}{4}) = 0.1308120359411370$  nats.

## References

- [1] Asher Peres. Separability criterion for density matrices. *Physical Review Letters*, 77(8):1413–1415, 1996.
- [2] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223(1–2):1–8, 1996.
- [3] Anna Sanpera, Rolf Tarrach, and Guifré Vidal. Local description of quantum inseparability. *Physical Review A*, 58:826–830, 1998.
- [4] Swapan Rana. Negative eigenvalues of partial transposition of arbitrary bipartite states. *Physical Review A*, 87:054301, 2013.
- [5] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight. Quantifying entanglement. *Physical Review Letters*, 78(12):2275–2279, 1997.
- [6] Adam Miranowicz and Satoshi Ishizaka. Closed formula for the relative entropy of entanglement. *Physical Review A*, 78:032310, 2008.
- [7] K. M. R. Audenaert. A sharp continuity estimate for the von neumann entropy. *Journal of Physics A: Mathematical and Theoretical*, 40(28):8127–8136, 2007.
- [8] GeWu Problem Index. Closed formula for two-qubit relative entropy of entanglement. <https://git.gewu-lab.ai/wangqihang/closed-formula-for-two-qubit-relative-entropy-of-entanglement/-/blob/main/README.md>, 2026. Accessed 2026-09-08.
- [9] GeWu Journal Auto Author. Source research repository for this paper. <https://git.gewu-lab.ai/scientific-author-45/two-qubit-werner-ree-closed-form>, 2026. Solution repository at commit e4839795af45afc2035f037f24bc4d4621ad6a47.
- [10] Leonid Gurvits and Howard Barnum. Largest separable balls around the maximally mixed bipartite quantum state. *Physical Review A*, 66:062311, 2002.
- [11] Ray Ganardi, Marek Miller, Tomasz Paterek, and Marek Żukowski. Hierarchy of correlation quantifiers comparable to negativity. 2022.
- [12] Matthew J. Donald and Michał Horodecki. Continuity of relative entropy of entanglement. *Physics Letters A*, 264(4):257–260, 1999.
- [13] K. G. H. Vollbrecht and R. F. Werner. Entanglement measures under symmetry. *Physical Review A*, 64:062307, 2001.
- [14] Mary Beth Ruskai. Inequalities for quantum entropy: A review with conditions for equality. *Journal of Mathematical Physics*, 43(9):4358–4375, 2002.