Convex optimization in quantum information

Volkher B. Scholz

based on joint work with M. Berta (Caltech) & O. Fawzi (ENS Lyon)

SIAM J. Opt., 26, 1529 (2016)

WIC Midwintermeeting 2017

GHENT UNIVERSITY

Convex optimization and information theory

- Many problems in (classical) information theory can be formulated as convex optimization problems
- Examples: Computing capacities, determining good coding schemes, finding upper bounds on errors (to be discussed),…
- There has already been a very fruitful relationship between classical information theory and optimization theory
- **Quantum information theory**: in very abstract terms, **commuting** variables have to be replaced by **non-commuting** variables -> optimization over non-commutative fields
- **This talk**: an introduction into the the subject, based on the problem of sending information over a noisy channel
- In particular, we will focus on constructing **relaxations** to obtain **upper bounds**

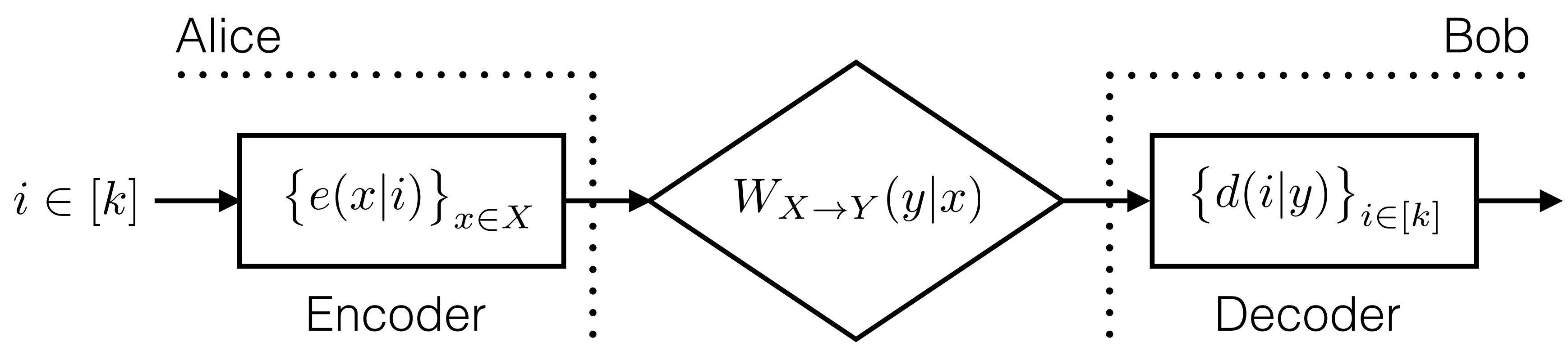
Overview

- Classical noisy channel coding
- Linear program relaxations and the meta-converse
- Entanglement-assisted *channel coding*
  (**quantum assistance**)
- Semidefinite programming (SDPs)
- SDP hierarchies for understanding (bounding) the difference between classical and quantum variables
- Conclusion / Outlook

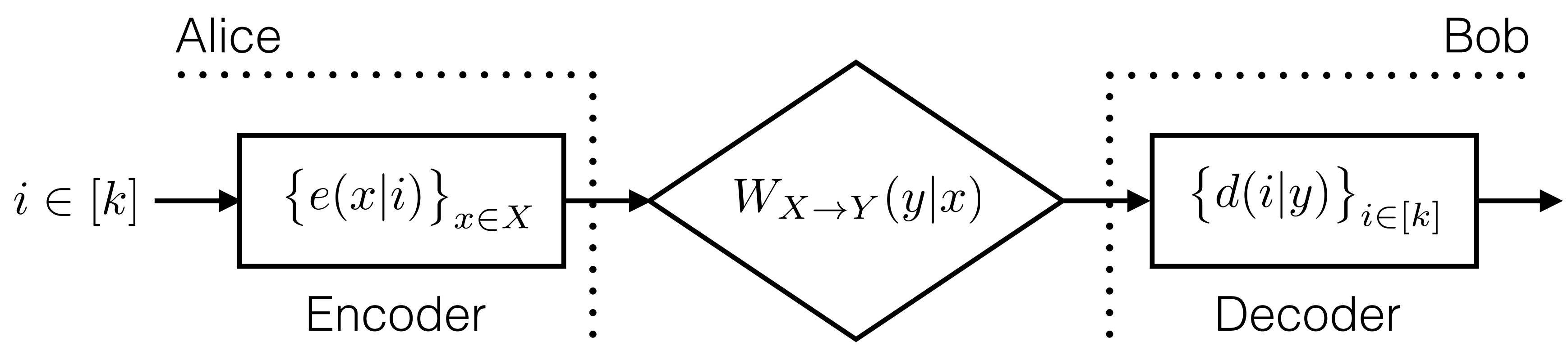
Overview

- Classical noisy channel coding
- Linear program relaxations and the meta-converse
- Entanglement-assisted *channel coding*
  (**quantum assistance**)
- Semidefinite programming (SDPs)
- SDP hierarchies for understanding (bounding) the difference between classical and quantum variables
- Conclusion / Outlook

Classical noisy channel coding (I)



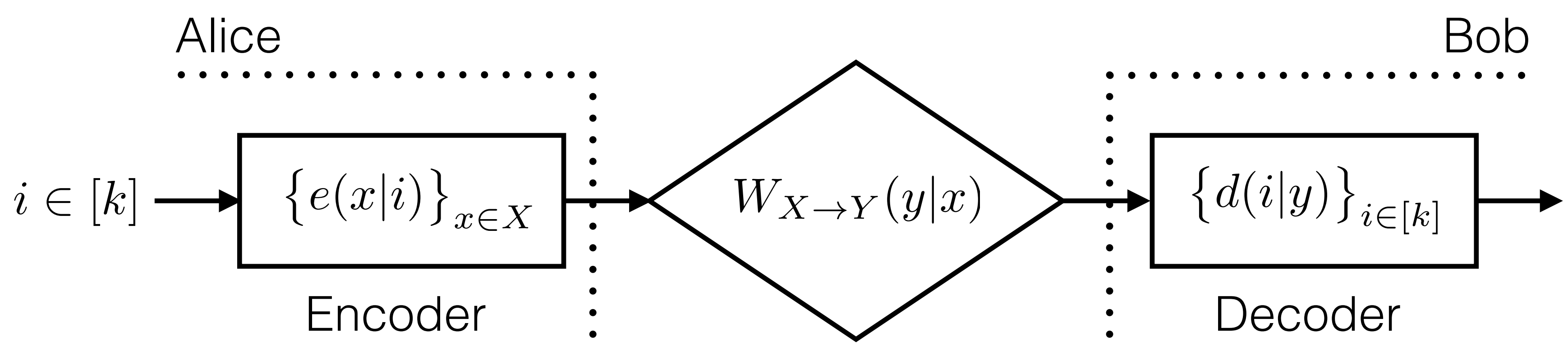
Classical noisy channel coding (I)



- Given noisy channel $W_{X \to Y}$ mapping X to Y with transition probability:

$$W_{X \to Y}(y|x) \; \forall (x, y) \in X \times Y$$

Classical noisy channel coding (I)

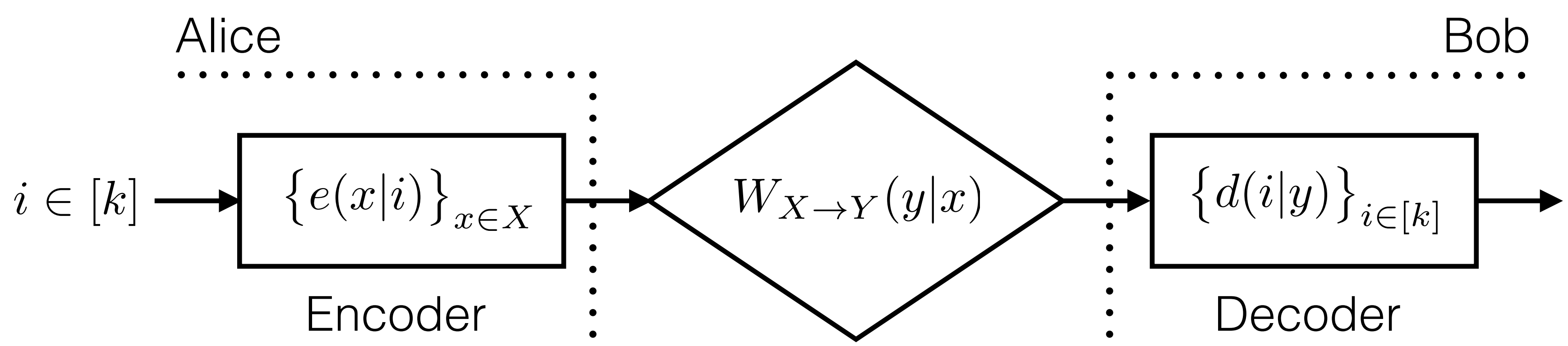


- Given noisy channel $W_{X \to Y}$ mapping X to Y with transition probability:

$$W_{X \to Y}(y|x) \; \forall (x, y) \in X \times Y$$

- The goal is to send k different messages using W while minimising the error probability for decoding:

Classical noisy channel coding (I)



- Given noisy channel $W_{X\to Y}$ mapping X to Y with transition probability:

$$W_{X\to Y}(y|x)\ \forall (x,y) \in X \times Y$$

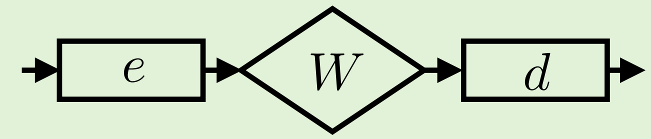
- The goal is to send k different messages using W while minimising the error probability for decoding:

"bilinear optimisation"

$$\begin{aligned} p_{\mathrm{succ}}(W,k) := \underset{(e,d)}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)e(x|i)d(i|y) \\ \text{subject to} \quad & \sum_x e(x|i) = 1 \quad \forall i \in [k]\,, \quad \sum_i d(i|y) = 1 \quad \forall y \in Y \\ & 0 \le e(x|i) \le 1 \quad \forall (x,i) \in X \times [k]\,, \quad 0 \le d(i|y) \le 1 \quad \forall (i,y) \in [k] \times Y\,. \end{aligned}$$

Classical noisy channel coding (II)

$$
\begin{aligned}
p_{\mathrm{succ}}(W,k) := \underset{(e,d)}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)e(x|i)d(i|y) \\
\text{subject to} \quad & \sum_x e(x|i) = 1 \quad \forall i \in [k]\,, \quad \sum_i d(i|y) = 1 \quad \forall y \in Y \\
& 0 \le e(x|i) \le 1 \quad \forall (x,i) \in X \times [k]\,, \quad 0 \le d(i|y) \le 1 \quad \forall (i,y) \in [k] \times Y\,.
\end{aligned}
$$

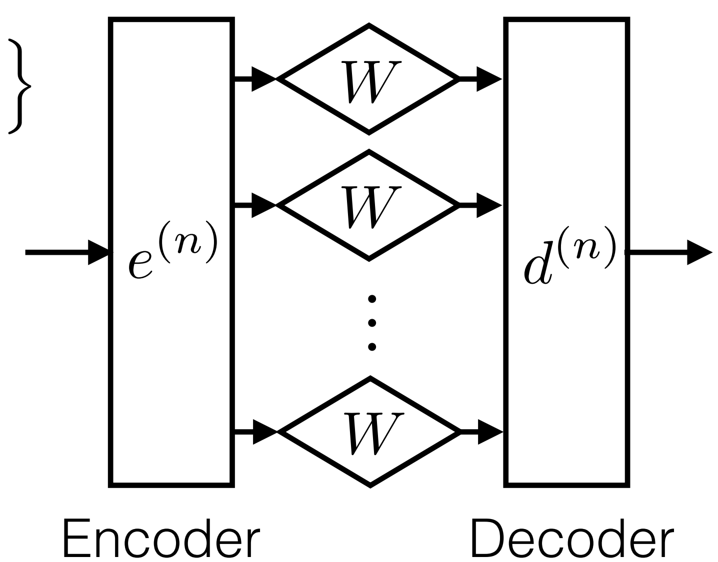


compared to

- Shannon's asymptotic independent and identical distributed (iid) channel capacity:

Definition: $C(W) := \sup\left\{R \,\middle|\, \forall\delta > 0 : \lim_{n\to\infty} p_{\mathrm{succ}}\left(W^{\times n}, [R(1-\delta)]^n\right) = 1\right\}$

Answer: $C(W) = \max_{P_X} I(X:Y)$ mutual information

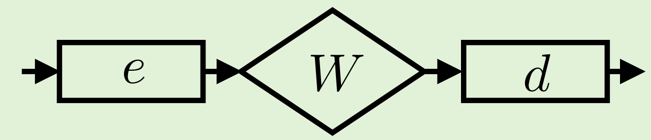


Overview

- Classical noisy channel coding
- Linear program relaxations and the meta-converse
- Entanglement-assisted *channel coding* (**quantum assistance**)
- Semidefinite programming (SDPs)
- SDP hierarchies for understanding (bounding) the difference between classical and quantum variables
- Conclusion / Outlook

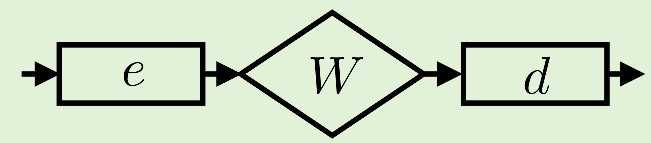
Linear program relaxations (I)

$$
\begin{aligned}
p_{\mathrm{succ}}(W,k) := \underset{(e,d)}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)e(x|i)d(i|y) \\
\text{subject to} \quad & \sum_{x} e(x|i) = 1 \quad \forall i \in [k]\,, \quad \sum_{i} d(i|y) = 1 \quad \forall y \in Y \\
& 0 \le e(x|i) \le 1 \quad \forall (x,i) \in X \times [k]\,, \quad 0 \le d(i|y) \le 1 \quad \forall (i,y) \in [k] \times Y\,.
\end{aligned}
$$



Linear program relaxations (I)

$$
\begin{aligned}
p_{\mathrm{succ}}(W,k) := \underset{(e,d)}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)e(x|i)d(i|y) \\
\text{subject to} \quad & \sum_x e(x|i) = 1 \quad \forall i \in [k]\,, \quad \sum_i d(i|y) = 1 \quad \forall y \in Y \\
& 0 \le e(x|i) \le 1 \quad \forall (x,i) \in X \times [k]\,, \quad 0 \le d(i|y) \le 1 \quad \forall (i,y) \in [k] \times Y\,.
\end{aligned}
$$



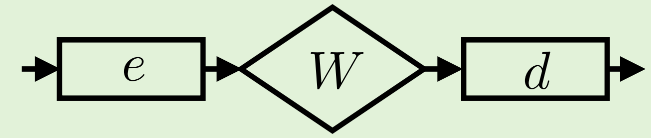
idea: consider this quantity as a matrix in x and y variables, and collect all its linear constraints

$$r(x,y) = \sum_i e(x|i)d(i|y)$$

entry-wise positive

Linear program relaxations (I)

$$p_{\text{succ}}(W,k) := \underset{(e,d)}{\text{maximize}} \quad \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)e(x|i)d(i|y)$$



$$\text{subject to} \quad \sum_x e(x|i) = 1 \quad \forall i \in [k]\,, \quad \sum_i d(i|y) = 1 \quad \forall y \in Y$$

$$0 \le e(x|i) \le 1 \quad \forall (x,i) \in X \times [k]\,, \quad 0 \le d(i|y) \le 1 \quad \forall (i,y) \in [k] \times Y\,.$$

idea: consider this quantity as a matrix in x and y variables, and collect all its linear constraints

$$r(x,y) = \sum_i e(x|i)d(i|y)$$

entry-wise positive

Then:

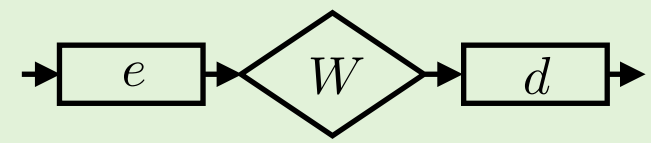
$$\sum_x r(x,y) = \sum_x e(x|i) \sum_i d(i|y) \le 1$$

Linear program relaxations (I)

$$p_{\text{succ}}(W,k) := \underset{(e,d)}{\text{maximize}} \quad \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)e(x|i)d(i|y)$$

$$\text{subject to} \quad \sum_x e(x|i) = 1 \quad \forall i \in [k]\,, \quad \sum_i d(i|y) = 1 \quad \forall y \in Y$$

$$0 \le e(x|i) \le 1 \quad \forall (x,i) \in X \times [k]\,, \quad 0 \le d(i|y) \le 1 \quad \forall (i,y) \in [k] \times Y\,.$$



idea: consider this quantity as a matrix in x and y variables, and collect all its linear constraints

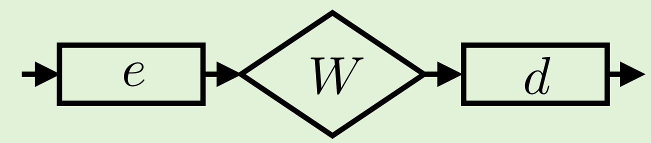
$$r(x,y) = \sum_i e(x|i)d(i|y)$$ entry-wise positive

Then: $$\sum_x r(x,y) = \sum_x e(x|i) \sum_i d(i|y) \le 1$$

Moreover, define $$p(x) = \sum_i e(x|i)$$ entry-wise positive

Linear program relaxations (I)

$$p_{\text{succ}}(W,k) := \underset{(e,d)}{\text{maximize}} \quad \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)e(x|i)d(i|y)$$



$$\text{subject to} \quad \sum_x e(x|i) = 1 \quad \forall i \in [k]\,, \quad \sum_i d(i|y) = 1 \quad \forall y \in Y$$

$$0 \le e(x|i) \le 1 \quad \forall (x,i) \in X \times [k]\,, \quad 0 \le d(i|y) \le 1 \quad \forall (i,y) \in [k] \times Y\,.$$

idea: consider this quantity as a matrix in x and y variables, and collect all its linear constraints

$$r(x,y) = \sum_i e(x|i)d(i|y)$$ entry-wise positive

Then: $$\sum_x r(x,y) = \sum_x e(x|i) \sum_i d(i|y) \le 1$$

Moreover, define $$p(x) = \sum_i e(x|i)$$ entry-wise positive

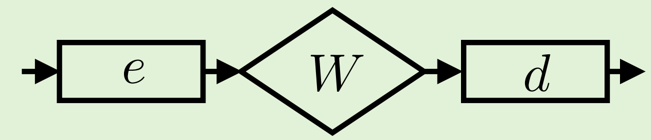
Then: $$\sum_x p(x) = \sum_i \sum_x e(x|i) = \sum_i = k$$

Linear program relaxations (I)

$$p_{\mathrm{succ}}(W,k) := \underset{(e,d)}{\mathrm{maximize}} \quad \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)e(x|i)d(i|y)$$

$$\text{subject to} \quad \sum_x e(x|i) = 1 \quad \forall i \in [k]\,, \quad \sum_i d(i|y) = 1 \quad \forall y \in Y$$

$$0 \le e(x|i) \le 1 \quad \forall (x,i) \in X \times [k]\,, \quad 0 \le d(i|y) \le 1 \quad \forall (i,y) \in [k] \times Y\,.$$



idea: consider this quantity as a matrix in x and y variables, and collect all its linear constraints

$$r(x,y) = \sum_i e(x|i)d(i|y)$$ entry-wise positive

Then: $$\sum_x r(x,y) = \sum_x e(x|i) \sum_i d(i|y) \le 1$$

Moreover, define $$p(x) = \sum_i e(x|i)$$ entry-wise positive

Then: $$\sum_x p(x) = \sum_i \sum_x e(x|i) = \sum_i = k$$

And also: $$r(x,y) = \sum_i e(x|i)d(i|y) \le \sum_i e(x|i) = p(x)$$

Linear program relaxations (II)

- The bilinear program can be relaxed to a linear program:

$$
\begin{aligned}
\underset{(r,p)}{\text{maximize}} \quad & \frac{1}{k} \sum_{x,y} W_{X\to Y}(y|x) r(x,y) \\
\text{subject to} \quad & r(x,y) \geq 0\,,\ p(x) \geq 0,\ \sum_x r(x,y) \leq 1, \\
& \sum_x p(x) = k,\ \ r(x,y) \leq p(x)\,.
\end{aligned}
$$

Linear program relaxations (II)

- The bilinear program can be relaxed to a linear program:

$$\begin{aligned}
\underset{(r,p)}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y} W_{X\to Y}(y|x)r(x,y) \\
\text{subject to} \quad & r(x,y) \geq 0\,,\ p(x) \geq 0,\ \sum_x r(x,y) \leq 1, \\
& \sum_x p(x) = k,\ \ r(x,y) \leq p(x)\,.
\end{aligned}$$

- This is the meta-converse of Polyanski and Verdu (c.f. Matthews)

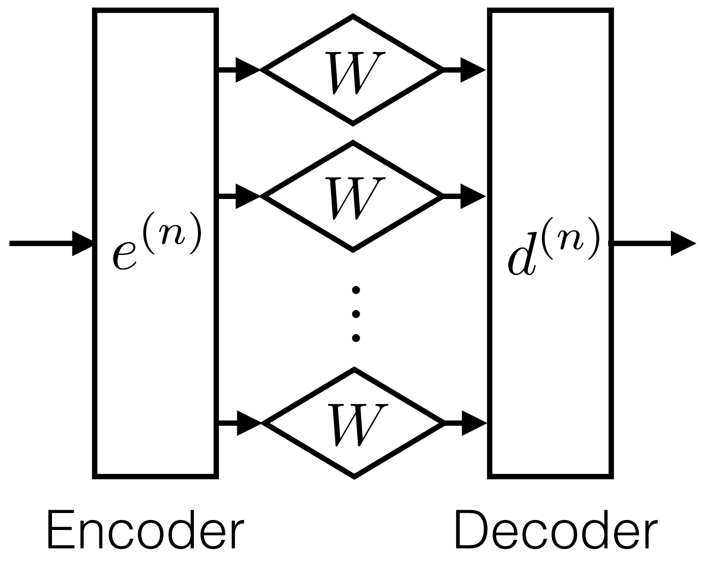
Linear program relaxations (II)

- The bilinear program can be relaxed to a linear program:

$$\begin{aligned} \underset{(r,p)}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y} W_{X\to Y}(y|x) r(x,y) \\ \text{subject to} \quad & r(x,y) \geq 0\,,\; p(x) \geq 0,\; \sum_x r(x,y) \leq 1, \\ & \sum_x p(x) = k,\;\; r(x,y) \leq p(x)\,. \end{aligned}$$

- This is the meta-converse of Polyanski and Verdu (c.f. Matthews)

- It converges to the exact value of the bilinear optimization program in the limit of many channel uses

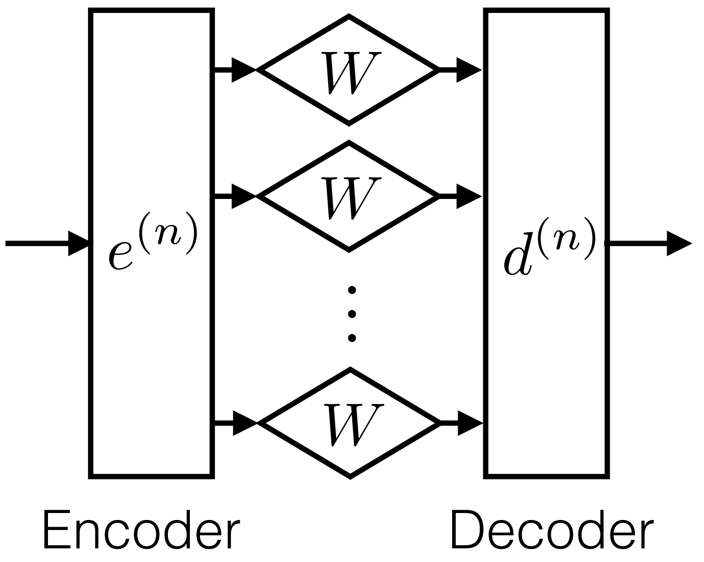


Linear program relaxations (II)

- The bilinear program can be relaxed to a linear program:

$$\begin{aligned} \underset{(r,p)}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y} W_{X\to Y}(y|x) r(x,y) \\ \text{subject to} \quad & r(x,y) \geq 0\,,\ p(x) \geq 0,\ \sum_{x} r(x,y) \leq 1, \\ & \sum_{x} p(x) = k,\ \ r(x,y) \leq p(x)\,. \end{aligned}$$

- This is the meta-converse of Polyanski and Verdu (c.f. Matthews)
- It converges to the exact value of the bilinear optimization program in the limit of many channel uses
- Can be used to obtain very sharp finite n bounds



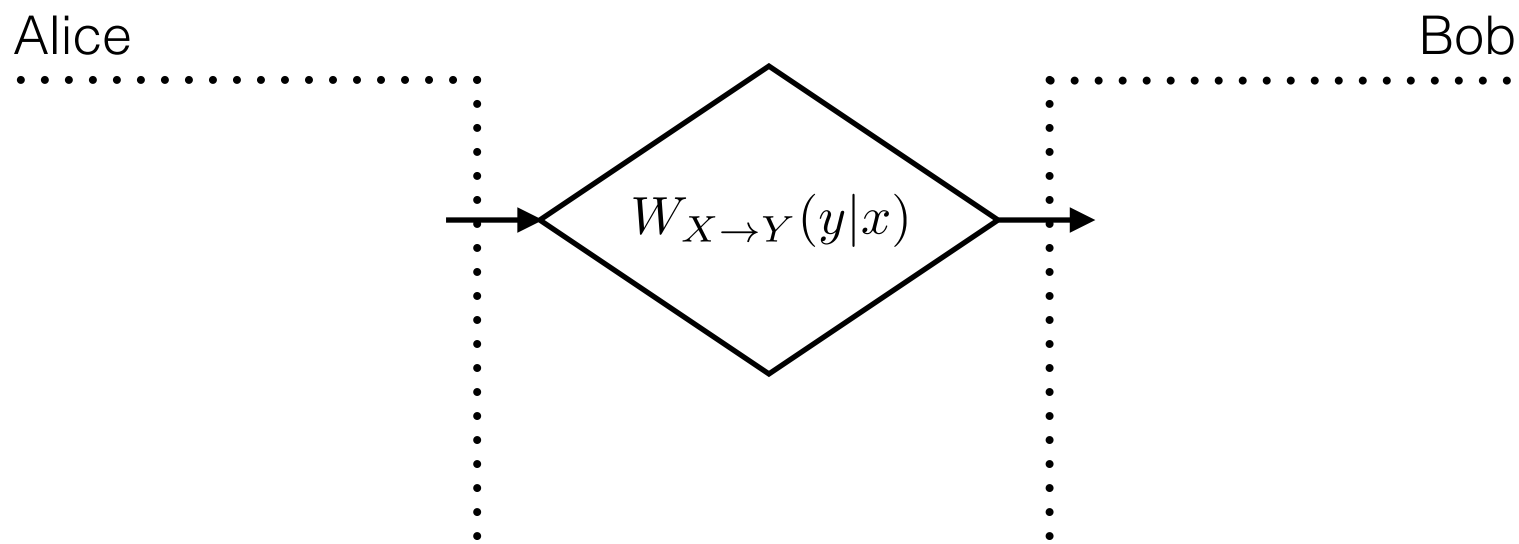
Overview

- Classical noisy channel coding
- Linear program relaxations and the meta-converse
- Entanglement-assisted *channel coding* (**quantum assistance**)
- Semidefinite programming (SDPs)
- SDP hierarchies for understanding (bounding) the difference between classical and quantum variables
- Conclusion / Outlook

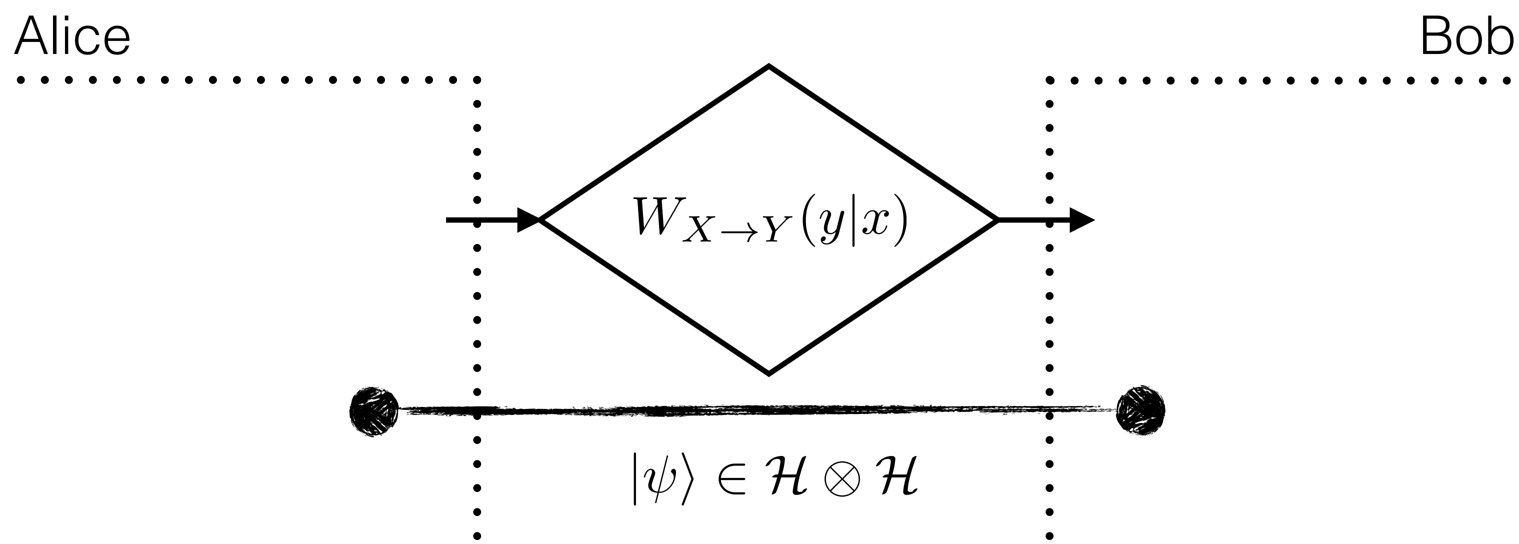
Adding quantum assistance….
but first recap what “quantum” means

- Classical variables are replaced by operators (“matrices”) on a Hilbert space $\mathcal{H}$
- Hilbert space $\mathcal{H}$ = linear vector space equipped with a scalar product (sesquilinear positive definite form)
- Configurations of the physical system are described by elements $|\psi\rangle$ (“states”) of unit norm of the Hilbert space
- Separated labaratories are described by taking the tensor product of their respective Hilbert spaces $\mathcal{H} \otimes \mathcal{H}$

Entanglement-assisted channel coding (I)

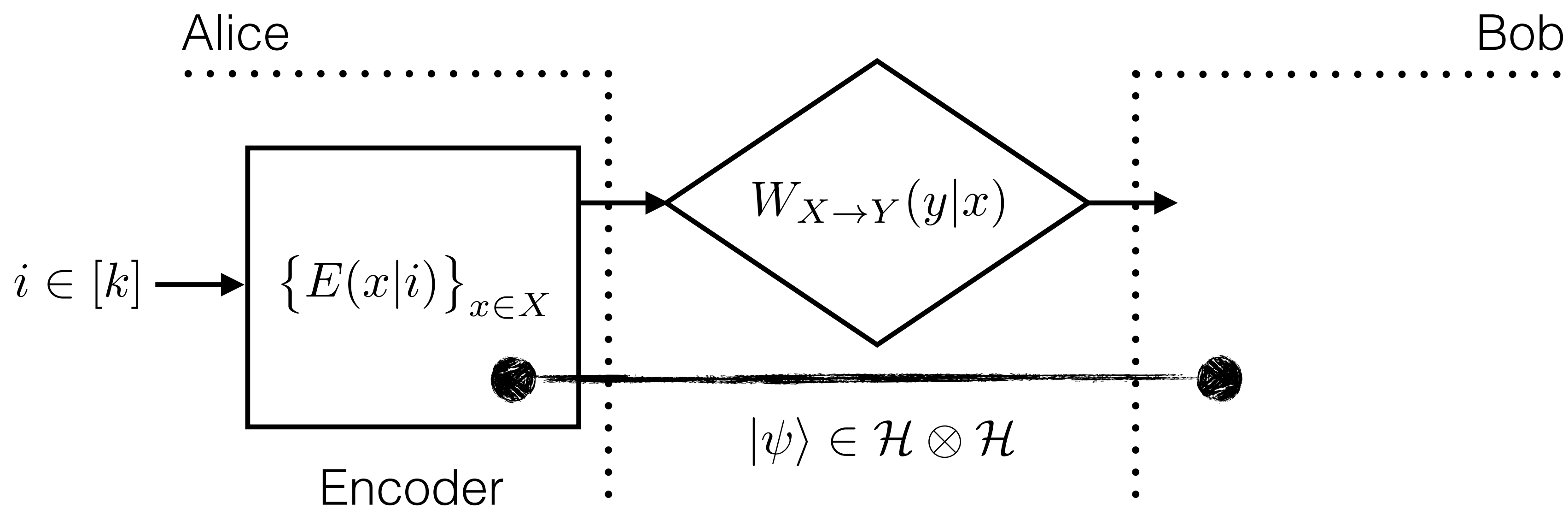


Entanglement-assisted channel coding (I)



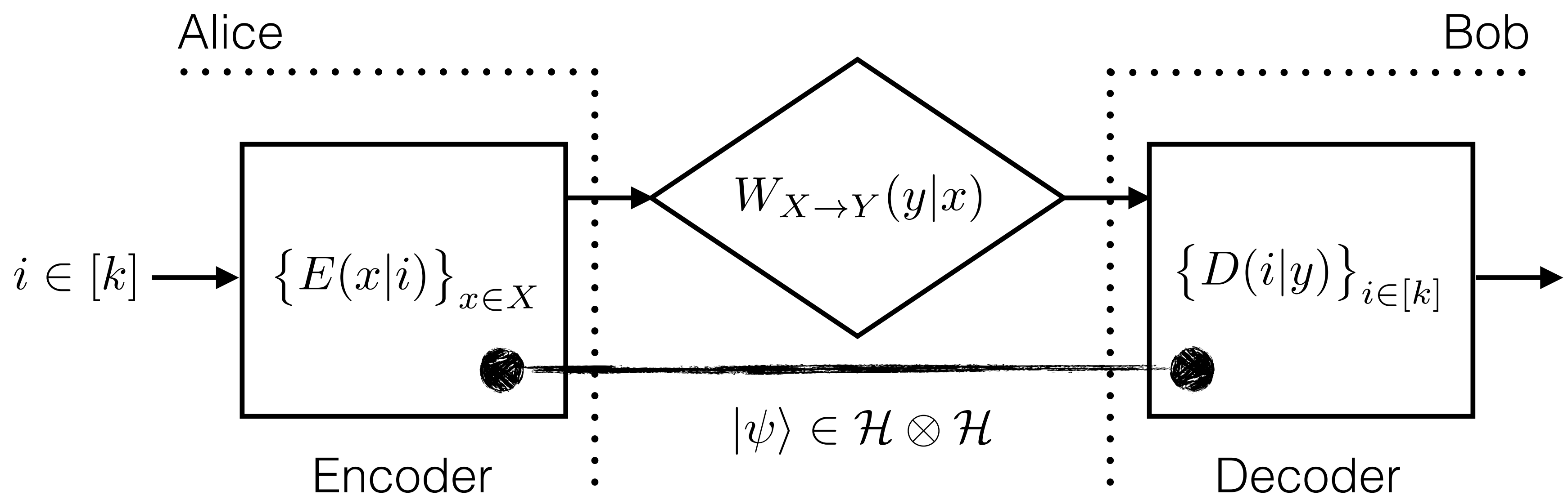
- The encoder and the decoder share an entangled quantum state

Entanglement-assisted channel coding (I)



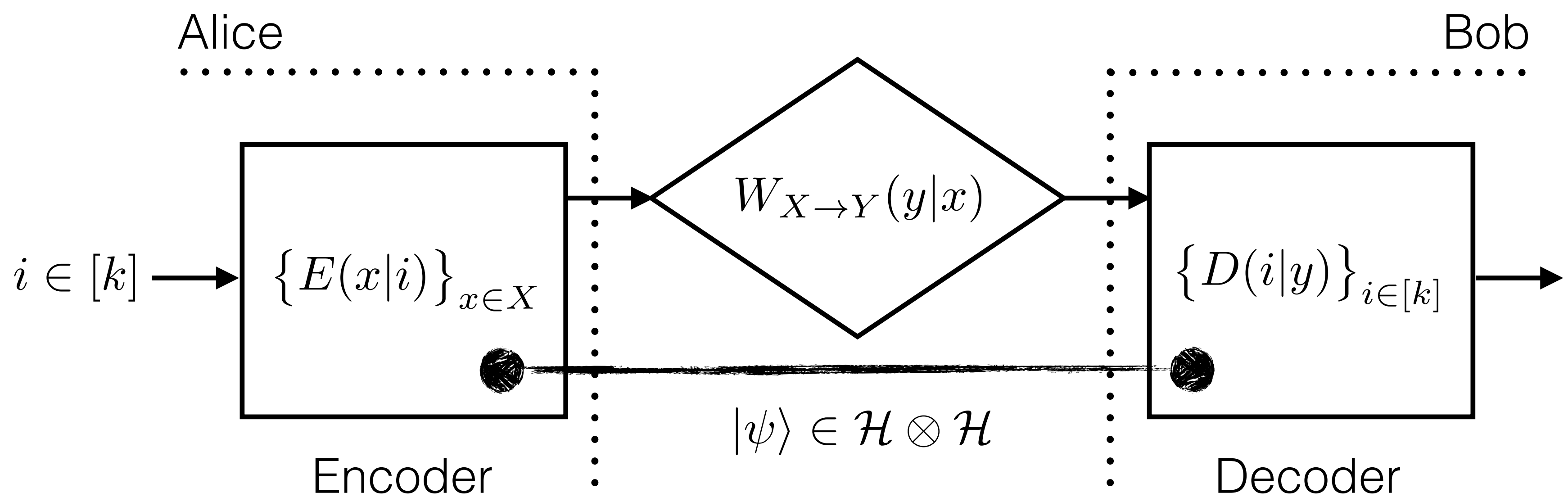
- The encoder and the decoder share an entangled quantum state
- The decoder can also measure his share and combine the outcome with his received output to guess the message

Entanglement-assisted channel coding (I)



- The encoder and the decoder share an entangled quantum state
- The decoder can also measure his share and combine the outcome with his received output to guess the message
- The encoder can first measure his part of the entangled pair and depending on the ou can choose an input into the channel

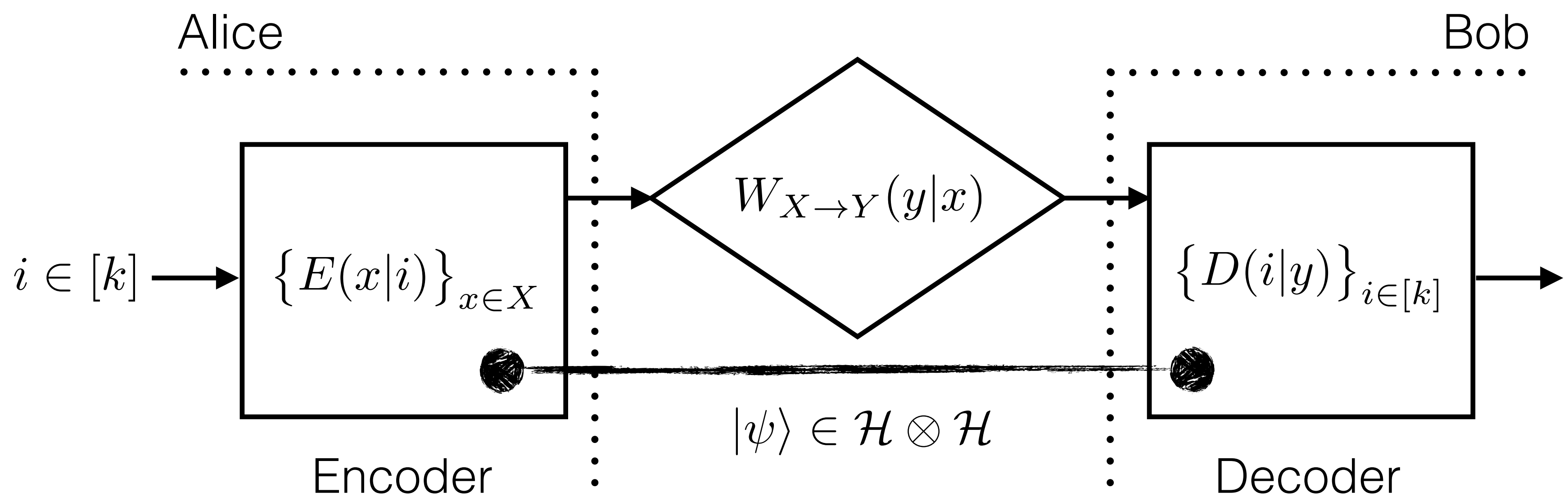
Entanglement-assisted channel coding (I)



"quantum bilinear optimisation"

$$
\begin{aligned}
p^*_{\text{succ}}(W,k) := \underset{(\mathcal{H},\psi,E,D)}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\langle\psi|E(x|i)\otimes D(i|y)|\psi\rangle \\
\text{subject to} \quad & \sum_x E(x|i) = 1_{\mathcal{H}} \quad \forall i\in[k]\,, \quad \sum_i D(i|y) = 1_{\mathcal{H}} \quad \forall y \in Y \\
& 0 \le E(x|i) \le 1_{\mathcal{H}}\ \forall (x,i)\in X\times[k]\,, \quad 0 \le D(i|y) \le 1_{\mathcal{H}}\ \forall (i,y)\in[k]\times Y\,.
\end{aligned}
$$

Entanglement-assisted channel coding (I)



"quantum bilinear optimisation"

$$p^*_{\mathrm{succ}}(W,k) := \underset{(\mathcal{H},\psi,E,D)}{\text{maximize}} \ \frac{1}{k} \sum_{x,y,i} W_{X\to Y}(y|x) \langle \psi | E(x|i) \otimes D(i|y) | \psi \rangle$$

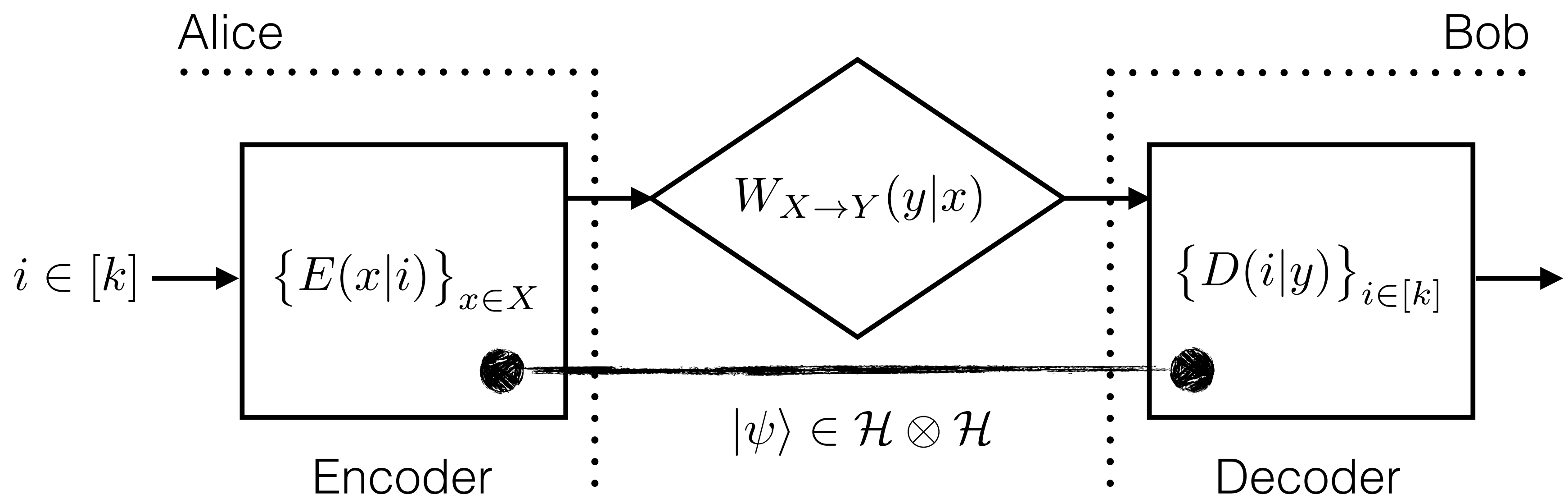
$$\text{subject to} \ \sum_x E(x|i) = 1_{\mathcal{H}} \quad \forall i \in [k]\,, \quad \sum_i D(i|y) = 1_{\mathcal{H}} \quad \forall y \in Y$$

$$0 \leq E(x|i) \leq 1_{\mathcal{H}} \ \forall (x,i) \in X \times [k]\,, \quad 0 \leq D(i|y) \leq 1_{\mathcal{H}} \ \forall (i,y) \in [k] \times Y\,.$$

- **Scalar** (commutative) versus **matrix** (non-commutative) variables:

$$p_{\mathrm{succ}}(W,k) := \underset{(e,d)}{\text{maximize}} \ \frac{1}{k} \sum_{x,y,i} W_{X\to Y}(y|x)\, e(x|i) d(i|y)$$

Entanglement-assisted channel coding (I)



"quantum bilinear optimisation"

$$
\begin{aligned}
p^*_{\mathrm{succ}}(W,k) := \underset{(\mathcal{H},\psi,E,D)}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\langle\psi|E(x|i)\otimes D(i|y)|\psi\rangle \\
\text{subject to} \quad & \sum_x E(x|i) = 1_{\mathcal{H}} \quad \forall i \in [k]\,, \quad \sum_i D(i|y) = 1_{\mathcal{H}} \quad \forall y \in Y \\
& 0 \le E(x|i) \le 1_{\mathcal{H}} \ \forall (x,i) \in X\times[k]\,, \quad 0 \le D(i|y) \le 1_{\mathcal{H}} \ \forall (i,y) \in [k]\times Y\,.
\end{aligned}
$$

- **Scalar** (commutative) versus **matrix** (non-commutative) variables:

$$p_{\mathrm{succ}}(W,k) := \underset{(e,d)}{\text{maximize}} \ \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\, e(x|i)d(i|y)$$

- Unknown if $p^*_{\mathrm{succ}}(W,k)$ is computable!

Entanglement-assisted channel coding (II)

- Understand the possible separation: $p_{\mathrm{succ}}(W, k) \quad \text{versus} \quad p^{*}_{\mathrm{succ}}(W, k)$

Entanglement-assisted channel coding (II)

- Understand the possible separation: $p_{\mathrm{succ}}(W,k) \quad \text{versus} \quad p^*_{\mathrm{succ}}(W,k)$

- For the asymptotic iid capacity entanglement (quantum) assistance does not help:

  $C(W) = C^*(W)$ [Bennett et al., PRL (1999)]

Entanglement-assisted channel coding (II)

- Understand the possible separation: $p_{\mathrm{succ}}(W,k) \quad \text{versus} \quad p^*_{\mathrm{succ}}(W,k)$

- ~~For the asymptotic iid capacity entanglement (quantum) assistance does not help:~~

  ~~$C(W) = C^*(W)$ [Bennett et al., PRL (1999)]~~

Entanglement-assisted channel coding (II)

- Understand the possible separation: $p_{\text{succ}}(W,k) \quad \text{versus} \quad p^*_{\text{succ}}(W,k)$

- ~~For the asymptotic iid capacity entanglement (quantum) assistance does not help:~~
  ~~$C(W) = C^*(W)$ [Bennett et al., PRL (1999)]~~

- In general, there is a **separation**:

$$Z = \begin{pmatrix} 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 0 \\ 0 & 1/3 & 0 & 1/3 \\ 1/3 & 0 & 0 & 1/3 \\ 0 & 1/3 & 1/3 & 0 \end{pmatrix}$$

$$p_{\text{succ}}(Z,2) = \frac{5}{6} \approx 0.833 \quad \text{vs.} \quad p^*_{\text{succ}}(Z,2) \geq \frac{2+2^{-1/2}}{3} \approx 0.902$$

[Prevedel et al., PRL (2011)]

Entanglement-assisted channel coding (II)

- Understand the possible separation: $p_{\text{succ}}(W,k) \quad \text{versus} \quad p^*_{\text{succ}}(W,k)$

- ~~For the asymptotic iid capacity entanglement (quantum) assistance does not help:~~
  ~~$C(W) = C^*(W)$ [Bennett et al., PRL (1999)]~~

- In general, there is a **separation**:

$$Z = \begin{pmatrix} 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 0 \\ 0 & 1/3 & 0 & 1/3 \\ 1/3 & 0 & 0 & 1/3 \\ 0 & 1/3 & 1/3 & 0 \end{pmatrix}$$

$$p_{\text{succ}}(Z,2) = \frac{5}{6} \approx 0.833 \quad \text{vs.} \quad p^*_{\text{succ}}(Z,2) \geq \frac{2+2^{-1/2}}{3} \approx 0.902$$

[Prevedel et al., PRL (2011)]

—> this is also optimal with two-dimensional assistance

[Hemenway et al., PRA (2013)]
[Williams and Bourdon, arXiv:1109.1029]

Entanglement-assisted channel coding (II)

- Understand the possible separation: $p_{\text{succ}}(W,k) \quad \text{versus} \quad p^*_{\text{succ}}(W,k)$

- ~~For the asymptotic iid capacity entanglement (quantum) assistance does not help:~~
  ~~$C(W) = C^*(W)$ [Bennett et al., PRL (1999)]~~

- In general, there is a **separation**:

$$Z = \begin{pmatrix} 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 0 \\ 0 & 1/3 & 0 & 1/3 \\ 1/3 & 0 & 0 & 1/3 \\ 0 & 1/3 & 1/3 & 0 \end{pmatrix}$$

$$p_{\text{succ}}(Z,2) = \frac{5}{6} \approx 0.833 \quad \text{vs.} \quad p^*_{\text{succ}}(Z,2) \geq \frac{2+2^{-1/2}}{3} \approx 0.902$$

[Prevedel et al., PRL (2011)]

—> this is also optimal with two-dimensional assistance

[Hemenway et al., PRA (2013)]
[Williams and Bourdon, arXiv:1109.1029]

- However: $[0.902, 1] \ni p^*_{\text{succ}}(Z,2) = ?$

Entanglement-assisted channel coding (II)

- Understand the possible separation: $p_{\mathrm{succ}}(W,k) \quad \text{versus} \quad p^*_{\mathrm{succ}}(W,k)$

- ~~For the asymptotic iid capacity entanglement (quantum) assistance does not help:~~
  ~~$C(W) = C^*(W)$ [Bennett et al., PRL (1999)]~~

- In general, there is a **separation**:

$$Z = \begin{pmatrix} 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 0 \\ 0 & 1/3 & 0 & 1/3 \\ 1/3 & 0 & 0 & 1/3 \\ 0 & 1/3 & 1/3 & 0 \end{pmatrix}$$

$$p_{\mathrm{succ}}(Z,2) = \frac{5}{6} \approx 0.833 \quad \text{vs.} \quad p^*_{\mathrm{succ}}(Z,2) \geq \frac{2+2^{-1/2}}{3} \approx 0.902$$

[Prevedel et al., PRL (2011)]

—> this is also optimal with two-dimensional assistance

[Hemenway et al., PRA (2013)]
[Williams and Bourdon, arXiv:1109.1029]

- However: $[0.902, 1] \ni p^*_{\mathrm{succ}}(Z,2) = ?$

How to obtain upper bounds for the entangled scenario?

Overview

- Classical noisy channel coding
- Linear program relaxations and the meta-converse
- Entanglement-assisted *channel coding* (**quantum assistance**)
- Semidefinite programming (SDPs)
- SDP hierarchies for understanding (bounding) the difference between classical and quantum variables
- Conclusion / Outlook

Semi-definite programming (SDPs) (I)

- A semi-definite program (SDP) is a triple (T, A, B) with T a hermitian preserving linear map on matrices, and A, B hermitian matrices

- We can associate to such a triple two optimization programs

$$\begin{aligned} \underset{X}{\text{maximize}} \quad & \mathrm{Trace}[X\,A] \\ \text{subject to} \quad & X \geq 0 \\ & T(X) \leq B \end{aligned} \qquad\qquad \begin{aligned} \underset{X}{\text{minimize}} \quad & \mathrm{Trace}[Y\,B] \\ \text{subject to} \quad & Y \geq 0 \\ & T(X) \geq A \end{aligned}$$

- Most often, their value agree; this leads to efficient (in terms of the size of the matrices and the approximation error) optimization algorithms

- SDPs can be used to obtain a hierarchy of outer approximations to convex optimization problems (Lassere and Parillo)

Semi-definite programming (SDPs) (II)

- SDPs can be used to obtain a hierarchy of outer approximations to convex optimization problems (Lassere and Parillo)

- Idea: finding a consistent way to construct SDP relaxations of convex optimization problems

- Can be generalized to non-commutative variables (Navascues et. al., Doherty et. al.)

- Our contribution: **new converging sequence of tighter SDP relaxations** for **quantum bilinear optimization** problems such as the channel coding problem with entanglement assistance

$p_{\mathrm{succ}}(W,k) \leq p^*_{\mathrm{succ}}(W,k) = \mathrm{sdp}_\infty(W,k) \leq \ldots \leq \mathrm{sdp}_1(W,k)$ <— *efficiently computable!*

Overview

- Classical noisy channel coding
- Linear program relaxations and the meta-converse
- Entanglement-assisted *channel coding*
  (**quantum assistance**)
- Semidefinite programming (SDPs)
- SDP hierarchies for understanding (bounding) the difference between classical and quantum variables
- Conclusion / Outlook

First level semidefinite programming relaxation (I)

- Quantum bilinear program:

$$
\begin{aligned}
p^*_{\mathrm{succ}}(W,k) := \underset{(\mathcal{H},\psi,E,D)}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\langle\psi|E(x|i)\otimes D(i|y)|\psi\rangle \\
\text{subject to} \quad & \sum_x E(x|i) = 1_{\mathcal{H}} \quad \forall i\in[k]\,, \quad \sum_i D(i|y) = 1_{\mathcal{H}} \quad \forall y\in Y \\
& 0 \le E(x|i) \le 1_{\mathcal{H}}\ \forall (x,i)\in X\times[k]\,, \quad 0\le D(i|y)\le 1_{\mathcal{H}}\ \forall (i,y)\in[k]\times Y\,.
\end{aligned}
$$

First level semidefinite programming relaxation (I)

- Quantum bilinear program:

$$p^*_{\text{succ}}(W,k) := \underset{(\mathcal{H},\psi,E,D)}{\text{maximize}} \quad \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\langle\psi|E(x|i)\otimes D(i|y)|\psi\rangle \leq \langle\psi|E(x|i)\cdot D(y|j)|\psi\rangle$$

$$\text{with}\left[E(x,i),D(y,j)\right]=0$$

$$\text{subject to} \quad \sum_x E(x|i) = 1_{\mathcal{H}} \quad \forall i\in[k]\,, \quad \sum_i D(i|y) = 1_{\mathcal{H}} \quad \forall y \in Y$$

$$0 \leq E(x|i) \leq 1_{\mathcal{H}}\ \forall (x,i)\in X\times[k]\,, \quad 0\leq D(i|y)\leq 1_{\mathcal{H}}\ \forall (i,y)\in[k]\times Y\,.$$

First level semidefinite programming relaxation (I)

- Quantum bilinear program:

idea: relaxation of this bilinear form

$$p^*_{\mathrm{succ}}(W,k) := \underset{(\mathcal{H},\psi,E,D)}{\text{maximize}} \quad \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\langle\psi|E(x|i)\otimes D(i|y)|\psi\rangle \leq \langle\psi|E(x|i)\cdot D(y|j)|\psi\rangle$$

$$\text{with}\left[E(x,i),D(y,j)\right]=0$$

$$\text{subject to} \quad \sum_x E(x|i) = 1_{\mathcal{H}} \quad \forall i\in[k]\,, \quad \sum_i D(i|y) = 1_{\mathcal{H}} \quad \forall y\in Y$$

$$0\leq E(x|i)\leq 1_{\mathcal{H}}\ \forall (x,i)\in X\times[k]\,, \quad 0\leq D(i|y)\leq 1_{\mathcal{H}}\ \forall (i,y)\in[k]\times Y\,.$$

*motivated by: "**NPA hierarchy**" (Bell inequalities)*

[Lasserre, SIAM (2001)], [Parrilo, Math. Program. (2003)], [Navascues et al., PRL (2007)], [Doherty et al., IEEE CCC (2008)], [Navascues et al., NJP (2008)], [Pironio et al., SIAM (2010)]

First level semidefinite programming relaxation (I)

- Quantum bilinear program:

idea: relaxation of this bilinear form

$$p^*_{\text{succ}}(W,k) := \underset{(\mathcal{H},\psi,E,D)}{\text{maximize}} \quad \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\langle\psi|E(x|i)\otimes D(i|y)|\psi\rangle \leq \langle\psi|E(x|i)\cdot D(y|j)|\psi\rangle$$

$$\text{with}\,\big[E(x,i),D(y,j)\big]=0$$

$$\text{subject to} \quad \sum_x E(x|i) = 1_{\mathcal{H}} \quad \forall i\in[k], \quad \sum_i D(i|y) = 1_{\mathcal{H}} \quad \forall y\in Y$$

$$0\leq E(x|i)\leq 1_{\mathcal{H}}\ \forall (x,i)\in X\times[k], \quad 0\leq D(i|y)\leq 1_{\mathcal{H}}\ \forall (i,y)\in[k]\times Y\,.$$

- First step: see as the part of the upper-right block of the Gram matrix

$$\Omega = \sum_{u,v}\langle\psi|X_u X_v|\psi\rangle|u\rangle\langle v| \quad \text{with} \quad X_u = \begin{cases} E(x|i) & u=(i,x) \\ D(j|y) & u=(j,y)\end{cases}$$

for i=j

$$\Omega = \left(\begin{array}{c:c} \langle\psi|E(x|i)\cdot E(x'|i')|\psi\rangle & \langle\psi|E(x|i)\cdot D(y|j)|\psi\rangle \\ \hdashline \langle\psi|E(x'|i')\cdot D(y'|j')|\psi\rangle & \langle\psi|D(y|j)\cdot D(y'|j')]|\psi\rangle \end{array}\right)$$

---

*motivated by: "**NPA hierarchy**" (Bell inequalities)*

[Lasserre, SIAM (2001)], [Parrilo, Math. Program. (2003)], [Navascues et al., PRL (2007)], [Doherty et al., IEEE CCC (2008)], [Navascues et al., NJP (2008)], [Pironio et al., SIAM (2010)]

First level semidefinite programming relaxation (I)

- Quantum bilinear program:

idea: relaxation of this bilinear form

$$p^*_{\mathrm{succ}}(W,k) := \underset{(\mathcal{H},\psi,E,D)}{\text{maximize}} \quad \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\langle\psi|E(x|i)\otimes D(i|y)|\psi\rangle \le \langle\psi|E(x|i)\cdot D(y|j)|\psi\rangle$$

$$\text{with } \big[E(x,i), D(y,j)\big] = 0$$

$$\text{subject to} \quad \sum_x E(x|i) = 1_{\mathcal{H}} \quad \forall i\in[k], \quad \sum_i D(i|y) = 1_{\mathcal{H}} \quad \forall y \in Y$$

$$0 \le E(x|i) \le 1_{\mathcal{H}}\ \forall (x,i)\in X\times[k], \quad 0 \le D(i|y) \le 1_{\mathcal{H}}\ \forall (i,y) \in [k]\times Y\,.$$

- First step: see as the part of the upper-right block of the Gram matrix

$$\Omega = \sum_{u,v}\langle\psi|X_u X_v|\psi\rangle|u\rangle\langle v| \quad \text{with} \quad X_u = \begin{cases} E(x|i) & u=(i,x)\\ D(j|y) & u=(j,y)\end{cases}$$

for i=j

$$\Omega = \left(\begin{array}{c|c} \langle\psi|E(x|i)\cdot E(x'|i')|\psi\rangle & \langle\psi|E(x|i)\cdot D(y|j)|\psi\rangle \\ \hline \langle\psi|E(x'|i')\cdot D(y'|j')|\psi\rangle & \langle\psi|D(y|j)\cdot D(y'|j')]|\psi\rangle \end{array}\right)$$

- Original constraints can be formulated as positivity conditions on $\Omega:\ \mathrm{sdp}_1(W,k)$

---

*motivated by: "**NPA hierarchy**" (Bell inequalities)*

[Lasserre, SIAM (2001)], [Parrilo, Math. Program. (2003)], [Navascues et al., PRL (2007)], [Doherty et al., IEEE CCC (2008)], [Navascues et al., NJP (2008)], [Pironio et al., SIAM (2010)]

First level semidefinite programming relaxation (II)

- First level relaxation: $p_{\text{succ}}(W,k) \leq p^*_{\text{succ}}(W,k) \leq \text{sdp}_1(W,k)$

$$
\begin{aligned}
\text{sdp}_1(W,k) = \underset{\Omega}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\Omega_{(i,x),(i,y)} \\
\text{subject to} \quad & \Omega \in \text{Pos}(1+k|X|+k|Y|)\,, \quad \Omega_{\emptyset,\emptyset} = 1 \quad \text{with } \emptyset \text{ the empty symbol} \\
& \Omega_{u,v} \geq 0 \quad \forall u,v \in X\times[k]\cup Y\times[k]\cup\{\emptyset\} \\
& \sum_x \Omega_{w,(i,x)} = \Omega_{w,\emptyset} \quad \forall i\in[k]\,, w\in X\times[k]\cup Y\times[k]\cup\{\emptyset\} \\
& \sum_i \Omega_{w,(i,y)} = \Omega_{w,\emptyset} \quad \forall y\in Y, w\in X\times[k]\cup Y\times[k]\cup\{\emptyset\}\,.
\end{aligned}
$$

First level semidefinite programming relaxation (II)

- First level relaxation: $p_{\text{succ}}(W,k) \leq p^*_{\text{succ}}(W,k) \leq \text{sdp}_1(W,k)$

$$
\begin{aligned}
\text{sdp}_1(W,k) = \underset{\Omega}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\Omega_{(i,x),(i,y)} \\
\text{subject to} \quad & \Omega \in \text{Pos}(1+k|X|+k|Y|)\,, \quad \Omega_{\emptyset,\emptyset} = 1 \quad \text{with } \emptyset \text{ the empty symbol} \\
& \Omega_{u,v} \geq 0 \quad \forall u,v \in X\times[k]\cup Y\times[k]\cup\{\emptyset\} \\
& \sum_x \Omega_{w,(i,x)} = \Omega_{w,\emptyset} \quad \forall i\in[k]\,, w\in X\times[k]\cup Y\times[k]\cup\{\emptyset\} \\
& \sum_i \Omega_{w,(i,y)} = \Omega_{w,\emptyset} \quad \forall y\in Y, w\in X\times[k]\cup Y\times[k]\cup\{\emptyset\}\,.
\end{aligned}
$$

First level semidefinite programming relaxation (II)

- First level relaxation: $p_{\text{succ}}(W,k) \leq p^*_{\text{succ}}(W,k) \leq \text{sdp}_1(W,k)$

$$
\begin{aligned}
\text{sdp}_1(W,k) = \underset{\Omega}{\text{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\Omega_{(i,x),(i,y)} \\
\text{subject to} \quad & \Omega \in \text{Pos}(1+k|X|+k|Y|)\,, \quad \Omega_{\emptyset,\emptyset} = 1 \quad \text{with } \emptyset \text{ the empty symbol} \\
& \Omega_{u,v} \geq 0 \quad \forall u,v \in X\times[k]\cup Y\times[k]\cup\{\emptyset\} \\
& \sum_x \Omega_{w,(i,x)} = \Omega_{w,\emptyset} \quad \forall i\in[k]\,, w\in X\times[k]\cup Y\times[k]\cup\{\emptyset\} \\
& \sum_i \Omega_{w,(i,y)} = \Omega_{w,\emptyset} \quad \forall y\in Y, w\in X\times[k]\cup Y\times[k]\cup\{\emptyset\}\,.
\end{aligned}
$$

- Going back to our example:

$$p_{\text{succ}}(Z,2) = \frac{5}{6} \approx 0.833$$

(known before, with two-dimensional assistance)

$$p^*_{\text{succ}}(Z,2) \geq \frac{2+2^{-1/2}}{3} \approx 0.902$$

$$
Z = \begin{pmatrix}
1/3 & 1/3 & 0 & 0 \\
0 & 0 & 1/3 & 1/3 \\
1/3 & 0 & 1/3 & 0 \\
0 & 1/3 & 0 & 1/3 \\
1/3 & 0 & 0 & 1/3 \\
0 & 1/3 & 1/3 & 0
\end{pmatrix}
$$

First level semidefinite programming relaxation (II)

- First level relaxation: $p_{\mathrm{succ}}(W,k) \leq p^*_{\mathrm{succ}}(W,k) \leq \mathrm{sdp}_1(W,k)$

$$
\begin{aligned}
\mathrm{sdp}_1(W,k) = \underset{\Omega}{\mathrm{maximize}} \quad & \frac{1}{k}\sum_{x,y,i} W_{X\to Y}(y|x)\Omega_{(i,x),(i,y)} \\
\text{subject to} \quad & \Omega \in \mathrm{Pos}(1+k|X|+k|Y|)\,, \quad \Omega_{\emptyset,\emptyset}=1 \quad \text{with } \emptyset \text{ the empty symbol} \\
& \Omega_{u,v} \geq 0 \quad \forall u,v \in X\times[k]\cup Y\times[k]\cup\{\emptyset\} \\
& \sum_x \Omega_{w,(i,x)} = \Omega_{w,\emptyset} \quad \forall i\in[k]\,, w\in X\times[k]\cup Y\times[k]\cup\{\emptyset\} \\
& \sum_i \Omega_{w,(i,y)} = \Omega_{w,\emptyset} \quad \forall y\in Y, w\in X\times[k]\cup Y\times[k]\cup\{\emptyset\}\,.
\end{aligned}
$$

- Going back to our example:

$$p_{\mathrm{succ}}(Z,2) = \frac{5}{6} \approx 0.833$$

(known before, with two-dimensional assistance)

$$p^*_{\mathrm{succ}}(Z,2) \geq \frac{2+2^{-1/2}}{3} \approx 0.902$$

$$Z = \begin{pmatrix} 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 0 \\ 0 & 1/3 & 0 & 1/3 \\ 1/3 & 0 & 0 & 1/3 \\ 0 & 1/3 & 1/3 & 0 \end{pmatrix}$$

- Relaxation: $p^*_{\mathrm{succ}}(Z,2) \leq \mathrm{sdp}_1(Z,2) \approx 0.908 = \frac{1}{2} + \frac{1}{\sqrt{6}}$
- Four-dimensional assistance: $p^*_{\mathrm{succ}}(Z,2) \geq \frac{1}{2} + \frac{1}{\sqrt{6}}$

Conclusions

- Fruitful interplay of optimization theory and (quantum) information theory
- Going quantum (= adding quantum assistance to classical tasks) roughly means that commutative variables have to replaced by non-commutative ones
- Optimization becomes harder, but semi-definite programming can be used to obtain converging upper bounds
- Many more directions to explore: fully quantum scenario, other information theoretic protocolls,…

Thanks!