Supplemental Material for "Necessary and sufficient quantum information characterization of Einstein-Podolsky-Rosen steering"

Marco Piani[1,2] and John Watrous[3]

[1]Institute for Quantum Computing & Department of Physics and Astronomy, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada
[2]Department of Physics and SUPA, University of Strathclyde, Glasgow G4 0NG, UK
[3]Institute for Quantum Computing & School of Computer Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada

ROBUSTNESS AS SEMIDEFINITE PROGRAM

Inspired by the work of Pusey [1] and Skrzypczyk et al. [2], we are going to prove that calculating the steering robustness $R(\mathcal{A})$ of an assemblage $\mathcal{A}=\{\rho_{a|x}\}_{a,x}$ falls under the umbrella of semidefinite programming (SDP) [3].

We recall that the steering robustness of $\mathcal{A}$ is defined as

$$R(\mathcal{A}) := \min\left\{ t\geq 0 \,\middle|\, \left\{\frac{\rho_{a|x}+t\,\tau_{a|x}}{1+t}\right\}_{a,x} \text{unsteerable}, \{\tau_{a|x}\} \text{ an assemblage}\right\}, \quad (1)$$

hence it corresponds to the minimum positive t such that

$$\rho_{a|x} = (1+t)\sigma^{\mathrm{US}}_{a|x} - t\tau_{a|x}, \quad \forall a,x,$$

with $\{\sigma^{\mathrm{US}}_{a|x}\}_{a,x}$ an unsteerable assemblage and $\{\tau_{a|x}\}_{a,x}$ an arbitrary assemblage. Notice that, since $\{\rho_{a|x}\}_{a,x}$ and $\{\sigma^{\mathrm{US}}_{a|x}\}_{a,x}$ are assemblages, $\tau_{a|x} = \big((1+t)\sigma^{\mathrm{US}}_{a|x} - \rho_{a|x}\big)/t$ is automatically an assemblage as long as

$$(1+t)\sigma^{\mathrm{US}}_{a|x} \geq \rho_{a|x},, \quad \forall a,x. \quad (2)$$

Since $\{\sigma^{\mathrm{US}}_{a|x}\}_{a,x}$ is unsteerable,

$$\rho^{\mathrm{US}}_{a|x} = \sum_\lambda D(a|x,\lambda)\sigma(\lambda) \quad \forall a,x, \quad (3)$$

we can rewrite Eq. (2) as the condition

$$(1+t)\sum_\lambda D(a|x,\lambda)\sigma_\lambda \geq \rho_{a|x}, \quad \forall a,x,$$

where the σ_λ's are subnormalized states, and the sum is over all the deterministic strategies to output a given x. If we consider that the factor $(1+t)$ can be absorbed into the σ_λ's (so that they are generally unnormalized, rather subnormalized), we realize that $R(\mathcal{A})+1$ can be characterized as the solution to

$$\begin{aligned}
\text{minimize} \quad & \sum_\lambda \mathrm{Tr}(\sigma_\lambda) \\
\text{subject to} \quad & \sum_\lambda D(a|x,\lambda)\sigma_\lambda \geq \rho_{a|x} \quad \forall a,x \\
& \sigma_\lambda \geq 0 \quad \forall \lambda
\end{aligned} \quad (4)$$

This is an example of SDP optimization problem [3]. For our purposes, the *primal problem* of an SDP is an optimization problem cast as

$$\begin{aligned}
\text{minimize} \quad & \langle C, X\rangle \\
\text{subject to} \quad & \Phi[X] \geq B \\
& X \geq 0,
\end{aligned}$$

where:

- $\langle C, X\rangle$ is the objective function;
- B and C are given Hermitian matrices;
- X is the matrix variable on which to optimize;
- $\langle X, Y\rangle := \mathrm{Tr}(X^\dagger Y)$ is the Hilbert-Schmidt inner product;
- Φ is a given Hermiticity-preserving linear map.

The *dual problem* provides a lower bound to the objective function of the primal problem. The dual problem is given by

$$\begin{aligned}
\text{maximize} \quad & \langle B, Y\rangle \\
\text{subject to} \quad & \Phi^\dagger[Y] \leq C \\
& Y \geq 0,
\end{aligned}$$

where $\Phi^\dagger$ is the dual of Φ with respect to the Hilbert-Schmidt inner product, and Y is another matrix variable.

One says that *strong duality* holds when the optimal values of the primal and dual problems coincide. Strong duality holds in many cases, and in particular under the Slater conditions that (i) the primal and dual problems are both feasible, and moreover the primal problem is *strictly feasible*, meaning that there is a positive definite $X>0$ such that $\Phi[X]>B$, or (ii) the primal and dual problems are both feasible, and moreover the dual problem is strictly feasible, meaning that there is a $Y>0$ such that $\Phi^\dagger[Y]<C$. In case (i), not only do the primal and dual values coincide, but there must exist Y_{opt} that achieves the optimal value for the dual problem; and similarly, in the case (ii), there must exist X_{opt} that achieves the optimal value in the primal problem.

In our case

$$C = \mathbb{1}, \quad B = \mathrm{diag}(\rho_{a|x})_{a,x},$$
$$\Phi[X] = \mathrm{diag}\left(\sum_\lambda D(a|x,\lambda)X_\lambda\right)_{a,x}$$

where $\mathrm{diag}(\cdot)_{a,x}$ indicates a block-diagonal matrix whose diagonal blocks are labeled by a, x, and the X_λ's are the diagonal blocks of X, labeled by λ. Thus, we have $\Phi^\dagger[Y] = \mathrm{diag}\left(\sum_{a,x} D(a|x,\lambda)Y_{a|x}\right)_\lambda$, and the dual of the primal problem (4) reads

$$\text{maximize} \quad \sum_{a,x} \mathrm{Tr}(F_{a|x}\rho_{a|x}) \tag{5a}$$
$$\text{subject to} \quad \sum_{a,x} D(a|x,\lambda)F_{a|x} \leq \mathbb{1} \quad \forall\lambda \tag{5b}$$
$$F_{a|x} \geq 0 \quad \forall a,x, \tag{5c}$$

It is easy to verify that both Slater conditions hold in our case. For instance, one can take $\sigma_\lambda = 2\mathbb{1}$ for all λ, and $F_{a|x} = \frac{\mathbb{1}}{|X|+1}$ for all a, x, with $|X|$ being the number of possible values for x. Thus, there exist $F_{a|x} = F^{\mathrm{opt}}_{a|x}$ satisfying the constraints of Eq. (5) and such that $\sum_{a,x} \mathrm{Tr}(F_{a|x}\rho_{a|x}) = 1 + R(\mathcal{A})$.

We remark that the optimal $F_{a|x}$ can always be chosen to saturate (5b). That is, there is a deterministic strategy $D(a|x,\lambda)$ and a normalized pure state $|\phi\rangle$ such that

$$\sum_{a,x} D(a|x,\lambda)\langle\phi|F_{a|x}|\phi\rangle = \langle\phi|\mathbb{1}|\phi\rangle = 1 \tag{6}$$

This is because otherwise it is always possible to increase (in operator sense) some $F_{a|x}$'s, still maintaining the optimal value for the objective function (which is operator monotone in the $F_{a|x}$'s).

DETAILS OF THE PROOF OF THEOREM 1

The claimed upper bound,

$$p_{\mathrm{corr}}(\mathcal{I}^B, \mathcal{M}^{B\to A}, \rho_{AB}) \leq (1 + R^{A\to B}_{\mathrm{steer}}(\rho_{AB}))p^{\mathrm{NE}}_{\mathrm{corr}}(\mathcal{I}),$$

can be proved using

$$p_{\mathrm{corr}}(\mathcal{I}^B, \mathcal{M}^{B\to A}, \rho_{AB}) = \sum_{a,x} \mathrm{Tr}_B(\Lambda^{\dagger B}_a[N^B_x]\rho_{a|x}), \tag{7}$$

and the definitions

$$R^{A\to B}_{\mathrm{steer}}(\rho_{AB}) := \sup_{\mathcal{MA}} R(\mathcal{A}), \tag{8}$$

and (1):

$$\begin{aligned}
&p_{\mathrm{corr}}(\mathcal{I}^B, \mathcal{M}^{B\to A}, \rho_{AB})\\
&\qquad = \sum_{a,x} \mathrm{Tr}_B(\Lambda^{\dagger B}_a[N^B_x]\rho_{a|x})\\
&\qquad \leq (1 + R(\mathcal{A}))\sum_{a,x} \mathrm{Tr}_B(\Lambda^{\dagger B}_a[N^B_x]\sigma^{\mathrm{US}}_{a|x})\\
&\qquad\quad - R(\mathcal{A})\sum_{a,x} \mathrm{Tr}_B(\Lambda^{\dagger B}_a[N^B_x]\tau^{\mathrm{US}}_{a|x})\\
&\qquad \leq (1 + R(\mathcal{A}))p^{\mathrm{NE}}_{\mathrm{corr}}(\mathcal{I})\\
&\qquad \leq (1 + R^{A\to B}_{\mathrm{steer}}(\rho_{AB}))p^{\mathrm{NE}}_{\mathrm{corr}}(\mathcal{I}).
\end{aligned}$$

On the other hand, suppose that $\mathcal{MA} = \{M_{a|x}\}_{a,x}$, where $a = 1, \ldots, |A|$ and $x = 1, \ldots, |X|$, is a measurement assemblage on A such that the corresponding assemblage $\mathcal{A} = \{\rho_{a|x} = \mathrm{Tr}_A(M^A_{a|x}\rho_{AB})\}_{a,x}$ is steerable. Let $F_{a|x} \geq 0$ be the operators optimal for (5), such that $\sum_{a,x} \mathrm{Tr}(F_{a|x}\rho_{a|x}) = 1 + R(\mathcal{A})$. In the proof of Theorem 1 of the main text we defined subchannels Λ_a that act as

$$\Lambda_a[\rho] = \begin{cases} \alpha\sum_{x=1}^{|X|} \mathrm{Tr}(\rho F_{a|x})|x\rangle\langle x| & 1 \leq a \leq |A| \\ \frac{1}{N}\mathrm{Tr}((\mathbb{1} - \sum_{a=1}^{|A|}\Lambda^\dagger_a[\mathbb{1}])\rho)\hat{\sigma}_a & |A|+1 \leq a \leq |A|+N, \end{cases} \tag{9}$$

where $\alpha = \|\sum_{a,x} F_{a|x}\|^{-1}_\infty > 0$, and the $\hat{\sigma}_a$, $a = |A|+1, \ldots, |A|+N$, are arbitrary (normalized) states in a two-dimensional subspace orthogonal to $\mathrm{span}\{|x\rangle \,|\, x = 1, \ldots, |X|\}$. It is immediate to check that $\mathrm{Tr}\left(\sum_{a=1}^{|A|+N}\Lambda_a[\rho]\right) = \mathrm{Tr}(\rho)$ (by construction), so $\mathcal{I} = \{\Lambda_a\}_{a=1,\ldots,|A|+N}$ is an instrument for the channel $\sum_{a=1}^{|A|+N}\Lambda_a$.

Let σ_{AB} be an arbitrary bipartite state on AB, and let $\mathcal{MA}^{B\to A} = \{Q_a\}^{B\to A}_a$ be an arbitrary one-way measurement from B to A, i.e., $Q^{B\to A}_a = \sum_y M'^A_{a|y} \otimes N'^B_y$, to guess which subchannel was actually realized. Notice that y in the latter expression potentially varies in an arbitrary range, different from the range $\{1, \ldots, |X|\}$ for the parameter x of the fixed measurement assemblage $\mathcal{MA}$. Nonetheless we observe that $\Lambda_a = \Pi'_X \circ \Lambda_a$ for $a = 1, \ldots, |A|+N$, where $\circ$ is composition, and

$$\Pi'_X[\tau] = \sum_{x=1}^{|X|} |x\rangle\langle x|\tau|x\rangle\langle x| + \Pi^\perp\tau\Pi^\perp,$$

with $\Pi^\perp$ the projector onto the two-dimensional space orthogonal to $\mathrm{span}\{|x\rangle \,|\, x = 1, \ldots, |X|\}$ that supports the arbitrary qubits states $\hat{\sigma}_a$, $a = |A|+1, \ldots, |A|+N$. Also,

$$\Lambda^B_a[\sigma_{AB}] = \frac{1}{N}\left(\sigma_A - \sum_{a'=1}^{|A|} \mathrm{Tr}_B(\Lambda^B_{a'}[\sigma_{AB}])\right) \otimes \hat{\sigma}^B_a,$$

for $a = |A|+1, \dots, |A|+N$. This implies that, for whatever input σ_{AB}, the optimal $Q_a^{B\to A}$ can be chosen to have the form

$$Q_a^{B\to A} = \begin{cases} \sum_{x=1}^{|X|} M_{a|x}'^{A} \otimes |x\rangle\langle x|^B & 1 \le a \le |A| \\ \mathbb{1}^A \otimes N_a^B, & |A|+1 \le a \le |A|+N, \end{cases} \tag{10}$$

with $\Pi^\perp N_a \Pi^\perp = N_a$, for $|A|+1 \le a \le |A|+N$, a POVM on the orthogonal qubit space. Omitting a detailed and straightforward proof of this, we instead provide the following intuition: For the subchannels (9), the best local measurement on the output probe is one that first of all discriminates between the space $\mathrm{span}\{|x\rangle \mid x = 1, \dots, |X|\}$ and the orthogonal qubit space. If the probe is found in the space $\mathrm{span}\{|x\rangle \mid x = 1, \dots, |X|\}$, the probe is then measured in the basis $\{|x\rangle \mid x = 1, \dots, |X|\}$ and the result if forwarded to decide which measurement to perform on the ancilla: this is optimal because, in this subspace, the output probe is already dephased in the basis $\{|x\rangle \mid x = 1, \dots, |X|\}$. If the probe is instead found in the orthogonal qubit space, there is no information to be gained from the ancilla, since, for the state of the probe to have support in the orthogonal qubit space, the probe must have been discarded and prepared in one of the random qubit states $\hat\sigma_a$. So, in this case, the ancilla is necessarily decorrelated and its state independent of the specific Λ_a, $a = |A|+1, \dots, |A|+N$, that has been realized; thus the optimal guess about said Λ_a can be made as soon as the output probe is measured.

Then, for an optimal $\mathcal{M}^{B\to A} = \{Q_a^{B\to A}\}_a$ of the form (10), we find in general

$$\begin{aligned}
&p_{\mathrm{corr}}(\mathcal{I}^B, \mathcal{M}^{B\to A}, \sigma_{AB}) \\
&= \sum_{a=1}^{|A|+N} \mathrm{Tr}(Q_a^{B\to A} \Lambda_a^B[\sigma_{AB}]) \\
&= \sum_{a=1}^{|A|} \mathrm{Tr}(Q_a^{B\to A} \Lambda_a^B[\sigma_{AB}]) + \sum_{a=|A|+1}^{|A|+N} \mathrm{Tr}(Q_a^{B\to A} \Lambda_a^B[\sigma_{AB}]) \\
&= \sum_{a=1}^{|A|} \sum_{x=1}^{|X|} \mathrm{Tr}(M_{a|x}'^{A} \otimes |x\rangle\langle x|_B \Lambda_a^B[\sigma_{AB}]) \\
&\quad + \left(1 - \sum_{a=1}^{|A|} \mathrm{Tr}(\Lambda_a^B[\sigma_{AB}])\right) \frac{1}{N} \sum_{a=|A|+1}^{|A|+N} \mathrm{Tr}(N_a \hat\sigma_a) \\
&= \sum_{a=1}^{|A|} \sum_{x=1}^{|X|} \mathrm{Tr}(\Lambda_a^\dagger[|x\rangle\langle x|] \sigma_{a|x}) \\
&\quad + \left(1 - \sum_{a=1}^{|A|} \mathrm{Tr}(\Lambda_a^B[\sigma_{AB}])\right) \frac{1}{N} \sum_{a=|A|+1}^{|A|+N} \mathrm{Tr}(N_a \hat\sigma_a),
\end{aligned}$$

with $\sigma_{a|x} = \mathrm{Tr}_A(M'_{a|x}\sigma_{AB})$. By construction it holds that $\Lambda_a^\dagger[|x\rangle\langle x|] = \alpha F_{a|x}$ for $1 \le a \le |A|$ and $1 \le x \le |X|$, therefore

$$\begin{aligned}
&p_{\mathrm{corr}}(\mathcal{I}^B, \mathcal{M}^{B\to A}, \rho_{AB}) \\
&= \alpha \sum_{a=1}^{|A|} \sum_{x=1}^{|X|} \mathrm{Tr}(F_{a|x}\sigma_{a|x}) \\
&\quad + \left(1 - \sum_{a=1}^{|A|} \mathrm{Tr}(\Lambda_a^B[\sigma_{AB}])\right) \frac{1}{N} \sum_{a=|A|+1}^{|A|+N} \mathrm{Tr}(N_a \hat\sigma_a) \\
&\le \alpha \sum_{a=1}^{|A|} \sum_{x=1}^{|X|} \mathrm{Tr}(F_{a|x}\sigma_{a|x}) + \frac{2}{N}.
\end{aligned} \tag{11}$$

In the last line we used

$$\left(1 - \sum_{a=1}^{|A|} \mathrm{Tr}(\Lambda_a^B[\sigma_{AB}])\right) \le 1$$

and

$$\begin{aligned}
\frac{1}{N} \sum_{a=|A|+1}^{|A|+N} \mathrm{Tr}(N_a \hat\sigma_a) &\le \frac{1}{N} \mathrm{Tr}\left(\sum_{a=|A|+1}^{|A|+N} N_a\right) \\
&\le \frac{1}{N} \mathrm{Tr}(\Pi^\perp) = \frac{2}{N}.
\end{aligned} \tag{12}$$

It is clear that if $\sigma_{AB} = \rho_{AB}$ and $M'_{a|x} = M_{a|x}$ in (10), so that $\sigma_{a|x} = \rho_{a|x}$, then we have

$$1 + R(\mathcal{A}) \le p_{\mathrm{corr}}(\mathcal{I}^B, \mathcal{M}^{B\to A}, \rho_{AB}) \le 1 + R(\mathcal{A}) + \frac{2}{N}.$$

It remains to prove that

$$\alpha \le p_{\mathrm{corr}}^{\mathrm{NE}}(\mathcal{I}) \le \alpha + \frac{2}{N}. \tag{13}$$

This is readily verified by considering that (5b) can be saturated, as argued at the end of the previous section (see (6)), for an optimal solution of the SDP problem. So we have that for some deterministic $D(a|x,\lambda)$ and some uncorrelated input state $|\phi\rangle$ to the channel,

$$\begin{aligned}
1 &= \sum_{a=1}^{|A|} \sum_{x=1}^{|X|} D(a|x,\lambda) \langle\phi|F_{a|x}|\phi\rangle \\
&= \frac{1}{\alpha} \sum_{a=1}^{|A|} \sum_{x=1}^{|X|} D(a|x,\lambda) \langle\phi|\Lambda_a^\dagger[|x\rangle\langle x|]|\phi\rangle \\
&= \frac{1}{\alpha} \sum_{a=1}^{|A|} \mathrm{Tr}\left(\left(\sum_{x:D(a|x,\lambda)=1} |x\rangle\langle x|\right) \Lambda_a[|\phi\rangle\langle\phi|]\right) \\
&= \frac{1}{\alpha} \sum_{a=1}^{|A|} \mathrm{Tr}\left(M_a'' \Lambda_a[|\phi\rangle\langle\phi|]\right),
\end{aligned}$$

having defined $M_a'' := \sum_{x:D(a|x,\lambda)=1} |x\rangle\langle x|$. Considering also the subchannels Λ_a, $a = |A|+1, \dots, |A|+N$, and bounding their contribution to the probability of success as in (11), we arrive at (13).

ON THE SCALING OF THE STEERABILITY OF MAXIMALLY ENTANGLED STATES

We have argued that $R_{\text{steer}}^{A\to B}(\rho_{AB}) \leq R_g(\rho_{AB})$, where $R_g(\rho_{AB})$ is the generalized entanglement robustness

$$R_g(\rho_{AB}) = \min\left\{t \geq 0 \,\middle|\, \frac{\rho_{AB} + t\,\tau_{AB}}{1+t} \text{ separable}, \tau \text{ a state}\right\}.$$

Indeed, let τ_{AB} be optimal for the generalized entanglement robustness, i.e., suppose

$$\sigma_{AB} = \frac{\rho_{AB} + R_g(\rho_{AB})\tau_{AB}}{1 + R_g(\rho_{AB})}$$

is separable. Then $\sigma_{a|x} = \text{Tr}_A(M_{a|x}\sigma_{AB})$ is unsteerable for any measurement assemblage $\{M_{a|x}\}_{a,x}$, proving that $R_g(\rho_{AB})$ is an upper bound to $R_{\text{steer}}^{A\to B}(\rho_{AB})$ (see Eq. (8)). This means that, if a state is weakly entangled with respect to R_g, it is also weakly steerable with respect to $R_{\text{steer}}^{A\to B}$. In [4] it was proven that, for any bipartite pure state

$$|\psi\rangle_{AB} = \sum_i \sqrt{p_i}|i\rangle_A|i\rangle_B,$$

here in its Schmidt decomposition, the generalized entanglement robustness is equal to

$$R_g(|\psi\rangle\langle\psi|_{AB}) = \left(\sum_i \sqrt{p_i}\right)^2 - 1 = 2\mathcal{N}(|\psi\rangle\langle\psi|_{AB}),$$

where $\mathcal{N}$ is the negativity of entanglement [5]. In particular, then, for a maximally entangled state in dimension $d\times d$, $|\psi_d^+\rangle_{AB} = \frac{1}{\sqrt{d}}\sum_{i=1}^d |i\rangle_A|i\rangle_B$, one has

$$R_{\text{steer}}^{A\to B}(\psi_{d,AB}^+) \leq R_g(\psi_{d,AB}^+) = d-1,$$

having used the notation $\psi_{d,AB}^+ = |\psi_d^+\rangle\langle\psi_d^+|_{AB}$.

We conclude by providing a lower bound on $R_{\text{steer}}^{A\to B}(\psi_{d,AB}^+)$ for d a power of a prime number. We will use techniques similar to the ones used in the examples of [6].

Fix d to be the power of a prime number. Then we know that there there are $d+1$ mutually unbiased bases, i.e., $d+1$ orthonormal sets $\{|\psi_{a|x}\rangle\}_{a=1,\ldots,d}$, one for each $x = 1,\ldots,d+1$, such that [7]

$$|\langle\psi_{a|x}|\psi_{b|y}\rangle| = \begin{cases}\delta_{a,b} & x = y\\ \frac{1}{\sqrt{d}} & x\neq y\end{cases}$$

We will consider a measurement assemblage $\{M_{a|x} = |\psi_{a|x}\rangle\langle\psi_{a|x}|\}_{a,x}$. Suppose $\rho_{AB} = \psi_{d,AB}^+$. We have

$$\rho_{a|x}^B = \text{Tr}_A(M_{a|x}^A\psi_{d,AB}^+) = \frac{1}{d}|\psi_{a|x}^*\rangle\langle\psi_{a|x}^*|$$

Here $|\psi_{a|x}^*\rangle$ indicates orthonormal vectors whose coefficients in the local basis $\{|i\rangle_B\}$ are the complex conjugate of the coefficients of $|\psi_{a|x}\rangle$ in the local basis $\{|i\rangle_A\}$. Thus, the bases $\{|\psi_{a|x}^*\rangle\}_{a=1,\ldots,d}$ are still mutually unbiased.

We want to lower bound the steering robustness of $\{\rho_{a|x}^B\}_{a,x}$, which in turn will give us a lower bound on $R_{\text{steer}}^{A\to B}(\psi_{d,AB}^+)$. To do this, we use a specific choice for the $F_{a|x}$'s in (5). We choose $F_{a|x} = \beta|\psi_{a|x}^*\rangle\langle\psi_{a|x}^*|$, where $\beta > 0$ will be fixed to satisfy (5b) (condition (5c) is satisfied for any $\beta \geq 0$), i.e.,

$$\left\|\sum_{a,x} D(a|x,\lambda)F_{a|x}\right\|_\infty \leq 1$$

for all deterministic $D(a|x,\lambda)$. With our choice of $F_{a|x}$, this can be achieved by taking

$$\beta \leq \left(\max_\lambda \left\|\sum_x |\psi_{f_\lambda(x)|x}^*\rangle\langle\psi_{f_\lambda(x)|x}^*|\right\|_\infty\right)^{-1} \tag{14}$$

where the maximum is over all functions $f_\lambda : \{1,\ldots,d+1\}\to\{1,\ldots,d\}$, labeled by λ. To estimate the right hand side of (14), we will use the fact [8] that, for

$$|\gamma\rangle_{CD} = \sum_{x=1}^{d+1}|\psi_{f_\lambda(x)|x}^*\rangle_C|x\rangle_D,$$

where $\{|x\rangle\}_{x=1,\ldots,d+1}$ is an orthonormal basis, the spectrum of

$$\text{Tr}_D(|\gamma\rangle\langle\gamma|_{CD}) = \sum_x |\psi_{f_\lambda(x)|x}^*\rangle\langle\psi_{f_\lambda(x)|x}^*|.$$

is the same as the spectrum of

$$\begin{aligned}\text{Tr}_C(|\gamma\rangle\langle\gamma|_{CD}) &= \sum_{x,y}\langle\psi_{f_\lambda(x)|x}^*|\psi_{f_\lambda(y)|y}^*\rangle|y\rangle\langle x|\\ &= \sum_x |x\rangle\langle x| + \frac{1}{\sqrt{d}}\sum_{x\neq y}e^{i\phi_{x,y}}|y\rangle\langle x|\\ &= \left(1-\frac{1}{\sqrt{d}}\right)\mathbb{1} + \frac{1}{\sqrt{d}}\sum_{x,y}e^{i\phi_{x,y}}|y\rangle\langle x|\end{aligned}$$

where $\phi_{x,y}$ are real numbers representing phases. Thus, we have

$$\begin{aligned}&\left\|\sum_x |\psi_{f_\lambda(x)|x}^*\rangle\langle\psi_{f_\lambda(x)|x}^*|\right\|_\infty\\ &\quad= \left\|\sum_{x,y}\langle\psi_{f_\lambda(x)|x}^*|\psi_{f_\lambda(y)|y}^*\rangle|y\rangle\langle x|\right\|_\infty\\ &\quad\leq \left(1-\frac{1}{\sqrt{d}}\right) + \frac{1}{\sqrt{d}}\left\|\sum_{x,y}e^{i\phi_{x,y}}|y\rangle\langle x|\right\|_\infty\\ &\quad\leq \left(1-\frac{1}{\sqrt{d}}\right) + \frac{1}{\sqrt{d}}\left\|\sum_{x,y}e^{i\phi_{x,y}}|y\rangle\langle x|\right\|_2\\ &\quad= \left(1-\frac{1}{\sqrt{d}}\right) + \frac{1}{\sqrt{d}}(d+1)\\ &\quad= 1+\sqrt{d}.\end{aligned}$$

Since this estimate is independent of λ, we can take $\beta = 1/(\sqrt{d}+1)$. Hence, we conclude that, for d the power of a prime number,

$$\begin{aligned}
&R^{A\to B}_{\text{steer}}(\psi^+_{d,AB}) \\
&\geq R\left(\left\{\frac{1}{d}|\psi^*_{a|x}\rangle\langle\psi^*_{a|x}|\right\}\right) \\
&\geq \sum_{a,x} \mathrm{Tr}\left(\left(\frac{1}{d}|\psi^*_{a|x}\rangle\langle\psi^*_{a|x}|\right)\left(\frac{1}{\sqrt{d}+1}|\psi^*_{a|x}\rangle\langle\psi^*_{a|x}|\right)\right) - 1 \\
&= \frac{1}{d(\sqrt{d}+1)}(d(d+1)) - 1 \\
&= \sqrt{d}\frac{\sqrt{d}-1}{\sqrt{d}+1} \\
&\geq \sqrt{d} - 2.
\end{aligned} \tag{15}$$

[1] M. F. Pusey, Phys. Rev. A **88**, 032313 (2013), URL `http://link.aps.org/doi/10.1103/PhysRevA.88.032313`.

[2] P. Skrzypczyk, M. Navascués, and D. Cavalcanti, Phys. Rev. Lett. **112**, 180404 (2014), URL `http://link.aps.org/doi/10.1103/PhysRevLett.112.180404`.

[3] S. P. Boyd and L. Vandenberghe, *Convex optimization* (Cambridge university press, 2004).

[4] A. Harrow and M. Nielsen, Phys. Rev. A **68**, 012308 (2003).

[5] G. Vidal and R. F. Werner, Phys. Rev. A **65**, 032314 (2002), URL `http://link.aps.org/doi/10.1103/PhysRevA.65.032314`.

[6] Z. Jiang, M. Piani, and C. M. Caves, Quantum information processing **12**, 1999 (2013).

[7] I. Bengtsson, arXiv preprint quant-ph/0610216 (2006).

[8] R. Jozsa and J. Schlienz, Phys. Rev. A **62**, 012301 (2000).