Quantum Simpson’s Paradox
and high order Bell-Tsirelson Inequalities

Yaoyun Shi

University of Michigan

June 3, 2013
Outline

Simpson’s Paradox

Models for Quantum Simpson’s Paradox

Main results
  The construction
  Limit on quantum inversion

Discussions
Outline

Simpson’s Paradox

Models for Quantum Simpson’s Paradox

Main results
   The construction
   Limit on quantum inversion

Discussions
Outline

Simpson’s Paradox

Models for Quantum Simpson’s Paradox

Main results
  The construction
  Limit on quantum inversion

Discussions
Outline

Simpson’s Paradox

Models for Quantum Simpson’s Paradox

Main results
  The construction
  Limit on quantum inversion

Discussions
### US Civil Rights Act of 1964

<table>
<thead>
<tr>
<th></th>
<th>Democrat</th>
<th>Republican</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All</strong></td>
<td>61%(152/248)</td>
<td>80%(138/172)</td>
</tr>
</tbody>
</table>

**Table:** Vote totals
### US Civil Rights Act of 1964

<table>
<thead>
<tr>
<th>House</th>
<th>Democrat</th>
<th>Republican</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern</td>
<td>94%(145/154)</td>
<td>85%(138/162)</td>
</tr>
<tr>
<td>Southern</td>
<td>7%(7/94)</td>
<td>0%(0/10)</td>
</tr>
<tr>
<td>All</td>
<td>61%(152/248)</td>
<td>80%(138/172)</td>
</tr>
</tbody>
</table>

**Table**: Vote totals
Did Berkeley discriminate women in graduate admission?

<table>
<thead>
<tr>
<th>Department</th>
<th><strong>Men</strong></th>
<th></th>
<th><strong>Women</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Applicants</td>
<td>Admitted</td>
<td>Applicants</td>
<td>Admitted</td>
</tr>
<tr>
<td>All</td>
<td>8442</td>
<td>44%</td>
<td>4321</td>
<td>35%</td>
</tr>
</tbody>
</table>

**Table:** Berkeley graduate admission Fall 1973 (Bickel *et al.*, 1973)
Did Berkeley discriminate women in graduate admission?

<table>
<thead>
<tr>
<th>Department</th>
<th>Men Applicants</th>
<th>Men Admitted</th>
<th>Women Applicants</th>
<th>Women Admitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>825</td>
<td>62%</td>
<td>108</td>
<td>82%</td>
</tr>
<tr>
<td>B</td>
<td>560</td>
<td>63%</td>
<td>25</td>
<td>68%</td>
</tr>
<tr>
<td>C</td>
<td>325</td>
<td>37%</td>
<td>593</td>
<td>34%</td>
</tr>
<tr>
<td>D</td>
<td>417</td>
<td>33%</td>
<td>375</td>
<td>35%</td>
</tr>
<tr>
<td>E</td>
<td>191</td>
<td>28%</td>
<td>393</td>
<td>24%</td>
</tr>
<tr>
<td>F</td>
<td>272</td>
<td>6%</td>
<td>341</td>
<td>7%</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>All</td>
<td>8442</td>
<td>44%</td>
<td>4321</td>
<td>35%</td>
</tr>
</tbody>
</table>

Table: Berkeley graduate admission Fall 1973 (Bickel et al., 1973)
P. J. Bickel, E. A. Hammel, and J. W. O’Connell (Science 187 (4175), 1973):

*Measuring bias is harder than is usually assumed, and the evidence is sometimes contrary to expectation.*
### Simpson’s thought experiment

<table>
<thead>
<tr>
<th>Gender</th>
<th>Untreated</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>Survived</td>
</tr>
<tr>
<td>Men</td>
<td>7</td>
<td>57%(4)</td>
</tr>
<tr>
<td>Women</td>
<td>5</td>
<td>40%(2)</td>
</tr>
</tbody>
</table>

**Table:** E. H. Simpson, 1951
## Simpson’s thought experiment

<table>
<thead>
<tr>
<th>Gender</th>
<th>Untreated</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>Survived</td>
</tr>
<tr>
<td>Men</td>
<td>7</td>
<td>57%(4)</td>
</tr>
<tr>
<td>Women</td>
<td>5</td>
<td>40%(2)</td>
</tr>
<tr>
<td>All</td>
<td>12</td>
<td>50%(6)</td>
</tr>
</tbody>
</table>

Table: E. H. Simpson, 1951

### Simpson's Paradox Models for Quantum Simpson’s Paradox

- **C1: Simpson's**
  - Untreated: $R_c^m = 4/7$
  - Treated: $R_t^m = 8/13$
  - Untreated: $R_c^f = 2/5$
  - Treated: $R_t^f = 12/27$

- **C2: Modified Simpson's**
  - Untreated: $R_c = 6/12$
  - Treated: $R_t = 20/40$

- **Q1: Quantum reversal**
  - Improves recovery rate:
    - Untreated: $R_c$ to $R_t$
    - Treated: $R_c$ to $R_t$

- **Q2: Quantum reversal**
  - Decreases recovery rate:
    - Untreated: $R_c^m$ to $R_t^m$
    - Treated: $R_c^f$ to $R_t^f$
### Simpson’s thought experiment

<table>
<thead>
<tr>
<th>Gender</th>
<th>Untreated</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients</td>
<td>Survived</td>
</tr>
<tr>
<td>Men</td>
<td>7</td>
<td>57%(4)</td>
</tr>
<tr>
<td>Women</td>
<td>5</td>
<td>40%(2)</td>
</tr>
<tr>
<td>All</td>
<td>12</td>
<td>50%(6)</td>
</tr>
</tbody>
</table>

Table: E. H. Simpson, 1951
Simpson’s Paradox aka Simpson-Yule Effect

A trend in the *aggregated* data may be reversed within *partitions* of the data.

- Karl Pearson, *et al.*, in 1899; Udny Yule in 1903; E. H. Simpson in 1953
- Numerous real-life examples
- Implications in decision making: which table should we use, aggregated or partitioned?
Simpson’s Paradox aka Simpson-Yule Effect

A trend in the *aggregated* data may be reversed within *partitions* of the data.

- Karl Pearson, *et al.*, in 1899; Udny Yule in 1903; E. H. Simpson in 1953
- Numerous real life examples
- Implications in decision making: which table should we use, aggregated or partitioned?
  - Depends on the causal dependency.
  - Statistics not sufficient for deciding which table.
Simpson’s Paradox aka Simpson-Yule Effect

A trend in the *aggregated* data may be reversed within *partitions* of the data.

- Karl Pearson, *et al.*, in 1899; Udny Yule in 1903; E. H. Simpson in 1953
- Numerous real life examples
- Implications in decision making: which table should we use, aggregated or partitioned?
  - Depends on the causal dependency.
  - Statistics not sufficient for deciding which table.
A trend in the aggregated data may be reversed within partitions of the data.

- Karl Pearson, *et al.*, in 1899; Udny Yule in 1903; E. H. Simpson in 1953
- Numerous real life examples
- Implications in decision making: which table should we use, aggregated or partitioned?
  - Depends on the causal dependency.
  - Statistics not sufficient for deciding which table.
Simpson’s Paradox aka Simpson-Yule Effect

A trend in the *aggregated* data may be reversed within *partitions* of the data.

- Karl Pearson, *et al.*, in 1899; Udny Yule in 1903; E. H. Simpson in 1953
- Numerous real life examples
- Implications in decision making: which table should we use, aggregated or partitioned?
  - Depends on the causal dependency.
  - Statistics not sufficient for deciding which table.
Simpson’s Paradox aka Simpson-Yule Effect

A trend in the *aggregated* data may be reversed within *partitions* of the data.

- Karl Pearson, *et al.*, in 1899; Udny Yule in 1903; E. H. Simpson in 1953
- Numerous real life examples
- Implications in decision making: which table should we use, aggregated or partitioned?
  - Depends on the causal dependency.
  - Statistics not sufficient for deciding which table.
Simpson’s Paradox aka Simpson-Yule Effect

A trend in the *aggregated* data may be reversed within *partitions* of the data.

- Karl Pearson, *et al.*, in 1899; Udny Yule in 1903; E. H. Simpson in 1953
- Numerous real-life examples
- Implications in decision making: which table should we use, aggregated or partitioned?
  - Depends on the causal dependency.
  - Statistics not sufficient for deciding which table.
Paris’ formulation (2012)

Start with two binary measurements

\{A, I - A\}, \quad \text{and} \quad \{B, I - B\},

and two pure states \(|\psi_1\rangle, |\psi_2\rangle\), s.t.

\langle \psi_i | A | \psi_i \rangle > \langle \psi_i | B | \psi_i \rangle, \quad i = 1, 2,

yet for some combined \(\psi\),

\[ \text{Tr } A \psi < \text{Tr } B \psi. \]

• Classical lurking variable

\[ \psi = \alpha |\psi_1\rangle \langle \psi_1| + \beta |\psi_2\rangle \langle \psi_2|. \]

• Quantum lurking variable

\[ |\psi\rangle = \alpha |\psi_1\rangle + \beta |\psi_2\rangle. \]
Paris’ formulation (2012)

Start with two binary measurements

\{A, \mathbb{I} - A\}, \quad \text{and} \quad \{B, \mathbb{I} - B\},

and two pure states \(|\psi_1\rangle, |\psi_2\rangle\), s.t.

\[ \langle \psi_i | A | \psi_i \rangle > \langle \psi_i | B | \psi_i \rangle, \quad i = 1, 2, \]

yet for some \textit{combined} \(\psi\),

\[ \text{Tr } A\psi < \text{Tr } B\psi. \]

- \textbf{Classical} lurking variable

\[ \psi = \alpha |\psi_1\rangle\langle \psi_1| + \beta |\psi_2\rangle\langle \psi_2|. \]

- \textbf{Quantum} lurking variable

\[ |\psi\rangle = \alpha |\psi_1\rangle + \beta |\psi_2\rangle. \]
Paris’ formulation (2012)

Start with two binary measurements

\[ \{A, \mathbb{I} - A\}, \quad \text{and} \quad \{B, \mathbb{I} - B\}, \]

and two pure states \(|\psi_1\rangle, |\psi_2\rangle\), s.t.

\[ \langle \psi_i | A | \psi_i \rangle > \langle \psi_i | B | \psi_i \rangle, \quad i = 1, 2, \]

yet for some combined \(\psi\),

\[ \text{Tr } A \psi < \text{Tr } B \psi. \]

- **Classical** lurking variable

\[ \psi = \alpha |\psi_1\rangle\langle \psi_1| + \beta |\psi_2\rangle\langle \psi_2|. \]

- **Quantum** lurking variable

\[ |\psi\rangle = \alpha |\psi_1\rangle + \beta |\psi_2\rangle. \]
This work: non-classical reversal?

- The *convexity property*:

\[ R_c \in [R_c^m, R_c^f] \quad \text{and} \quad R_t \in [R_t^m, R_c^f]. \]
This work: non-classical reversal?

- The *convexity property*:

$$R_c \in [R_c^m, R_c^f] \quad \text{and} \quad R_t \in [R_t^m, R_t^f].$$
This work: non-classical reversal?

- The *convexity property*:
  \[ R_c \in [R^m_c, R^f_c] \quad \text{and} \quad R_t \in [R^m_t, R^f_t]. \]

- Reversal is possible if and only if \([R^m_c, R^f_c]\) and \([R^m_t, R^f_t]\) have nontrivial intersection.
This work: non-classical reversal?

- The *convexity property*:
  \[ R_c \in [R_c^m, R_c^f] \text{ and } R_t \in [R_t^m, R_t^f]. \]

- Reversal is possible if and only if
  \[ [R_c^m, R_c^f] \text{ and } [R_t^m, R_t^f] \] have nontrivial intersection.

- **Quantum** reversal when the intersection is empty?

---

**C: Classical reversal possible**

**Cl: Classically no reversal**
Measurement scenario

Definition

A measurement scenario is a tuple

\[ \mathcal{M} := (|\phi\rangle, G, E, R), \]

where

- \(|\phi\rangle\): a quantum state,
- \(G, E, R\): binary measurements on \(|\phi\rangle\) for the outcome variables
Measurement scenario

Definition
A measurement scenario is a tuple

\[ \mathcal{M} := (|\phi\rangle, G, E, R), \]

where

- \(|\phi\rangle\): a quantum state,
- \(G, E, R\): binary measurements on \(|\phi\rangle\) for the outcome variables
  - Gender \(G \in \{F := \text{Female}, M := \text{Male}\}\),
  - Experiment \(E \in \{U := \text{Untreated}, T := \text{Treated}\}\),
  - Result \(R \in \{D := \text{Dead}, A := \text{Alive}\}\).
Measurement scenario

Definition

A measurement scenario is a tuple

\[ M := (|\phi\rangle, G, E, R), \]

where

- \(|\phi\rangle\): a quantum state,
- \(G, E, R\): binary measurements on \(|\phi\rangle\) for the outcome variables
  - Gender \(G \in \{F := \text{Female}, M := \text{Male}\}\),
  - Experiment \(E \in \{U := \text{Untreated}, T := \text{Treated}\}\),
  - Result \(R \in \{D := \text{Dead}, A := \text{Alive}\}\).
Measurement scenario

Definition
A measurement scenario is a tuple
\[ \mathcal{M} := (|\phi\rangle, G, E, R), \]
where
- \(|\phi\rangle\): a quantum state,
- \(G, E, R\): binary measurements on \(|\phi\rangle\) for the outcome variables
  - Gender \(G \in \{F := \text{Female}, M := \text{Male}\}\),
  - Experiment \(E \in \{U := \text{Untreated}, T := \text{Treated}\}\),
  - Result \(R \in \{D := \text{Dead}, A := \text{Alive}\}\).
Simpson’s Paradox Models for Quantum Simpson’s Paradox

Main results

Discussions

Measurement scenario

Definition

A measurement scenario is a tuple

\[ \mathcal{M} := (|\phi\rangle, G, E, R), \]

where

- \( |\phi\rangle \): a quantum state,
- \( G, E, R \): binary measurements on \( |\phi\rangle \) for the outcome variables
  - Gender \( G \in \{ F := \text{Female}, M := \text{Male} \} \),
  - Experiment \( E \in \{ U := \text{Untreated}, T := \text{Treated} \} \),
  - Result \( R \in \{ D := \text{Dead}, A := \text{Alive} \} \).
Measurement scenario

Definition
A measurement scenario is a tuple

\[ \mathcal{M} := (|\phi\rangle, G, E, R), \]

where

- \( |\phi\rangle \): a quantum state,
- \( G, E, R \): binary measurements on \( |\phi\rangle \) for the outcome variables
  - Gender \( G \in \{ F := \text{Female}, M := \text{Male} \} \),
  - Experiment \( E \in \{ U := \text{Untreated}, T := \text{Treated} \} \),
  - Result \( R \in \{ D := \text{Dead}, A := \text{Alive} \} \).
Two experiments

• **Aggregated Experiment**: Measure $E$, then $R$.
  Define the control and the treatment survival rates $R_c$ and $R_t$:
  
  \[
  R_c := \text{Prob}[R = \text{Alive}|E = \text{Untreated}],
  \]
  \[
  R_t := \text{Prob}[R = \text{Alive}|E = \text{Treated}].
  \]

• **Partitioned Experiment**: Measure $G$, then $E$, then $R$.
  Define the partitioned survival rates $R^m_c$, $R^f_c$, $R^m_t$, $R^f_t$:
  
  \[
  R^m_c := \text{Prob}[R = \text{Alive}|G = \text{Male}, E = \text{Untreated}]
  \]
  ...
Two experiments

- **Aggregated Experiment**: Measure $E$, then $R$.
  Define the control and the treatment survival rates $R_c$ and $R_t$:

  \[
  R_c := \text{Prob}[R = \text{Alive} | E = \text{Untreated}], \\
  R_t := \text{Prob}[R = \text{Alive} | E = \text{Treated}].
  \]

- **Partitioned Experiment**: Measure $G$, then $E$, then $R$.
  Define the partitioned survival rates $R_{c}^{m}$, $R_{c}^{f}$, $R_{t}^{m}$, $R_{t}^{f}$:

  \[
  R_{c}^{m} := \text{Prob}[R = \text{Alive} | G = \text{Male}, E = \text{Untreated}]
  \]
  \[
  \ldots
  \]
Two experiments

- **Aggregated Experiment**: Measure $E$, then $R$.
  Define the control and the treatment survival rates $R_c$ and $R_t$:

  \[
  R_c \ := \ \text{Prob}[R = \text{Alive} | E = \text{Untreated}], \\
  R_t \ := \ \text{Prob}[R = \text{Alive} | E = \text{Treated}].
  \]

- **Partitioned Experiment**: Measure $G$, then $E$, then $R$.
  Define the partitioned survival rates $R_m^c$, $R_m^t$, $R_f^c$, $R_f^t$:

  \[
  R_m^c \ := \ \text{Prob}[R = \text{Alive} | G = \text{Male}, E = \text{Untreated}]
  \]
  ...
Two experiments

• **Aggregated Experiment:** Measure $E$, then $R$.
  Define the control and the treatment survival rates $R_c$ and $R_t$:

  \[
  R_c := \Pr[R = \text{Alive}|E = \text{Untreated}], \quad R_t := \Pr[R = \text{Alive}|E = \text{Treated}].
  \]

• **Partitioned Experiment:** Measure $G$, then $E$, then $R$.
  Define the partitioned survival rates $R_{c_m}, R_{c_f}, R_{t_m}, R_{t_f}$:

  \[
  R_{c_m} := \Pr[R = \text{Alive}|G = \text{Male}, E = \text{Untreated}]
  \]

  ...
Two experiments

- **Aggregated Experiment**: Measure $E$, then $R$.
  Define the control and the treatment survival rates $R_c$ and $R_t$:
  \[
  R_c := \text{Prob}[R = \text{Alive} | E = \text{Untreated}], \\
  R_t := \text{Prob}[R = \text{Alive} | E = \text{Treated}].
  \]

- **Partitioned Experiment**: Measure $G$, then $E$, then $R$.
  Define the partitioned survival rates $R_{c}^{m}$, $R_{c}^{f}$, $R_{t}^{m}$, $R_{t}^{f}$:
  \[
  R_{c}^{m} := \text{Prob} [R = \text{Alive} | G = \text{Male}, E = \text{Untreated}] \\
  \ldots
  \]
Classical case

- Notation: $G^M, G^F$, etc., projections to the corresponding outcome subspace.

- $R_c = \alpha \cdot R^m_c + \beta \cdot R^f_c$, where

$$\alpha := \frac{\langle \phi | G^M E^U R^A E^U G^M | \phi \rangle}{\langle \phi | E^U R^A E^U | \phi \rangle}, \quad \beta := \frac{\langle \phi | G^F E^U R^A E^U G^F | \phi \rangle}{\langle \phi | E^U R^A E^U | \phi \rangle}.$$

- If the measurements commute, $\alpha + \beta = 1$: the convexity property holds.

- Thus can model classical instances with separate Hilbert space for Gender, Experiment, and Result.
Classical case

- Notation: $G^M, G^F$, etc., projections to the corresponding outcome subspace.

- $R_c = \alpha \cdot R_m^c + \beta \cdot R_f^c$, where

  $$\alpha := \frac{\langle \phi | G^M E^U R^A E^U G^M | \phi \rangle}{\langle \phi | E^U R^A E^U | \phi \rangle}, \quad \beta := \frac{\langle \phi | G^F E^U R^A E^U G^F | \phi \rangle}{\langle \phi | E^U R^A E^U | \phi \rangle}.$$  

- If the measurements commute, $\alpha + \beta = 1$: the convexity property holds.

- Thus can model classical instances with separate Hilbert space for Gender, Experiment, and Result.
Classical case

- Notation: $G^M, G^F$, etc., projections to the corresponding outcome subspace.
- $R_c = \alpha \cdot R^m_c + \beta \cdot R^f_c$, where
  \[
  \alpha := \frac{\langle \phi | G^M E^U R^A E^U G^M | \phi \rangle}{\langle \phi | E^U R^A E^U | \phi \rangle}, \quad \beta := \frac{\langle \phi | G^F E^U R^A E^U G^F | \phi \rangle}{\langle \phi | E^U R^A E^U | \phi \rangle}.
  \]
- If the measurements commute, $\alpha + \beta = 1$: the convexity property holds.
- Thus can model classical instances with separate Hilbert space for Gender, Experiment, and Result.
Classical case

- Notation: $G^M, G^F$, etc., projections to the corresponding outcome subspace.
- $R_c = \alpha \cdot R_c^m + \beta \cdot R_c^f$, where
  \[
  \alpha := \frac{\langle \phi | G^M E^U R^A E^U G^M | \phi \rangle}{\langle \phi | E^U R^A E^U | \phi \rangle}, \quad \beta := \frac{\langle \phi | G^F E^U R^A E^U G^F | \phi \rangle}{\langle \phi | E^U R^A E^U | \phi \rangle}.
  \]
- If the measurements commute, $\alpha + \beta = 1$: the convexity property holds.
- Thus can model classical instances with separate Hilbert space for Gender, Experiment, and Result.
Non-commuting case

- **Not true** for a general $P$: $G^M P G^M + G^F P G^F \leq P$.

- Thus $\alpha + \beta$ may not $\leq 1$ (actually can be arbitrary as long as nonnegative.)

- Make it possible for non-classical reversal.
Non-commuting case

- **Not true** for a general $P$: $G^M P G^M + G^F P G^F \leq P$.
- Thus $\alpha + \beta$ may not $\leq 1$ (actually can be arbitrary as long as nonnegative.)
- Make it possible for non-classical reversal.
Non-commuting case

- Not true for a general $P$: $G^M P G^M + G^F P G^F \leq P$.
- Thus $\alpha + \beta$ may not $\leq 1$ (actually can be arbitrary as long as nonnegative.)
- Make it possible for non-classical reversal.
Quantifying non-classical reversal

Definition
The **reversal quantity** \( S := S_M \) of a measurement scenario \( M \) is

\[
S = (R_t^f + R_t^m - R_t) - (R_c^f + R_c^m - R_c).
\]

Write \( S = d_t - d_c \), where

\[
d_t = R_t^f + R_t^m - R_t, d_c = R_c^f + R_c^m - R_c.
\]

If the convexity property holds, \( 0 \leq d_t, d_c \leq 1 \), thus \( |S| \leq 1 \).

**Proposition**

For all classical measurement scenario (i.e. all measurements commute), \( |S| \leq 1 \).
Quantifying non-classical reversal

**Definition**

The reversal quantity $S := S_M$ of a measurement scenario $\mathcal{M}$ is

$$S = (R_t^f + R_t^m - R_t) - (R_c^f + R_c^m - R_c).$$

Write $S = d_t - d_c$, where

$$d_t = R_t^f + R_t^m - R_t, d_c = R_c^f + R_c^m - R_c.$$

If the convexity property holds, $0 \leq d_t, d_c \leq 1$, thus $|S| \leq 1$.

**Proposition**

*For all classical measurement scenario (i.e. all measurements commute), $|S| \leq 1$.***
Quantifying non-classical reversal

Definition
The reversal quantity $S := S_M$ of a measurement scenario $M$ is

$$S = (R_t^f + R_t^m - R_t) - (R_c^f + R_c^m - R_c).$$

Write $S = d_t - d_c$, where

$$d_t = R_t^f + R_t^m - R_t, d_c = R_c^f + R_c^m - R_c.$$ If the convexity property holds, $0 \leq d_t, d_c \leq 1$, thus $|S| \leq 1$.

Proposition
For all classical measurement scenario (i.e. all measurements commute), $|S| \leq 1$. 

CI: Classically no reversal

<table>
<thead>
<tr>
<th>Recovery Rate</th>
<th>Untreated</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_c^f$</td>
<td>0.40</td>
<td>0.44</td>
</tr>
<tr>
<td>$R_c^m$</td>
<td>0.50</td>
<td>0.57</td>
</tr>
<tr>
<td>$R_t^f$</td>
<td>0.62</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Q1: Quantum reversal

<table>
<thead>
<tr>
<th>Recovery Rate</th>
<th>Untreated</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_c^f$</td>
<td>0.00</td>
<td>0.33</td>
</tr>
<tr>
<td>$R_c^m$</td>
<td>0.50</td>
<td>0.99</td>
</tr>
<tr>
<td>$R_t^f$</td>
<td>0.99</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Q2: Quantum reversal

<table>
<thead>
<tr>
<th>Recovery Rate</th>
<th>Untreated</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_c^f$</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$R_c^m$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$R_t^f$</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$R_t^m$</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
Question: what values possible for $S$ in general?

Theorem

For each sufficiently small $\epsilon > 0$, there exists a measurement scenario such that $|S| \geq 2 - \epsilon$. 
Question: what values possible for $S$ in general?

Theorem

For each sufficiently small $\epsilon > 0$, there exists a measurement scenario such that $|S| \geq 2 - \epsilon$. 

Extremal quantum reversal
Question: what values possible for $S$ in general?

**Theorem**

For each sufficiently small $\epsilon > 0$, there exists a measurement scenario such that $|S| \geq 2 - \epsilon$. 

---

<table>
<thead>
<tr>
<th>Recovery Rate</th>
<th>Untreated</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.33</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>0.40</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>0.44</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>0.50</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>0.57</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>0.62</td>
<td>32</td>
<td>24</td>
</tr>
</tbody>
</table>

---

The "quantum drug" is life-saving for both men and women but
Question: what values possible for $S$ in general?

Theorem

For each sufficiently small $\epsilon > 0$, there exists a measurement scenario such that $|S| \geq 2 - \epsilon$.

The “quantum drug” is life-saving for both men and women but deadly for all.
Theorem

For any measurement scenario, $|S| < 2$. 
Constructing the quantum extremal reversal

Put probabilities on the cube

- Vertexes: partitioned measurement probabilities.
- Edges: aggregated measurement probabilities.
Constructing the quantum extremal reversal

Put probabilities on the cube

- **Vertexes**: partitioned measurement probabilities.
- **Edges**: aggregated measurement probabilities.

- $R_1 = 0$
- $R_1 = R'' = \sqrt{\frac{1}{2} - c} = 1 - c^2$
- $R_1 = R'' = \sqrt{\frac{1}{2} - 2\epsilon}$
- $R_1 = R'' = \sqrt{\frac{1}{2} - 2\epsilon}$
- $\epsilon = R + R'' = R_1 \approx 2 - 2\epsilon + 2\epsilon$
- $\epsilon = R_1 + R'' = R_1 \approx 2 - 2\epsilon + 2\epsilon$
- $\epsilon = R_1 + R'' = R_1 \approx 2 - 2\epsilon + 2\epsilon$
- $S = \epsilon_1 - d_1 \approx 2 - 2\epsilon - 2\epsilon$
Constructing the quantum extremal reversal

Put probabilities on the cube

- Vertexes: partitioned measurement probabilities.
- Edges: aggregated measurement probabilities.

\[ R_1 = \frac{\epsilon^2}{4} \]
\[ R'_1 = R''_1 = \frac{\epsilon^2}{4 + 4} \approx 1 - \epsilon + \epsilon^2 \]
\[ R_1 = R'_1 = R''_1 = 1/2 \]

Thus
\[ d_1 = R_1 + R''_1 = R_1 + 2 - 2\epsilon + 2\epsilon \]
\[ d_2 = R_1 + R''_1 = R_1 = \epsilon/2 \]
\[ S = d_1 - d_2 = 2 - 2\epsilon \]
Constructing the quantum extremal reversal

Put probabilities on the cube

- Vertexes: partitioned measurement probabilities.
- Edges: aggregated measurement probabilities.

- \( R_t = 0 \)
- \( R_t^f = R_t^m = \frac{\epsilon}{4} + \frac{\epsilon^2}{4} \approx 1 - \epsilon + \epsilon^2 \)
- \( R_c = \frac{\epsilon}{1 + \epsilon} \approx \epsilon - \epsilon^2 \)
- \( R_c^f = R_c^m = \frac{\epsilon}{4} + \frac{\epsilon^2 + 2}{4} \approx \frac{\epsilon}{2} - \frac{\epsilon^2}{4} \)
- Thus
  \[ d_t = R_t^f + R_t^m - R_t \approx 2 - 2\epsilon + 2\epsilon^2, \]
  \[ d_c = R_c^f + R_c^m - R_c \approx \epsilon^2 / 2, \]
  \[ S = d_t - d_s \approx 2 - 2\epsilon \rightarrow 2. \]
Constructing the quantum extremal reversal

Put probabilities on the cube

- Vertexes: partitioned measurement probabilities.
- Edges: aggregated measurement probabilities.

- \( R_t = 0 \)
- \( R_t^f = R_t^m = \frac{\epsilon}{4} + \frac{\epsilon^2}{4} \approx 1 - \epsilon + \epsilon^2 \)
- \( R_c = \frac{\epsilon}{4 + \epsilon} \approx \epsilon - \epsilon^3 \)
- \( R_c^f = R_c^m = \frac{\epsilon}{4} + \frac{\epsilon^2 + \epsilon^4}{4} \approx \epsilon - \epsilon^2 \)
- Thus
  \[
  d_t = R_t^f + R_t^m - R_t \approx 2 - 2\epsilon + 2\epsilon^2, \\
  d_c = R_c^f + R_c^m - R_c \approx \epsilon^2 / 2, \\
  S = d_t - d_c \approx 2 - 2\epsilon \rightarrow 2.
  \]
Constructing the quantum extremal reversal

Put probabilities on the cube

- Vertexes: partitioned measurement probabilities.
- Edges: aggregated measurement probabilities.

- $R_t = 0$
- $R_t^f = R_t^m = \frac{\frac{\epsilon}{4} + \frac{\epsilon^2}{4}}{\epsilon^4 + \epsilon^2} \approx 1 - \epsilon + \epsilon^2$
- $R_c = \frac{\epsilon}{1 + \epsilon} \approx \epsilon - \epsilon^2$
- $R_c^f = R_c^m = \frac{\frac{\epsilon}{4} + \frac{\epsilon^2}{4}}{\epsilon^4 + \epsilon^2 + \frac{\epsilon^4}{4}} \approx \frac{\epsilon}{2} - \frac{\epsilon^2}{4}$

Thus

$d_t = R_t^f + R_t^m - R_t \approx 2 - 2\epsilon + 2\epsilon^2$,

$d_c = R_c^f + R_c^m - R_c \approx \epsilon^2 / 2$,

$S = d_t - d_s \approx 2 - 2\epsilon \rightarrow 2$. 

FIG. 2: Each of the three axes corresponds to one of the three measurements and each face is annotated with the squared lengths (for those projections in three, respectively, measurements, for the measurement outcomes determined by the faces incident to it, with a measurement outcome. Each face, edge, and vertex represent the resulting state after one, two, or...
Constructing the quantum extremal reversal

Put probabilities on the cube
- Vertexes: partitioned measurement probabilities.
- Edges: aggregated measurement probabilities.

- $R_t = 0$
- $R_t^f = R_t^m = \frac{\epsilon}{4 + \epsilon^2} \approx 1 - \epsilon + \epsilon^2$
- $R_c = \frac{\epsilon}{1 + \epsilon} \approx \epsilon - \epsilon^2$
- $R_c^f = R_c^m = \frac{\epsilon}{\frac{\epsilon}{4} + \frac{\epsilon^2 + 2}{4}} \approx \frac{\epsilon}{2} - \frac{\epsilon^2}{4}$

Thus

\[
d_t = R_t^f + R_t^m - R_t \approx 2 - 2\epsilon + 2\epsilon^2,
\]
\[
d_c = R_c^f + R_c^m - R_c \approx \epsilon^2 / 2,
\]
\[
S = d_t - d_s \approx 2 - 2\epsilon \rightarrow 2.
\]
Constructing the quantum extremal reversal

Put probabilities on the cube

- Vertexes: partitioned measurement probabilities.
- Edges: aggregated measurement probabilities.

$$R_t = 0$$

$$R_t^f = R_t^m = \frac{\epsilon}{\epsilon^2 + \frac{\epsilon}{4}} \approx 1 - \epsilon + \epsilon^2$$

$$R_c = \frac{\epsilon}{1+\epsilon} \approx \epsilon - \epsilon^2$$

$$R_c^f = R_c^m = \frac{\epsilon}{\epsilon^2 + \frac{\epsilon}{4} + \frac{\epsilon^2}{2}} \approx \frac{\epsilon}{2} - \frac{\epsilon^2}{4}$$

Thus

$$d_t = R_t^f + R_t^m - R_t \approx 2 - 2\epsilon + 2\epsilon^2,$$

$$d_c = R_c^f + R_c^m - R_c \approx \epsilon^2 / 2,$$

$$S = d_t - d_s \approx 2 - 2\epsilon \to 2.$$
The construction

- **State space:** \( H = V \otimes W \).
  - \( \mathbb{R} \) acts on \( W \), \( \text{dim}(W) = 2 \), spanned by \( |D\rangle, |A\rangle \),
  - \( E, G \) act on \( V \), \( \text{dim}(V) = 4 \).
- Untreated subspace spanned by \( |U\rangle \);
  Treated by \( |T_i\rangle, i = 0, 1, 2 \).
- Female subspace spanned by \( |F_i\rangle, i = 0, 1 \);
  Male by \( |M_i\rangle, i = 0, 1 \).
- Relation between the orthonormal vectors
  - Notation: \( |(\psi_1 \pm \psi_2)\rangle := \sqrt{\frac{1}{2}} (|\psi_1\rangle \pm |\psi_2\rangle) \).
The construction

- **State space:** $H = V \otimes W$.
  - $\mathbb{R}$ acts on $W$, $\dim(W) = 2$, spanned by $|D\rangle, |A\rangle$,
  - $E, G$ act on $V$, $\dim(V) = 4$.

- Untreated subspace spanned by $|U\rangle$;
  Treated by $|T_i\rangle$, $i = 0, 1, 2$.

- Female subspace spanned by $|F_i\rangle$, $i = 0, 1$;
  Male by $|M_i\rangle$, $i = 0, 1$.

- Relation between the orthonormal vectors
  - Notation: $|(\psi_1 \pm \psi_2)\rangle := \sqrt{2} (|\psi_1\rangle \pm |\psi_2\rangle)$.
The construction

- State space: \( H = V \otimes W \).
  - \( \mathbb{R} \) acts on \( W \), \( \text{dim}(W) = 2 \), spanned by \( |D\rangle, |A\rangle \).
  - \( E, G \) act on \( V \), \( \text{dim}(V) = 4 \).
- Untreated subspace spanned by \( |U\rangle \);
  Treated by \( |T_i\rangle, i = 0, 1, 2 \).
- Female subspace spanned by \( |F_i\rangle, i = 0, 1 \);
  Male by \( |M_i\rangle, i = 0, 1 \).
- Relation between the orthonormal vectors
  - \( |(\psi_1 \pm \psi_2)\rangle := \frac{1}{\sqrt{2}} (|\psi_1\rangle \pm |\psi_2\rangle) \).
The construction

- **State space**: $H = V \otimes W$.
  - $\mathbb{R}$ acts on $W$, $\dim(W) = 2$, spanned by $|D\rangle, |A\rangle$.
  - $E, G$ act on $V$, $\dim(V) = 4$.
- **Untreated subspace** spanned by $|U\rangle$;
  Treated by $|T_i\rangle$, $i = 0, 1, 2$.
- **Female subspace** spanned by $|F_i\rangle$, $i = 0, 1$;
  Male by $|M_i\rangle$, $i = 0, 1$.
- **Relation between the orthonormal vectors**
  - Notation: $|(\psi_1 \pm \psi_2)\rangle := \frac{1}{\sqrt{2}} (|\psi_1\rangle \pm |\psi_2\rangle)$.
The construction

- State space: $H = V \otimes W$.
  - $\mathbb{R}$ acts on $W$, $\dim(W) = 2$, spanned by $|D\rangle, |A\rangle$.
  - $E, G$ act on $V$, $\dim(V) = 4$.
- Untreated subspace spanned by $|U\rangle$; Treated by $|T_i\rangle$, $i = 0, 1, 2$.
- Female subspace spanned by $|F_i\rangle$, $i = 0, 1$; Male by $|M_i\rangle$, $i = 0, 1$.
- Relation between the orthonormal vectors
  - Notation: $|(\psi_1 \pm \psi_2)\rangle := \frac{1}{\sqrt{2}} (|\psi_1\rangle \pm |\psi_2\rangle)$.
The construction

- State space: $H = V \otimes W$.
  - $\mathbb{R}$ acts on $W$, $\dim(W) = 2$, spanned by $|D\rangle, |A\rangle$.
  - $E, G$ act on $V$, $\dim(V) = 4$.

- Untreated subspace spanned by $|U\rangle$; Treated by $|T_i\rangle$, $i = 0, 1, 2$.

- Female subspace spanned by $|F_i\rangle$, $i = 0, 1$; Male by $|M_i\rangle$, $i = 0, 1$.

- Relation between the orthonormal vectors
  - Notation: $|(\psi_1 \pm \psi_2)\rangle := \frac{1}{\sqrt{2}} (|\psi_1\rangle \pm |\psi_2\rangle)$.

$$|F_0\rangle := |(U_0 + U_1) + t\rangle, \quad |F_1\rangle := |u_2\rangle$$
$$|M_0\rangle := |(U_0 + U_1) - t\rangle, \quad |M_1\rangle := |U_0 - U_1\rangle.$$
The construction

- State space: $H = V \otimes W$.
  - $\mathbb{R}$ acts on $W$, $\dim(W) = 2$, spanned by $|D\rangle, |A\rangle$.
  - $E, G$ act on $V$, $\dim(V) = 4$.
- Untreated subspace spanned by $|U\rangle$;
  Treated by $|T_i\rangle$, $i = 0, 1, 2$.
- Female subspace spanned by $|F_i\rangle$, $i = 0, 1$;
  Male by $|M_i\rangle$, $i = 0, 1$.
- Relation between the orthonormal vectors
  - Notation: $|(\psi_1 \pm \psi_2)\rangle := \frac{1}{\sqrt{2}} (|\psi_1\rangle \pm |\psi_2\rangle)$.

$$|F_0\rangle := |(U_0 + U_1) + t\rangle, \quad |F_1\rangle := |u_2\rangle, \quad |M_0\rangle := |(U_0 + U_1) - t\rangle, \quad |M_1\rangle := |U_0 - U_1\rangle.$$
The construction

- **State space:** $H = V \otimes W$.
  - $\mathbb{R}$ acts on $W$, $\text{dim}(W) = 2$, spanned by $|D\rangle, |A\rangle$.
  - $E, G$ act on $V$, $\text{dim}(V) = 4$.
- Untreated subspace spanned by $|U\rangle$; Treated by $|T_i\rangle$, $i = 0, 1, 2$.
- Female subspace spanned by $|F_i\rangle$, $i = 0, 1$; Male by $|M_i\rangle$, $i = 0, 1$.
- Relation between the orthonormal vectors
  - Notation: $|(\psi_1 \pm \psi_2)\rangle := \frac{1}{\sqrt{2}} (|\psi_1\rangle \pm |\psi_2\rangle)$.

\[
|F_0\rangle := |(U_0 + U_1) + t\rangle, \quad |F_1\rangle := |u_2\rangle, \\
|M_0\rangle := |(U_0 + U_1) - t\rangle, \quad |M_1\rangle := |U_0 - U_1\rangle.
\]
The construction

- **State space:** $H = V \otimes W$.
  - $\mathbb{R}$ acts on $W$, $\dim(W) = 2$, spanned by $|D\rangle, |A\rangle$.
  - $E, G$ act on $V$, $\dim(V) = 4$.
- Untreated subspace spanned by $|U\rangle$; Treated by $|T_i\rangle$, $i = 0, 1, 2$.
- Female subspace spanned by $|F_i\rangle$, $i = 0, 1$; Male by $|M_i\rangle$, $i = 0, 1$.
- Relation between the orthonormal vectors
  - Notation: $|(\psi_1 \pm \psi_2)\rangle := \frac{1}{\sqrt{2}} (|\psi_1\rangle \pm |\psi_2\rangle)$.

$|F_0\rangle := |(U_0 + U_1) + t\rangle$, $|F_1\rangle := |u_2\rangle$, $|M_0\rangle := |(U_0 + U_1) - t\rangle$, $|M_1\rangle := |U_0 - U_1\rangle$. 
The construction, cont’d

Define the unnormalized vector

$$|\phi_A\rangle := \sqrt{\epsilon} |U_0 + U_1\rangle,$$

and

$$|\phi_D\rangle := |(U_0 - U_1) + U_2\rangle + \epsilon |T\rangle.$$

Finally, define the unnormalized

$$|\phi\rangle := |\phi_A\rangle \otimes |A\rangle + |\phi_D\rangle \otimes |D\rangle.$$
Before we present the details for the construction and the proof, we sketch below the underlying intuitions. A fundamental feature of quantum mechanics is that measuring a quantum system may in general change the state of the system. Furthermore, two different quantum measurements may be non-commuting, a consequence of which is that the state change incurred by one measurement would alter the outcome statistics of subsequently applying the other. This is the critical property that underlies our constructions, as well as the Heisenberg Uncertainty Principle, and other well-known quantum paradoxes such as the EPR Paradox [?1] and the GHZ Paradox [?2].

In our construction, the Gender and the Treatment measurements are non-commuting (though both commute with the Result measurement). Thus adding the Gender measurement may change dramatically the outcome statistics of the subsequent measurements. More specifically, the extremal violation of the convexity property $d_t \to 2$ is a consequence of the following features (illustrated in Fig. 2): (1) The Treatment portion of the quantum state has a small amplitude in the Dead subspace, but a vanishing amplitude in the Alive subspace. This implies $R_t = 0$. (2) After measuring Gender, both eigenstates retain a tiny amplitude in the Dead subspace, but a much larger, even though still small, amplitude in the Alive subspace. This gives $R_f t \to 1$, and with (1) $d_t \to 2$. Q1 and Q2 are constructed so that $d_t \to 2$ and at the same time $d_c$ is made small. We note that unlike in the violation of spatial Bell inequalities, entanglement is perhaps not relevant in our setting.

That $|S| < 2$ follows from the necessary tradeoff between $d_t$ and $d_c$. The tradeoff is most intuitive when considering why extremal violation of the convexity property cannot take place simultaneously on $d_t$ and $d_c$ (in opposite directions). An extremal violation on $d_t$ require the above two features (though “vanishing” can be replaced by “very small”). Consequently, $\Pr(\text{AT})$ and $\Pr(\text{DT})$ are small. Similarly, that $d_c \to -2$ implies $\Pr(\text{AU})$ and $\Pr(\text{DU})$ are also small, contradicting that the sum of those four probabilities is 1.

### Two geometric constraints

- Define lengths $\ell_{\text{FTD}} := |R^E T^G F|\phi\rangle|$, $\ell_{\text{TD}} := |R^E T|\phi\rangle|$, etc.
- Triangle inequalities, e.g. $\ell_{\text{TD}} \leq \ell_{\text{FTD}} + \ell_{\text{MTD}}$
- Conservation of total probabilities
  - Sum of colored dots = sum of colored edges
- Cause tradeoff between $d_t$ and $d_c$.
Intuition for $|S| < 2$

Two geometric constraints

- Define lengths
  \[ \ell_{\text{FTD}} := \| R^D E^T G^F |\phi\rangle \|, \]
  \[ \ell_{\text{TD}} := \| R^D E^T |\phi\rangle \|, \text{ etc.} \]
- Triangle inequalities, e.g.
  \[ \ell_{\text{TD}} \leq \ell_{\text{FTD}} + \ell_{\text{MTD}} \]
- Conservation of total probabilities
  Sum of colored dots = sum of colored edges.
- Cause tradeoff between $d_t$ and $d_c$. 

Before we present the details for the construction and the proof, we sketch below the underlying intuitions. A fundamental feature of quantum mechanics is that measuring a quantum system may in general change the state of the system. Furthermore, two different quantum measurements may be non-commuting, a consequence of which is that the state change incurred by one measurement would alter the outcome statistics of subsequently applying the other. This is the critical property that underlies our constructions, as well as the Heisenberg Uncertainty Principle, and other well-known quantum paradoxes such as the EPR Paradox \[?\] and the GHZ Paradox \[?\].

In our construction, the Gender and the Treatment measurements are non-commuting (though both commute with the Result measurement). Thus adding the Gender measurement may change dramatically the outcome statistics of the subsequent measurements. More specifically, the extremal violation of the convexity property \[d_t \rightarrow 2\] is a consequence of the following features (illustrated in Fig. 2):

1. The Treatment portion of the quantum state has a small amplitude in the Dead subspace, but a vanishing amplitude in the Alive subspace. This implies \[R_t = 0\].
2. After measuring Gender, both eigenstates retain a tiny amplitude in the Dead subspace, but a much larger, even though still small, amplitude in the Alive subspace. This gives \[R_{f,t} = R_{m,t} \rightarrow 1\], and with (1) \[d_t \rightarrow 2\].

Q1 and Q2 are constructed so that \[d_t \rightarrow 2\] and at the same time \[d_c\] is made small. We note that unlike in the violation of spatial Bell inequalities, entanglement is perhaps not relevant in our setting.

That \[|S| < 2\] follows from the necessary tradeoff between \[d_t\] and \[d_c\]. The tradeoff is most intuitive when considering why extremal violation of the convexity property cannot take place simultaneously on \[d_t\] and \[d_c\](in opposite directions). An extremal violation on \[d_t\] require the above two features (though “vanishing” can be replaced by “very small”). Consequently, \(\Pr(\text{AT})\) and \(\Pr(\text{DT})\) are small. Similarly, that \[d_c \rightarrow -2\] implies \(\Pr(\text{AU})\) and \(\Pr(\text{DU})\) are also small, contradicting that the sum of those four probabilities is 1.
Before we present the details for the construction and the proof, we sketch below the underlying intuitions. A fundamental feature of quantum mechanics is that measuring a quantum system may in general change the state of the system. Furthermore, two different quantum measurements may be non-commuting, a consequence of which is that the state change incurred by one measurement would alter the outcome statistics of subsequently applying the other. This is the critical property that underlies our constructions, as well as the Heisenberg Uncertainty Principle, and other well-known quantum paradoxes such as the EPR Paradox [1] and the GHZ Paradox [2].

In our construction, the Gender and the Treatment measurements are non-commuting (though both commute with the Result measurement). Thus adding the Gender measurement may change dramatically the outcome statistics of the subsequent measurements. More specifically, the extremal violation of the convexity property $d_t \to 2$ is a consequence of the following features (illustrated in Fig. 2): (1) The Treatment portion of the quantum state has a small amplitude in the Dead subspace, but a vanishing amplitude in the Alive subspace. This implies $R_t = 0$. (2) After measuring Gender, both eigenstates retain a tiny amplitude in the Dead subspace, but a much larger, even though still small, amplitude in the Alive subspace. This gives $R_f t = R_m t \to 1$, and with (1) $d_t \to 2$. Q1 and Q2 are constructed so that $d_t \to 2$ and at the same time $d_c$ is made small. We note that unlike in the violation of spatial Bell inequalities, entanglement is perhaps not relevant in our setting.

That $|S| < 2$ follows from the necessary tradeoff between $d_t$ and $d_c$. The tradeoff is most intuitive when considering why extremal violation of the convexity property cannot take place simultaneously on $d_t$ and $d_c$ (in opposite directions). An extremal violation on $d_t$ require the above two features (though “vanishing” can be replaced by “very small”). Consequently, $\Pr(\text{AT})$ and $\Pr(\text{DT})$ are small. Similarly, that $d_c \to -2$ implies $\Pr(\text{AU})$ and $\Pr(\text{DU})$ are also small, contradicting that the sum of those four probabilities is 1.

Two geometric constraints
- Define lengths
  \[ \ell_{\text{FTD}} := \| R^D E^T G^F |\phi\rangle \|, \]
  \[ \ell_{\text{TD}} := \| R^D E^T |\phi\rangle \|, \]
  etc.
- Triangle inequalities, e.g.
  \[ \ell_{\text{TD}} \leq \ell_{\text{FTD}} + \ell_{\text{MTD}} \]
- Conservation of total probabilities
  Sum of colored dots = sum of colored edges.
- Cause tradeoff between $d_t$ and $d_c$.  

Intuition for $|S| < 2$
Intuition for $|S| < 2$

Two geometric constraints

- Define lengths
  \[ \ell_{FTD} := \| R^D E^T G^F |\phi> \|, \]
  \[ \ell_{TD} := \| R^D E^T |\phi> \|, \] etc.

- Triangle inequalities, e.g.
  \[ \ell_{TD} \leq \ell_{FTD} + \ell_{MTD} \]

- Conservation of total probabilities
  \[ \text{Sum of colored dots} = \text{sum of colored edges}. \]

- Cause tradeoff between $d_t$ and $d_c$. 

Before we present the details for the construction and the proof, we sketch below the underlying intuitions. A fundamental feature of quantum mechanics is that measuring a quantum system may in general change the state of the system. Furthermore, two different quantum measurements may be non-commuting, a consequence of which is that the state change incurred by one measurement would alter the outcome statistics of subsequently applying the other. This is the critical property that underlies our constructions, as well as the Heisenberg Uncertainty Principle, and other well-known quantum paradoxes such as the EPR Paradox and the GHZ Paradox.

In our construction, the Gender and the Treatment measurements are non-commuting (though both commute with the Result measurement). Thus adding the Gender measurement may change dramatically the outcome statistics of the subsequent measurements. More specifically, the extremal violation of the convexity property $d_t \to 2$ is a consequence of the following features (illustrated in Fig. 2): (1) The Treatment portion of the quantum state has a small amplitude in the Dead subspace, but a vanishing amplitude in the Alive subspace. This implies $R_t = 0$. (2) After measuring Gender, both eigenstates retain a tiny amplitude in the Dead subspace, but a much larger, even though still small, amplitude in the Alive subspace. This gives $R_f = R_m \to 1$, and with (1) $d_t \to 2$. Q1 and Q2 are constructed so that $d_t \to 2$ and at the same time $d_c$ is made small. We note that unlike in the violation of spatial Bell inequalities, entanglement is perhaps not relevant in our setting.

That $|S| < 2$ follows from the necessary tradeoff between $d_t$ and $d_c$. The tradeoff is most intuitive when considering why extremal violation of the convexity property cannot take place simultaneously on $d_t$ and $d_c$ (in opposite directions). An extremal violation on $d_t$ requires the above two features (though “vanishing” can be replaced by “very small”). Consequently, $\text{Pr}(AT)$ and $\text{Pr}(DT)$ are small. Similarly, that $d_c \to -2$ implies $\text{Pr}(AU)$ and $\text{Pr}(DU)$ are also small, contradicting that the sum of those four probabilities is 1.
Intuition for $|S| < 2$

**Two geometric constraints**

- **Define lengths**
  \[ \ell_{FTD} := \| R^T E^T G F | \phi \rangle \|, \]
  \[ \ell_{TD} := \| R^T E | \phi \rangle \|, \]
  etc.

- **Triangle inequalities**, e.g.
  \[ \ell_{TD} \leq \ell_{FTD} + \ell_{MTD} \]

- **Conservation of total probabilities**
  \[ \text{Sum of colored dots} = \text{sum of colored edges.} \]

- **Cause tradeoff between** $d_t$ **and** $d_c$. 
Intuition for $|S| < 2$, cont’d

Best demonstrated in proving $S < 4$

If $S \approx 4$

- Force each term to be close to 0/1,
- implying large disparities in the paired vertex/edge sizes
- Triangle inequalities force all edges to be small
- Contradicting conservation of probabilities.

More careful argument leads to $|S| < 2$. 
Intuition for $|S| < 2$, cont’d

Best demonstrated in proving $S < 4$

If $S \approx 4$

- Force each term to be close to 0/1,
- implying large disparities in the paired vertex/edge sizes
- Triangle inequalities force all edges to be small
- Contradicting conservation of probabilities.

More careful argument leads to $|S| < 2$. 

Before we present the details for the construction and the proof, we sketch below the underlying intuitions. A fundamental feature of quantum mechanics is that measuring a quantum system may in general change the state of the system. Furthermore, two different quantum measurements may be non-commuting, a consequence of which is that the state change incurred by one measurement would alter the outcome statistics of subsequently applying the other. This is the critical property that underlies our constructions, as well as the Heisenberg Uncertainty Principle, and other well-known quantum paradoxes such as the EPR Paradox [? ] and the GHZ Paradox [? ].

In our construction, the Gender and the Treatment measurements are non-commuting (though both commute with the Result measurement). Thus adding the Gender measurement may change dramatically the outcome statistics of the subsequent measurements. More specifically, the extremal violation of the convexity property $d_t \to 2$ is a consequence of the following features (illustrated in Fig. 2): (1) The Treatment portion of the quantum state has a small amplitude in the Dead subspace, but a vanishing amplitude in the Alive subspace. This implies $R_t = 0$. (2) After measuring Gender, both eigenstates retain a tiny amplitude in the Dead subspace, but a much larger, even though still small, amplitude in the Alive subspace. This gives $R_f_t \to 1$ and $R_m_t \to 1$, and with (1) $d_t \to 2$. 

Q1 and Q2 are constructed so that $d_t \to 2$ and at the same time $d_c$ is made small. We note that unlike in the violation of spatial Bell inequalities, entanglement is perhaps not relevant in our setting.

That $|S| < 2$ follows from the necessary tradeoff between $d_t$ and $d_c$. The tradeoff is most intuitive when considering why extremal violation of the convexity property cannot take place simultaneously on $d_t$ and $d_c$ (in opposite directions). An extremal violation on $d_t$ requires the above two features (though “vanishing” can be replaced by “very small”). Consequently, $Pr(\text{AT})$ and $Pr(\text{DT})$ are small. Similarly, that $d_c \to -2$ implies $Pr(\text{AU})$ and $Pr(\text{DU})$ are also small, contradicting that the sum of those four probabilities is 1.
Intuition for $|S| < 2$, cont’d

Best demonstrated in proving $S < 4$

If $S \approx 4$

- Force each term to be close to 0/1,
- implying large disparities in the paired vertex/edge sizes
- Triangle inequalities force all edges to be small
- Contradicting conservation of probabilities.

More careful argument leads to $|S| < 2$. 
Intuition for $|S| < 2$, cont’d

Best demonstrated in proving $S < 4$

If $S \approx 4$

- Force each term to be close to $0/1$,
- implying large disparities in the paired vertex/edge sizes
- Triangle inequalities force all edges to be small
- Contradicting conservation of probabilities.

More careful argument leads to $|S| < 2$. 
Intuition for $|S| < 2$, cont’d

Best demonstrated in proving $S < 4$

If $S \approx 4$

- Force each term to be close to $0/1$,
- implying large disparities in the paired vertex/edge sizes
- Triangle inequalities force all edges to be small
- Contradicting conservation of probabilities.

More careful argument leads to $|S| < 2$. 
Intuition for $|S| < 2$, cont’d

Best demonstrated in proving $S < 4$

If $S \approx 4$

- Force each term to be close to 0/1,
- implying large disparities in the paired vertex/edge sizes
- Triangle inequalities force all edges to be small
- Contradicting conservation of probabilities.

More careful argument leads to $|S| < 2$. 
Intuition for $|S| < 2$, cont’d

Best demonstrated in proving $S < 4$

If $S \approx 4$

- Force each term to be close to $0/1$,
- implying large disparities in the paired vertex/edge sizes
- Triangle inequalities force all edges to be small
- Contradicting conservation of probabilities.

More careful argument leads to $|S| < 2$. 
Conclusion

- Introduced a model to discuss Simpson’s Paradox in the quantum setting
- Found non-classical reversal (violation of a Bell-type inequality)
- Proved limit on the non-classical reversal (Tsirelson-type inequality)
- Constructed instances approaching the limit
- Our inequalities are linear in conditional probabilities involving repeated measurements
  - Different from temporal inequalities, also involving repeated measurements (Leggett and Garg 1985, Brukner et al. 2004)
- Our inequalities are linear in conditional probabilities but degree-6 in entries of the density operator
  - Quadratic inequalities studied before (Uffink 2002)
Conclusion

• Introduced a model to discuss Simpson’s Paradox in the quantum setting
• Found non-classical reversal (violation of a Bell-type inequality)
• Proved limit on the non-classical reversal (Tsirelson-type inequality)
• Constructed instances approaching the limit
• Our inequalities are linear in conditional probabilities involving repeated measurements
  • Different from temporal inequalities, also involving repeated measurements (Leggett and Garg 1985, Brukner et al. 2004)
• Our inequalities are linear in conditional probabilities but degree-6 in entries of the density operator.
  • Quadratic inequalities studied before (Uffink 2002).
Conclusion

- Introduced a model to discuss Simpson’s Paradox in the quantum setting
- Found non-classical reversal (violation of a Bell-type inequality)
- Proved limit on the non-classical reversal (Tsirelson-type inequality)
- Constructed instances approaching the limit
- Our inequalities are linear in conditional probabilities involving repeated measurements
  - Different from temporal inequalities, also involving repeated measurements (Leggett and Garg 1985, Brukner et al. 2004)
- Our inequalities are linear in conditional probabilities but degree-6 in entries of the density operator.
  - Quadratic inequalities studied before (Uffink 2002)
Conclusion

- Introduced a model to discuss Simpson’s Paradox in the quantum setting
- Found non-classical reversal (violation of a Bell-type inequality)
- Proved limit on the non-classical reversal (Tsirelson-type inequality)
- Constructed instances approaching the limit
- Our inequalities are linear in conditional probabilities involving repeated measurements
  - Different from temporal inequalities, also involving repeated measurements (Legget and Garg 1985, Brukner et al. 2004)
- Our inequalities are linear in conditional probabilities but degree-6 in entries of the density operator
  - Quadratic inequalities studied before (Uffink 2002)
Conclusion

- Introduced a model to discuss Simpson’s Paradox in the quantum setting
- Found non-classical reversal (violation of a Bell-type inequality)
- Proved limit on the non-classical reversal (Tsirelson-type inequality)
- Constructed instances approaching the limit
- Our inequalities are linear in conditional probabilities involving repeated measurements
  - Different from temporal inequalities, also involving repeated measurements (Legget and Garg 1985, Brukner et al. 2004).
  - Our inequalities are linear in conditional probabilities but degree-6 in entries of the density operator.
  - Quadratic inequalities studied before (Uffink 2002).
Conclusion

• Introduced a model to discuss Simpson’s Paradox in the quantum setting
• Found non-classical reversal (violation of a Bell-type inequality)
• Proved limit on the non-classical reversal (Tsirelson-type inequality)
• Constructed instances approaching the limit
• Our inequalities are linear in conditional probabilities involving repeated measurements
  • Different from temporal inequalities, also involving repeated measurements (Legget and Garg 1985, Brukner et al. 2004).
• Our inequalities are linear in conditional probabilities but degree-6 in entries of the density operator.
  • Quadratic inequalities studied before (Uffink 2002).
Conclusion

- Introduced a model to discuss Simpson’s Paradox in the quantum setting
- Found non-classical reversal (violation of a Bell-type inequality)
- Proved limit on the non-classical reversal (Tsirelson-type inequality)
- Constructed instances approaching the limit
- Our inequalities are linear in conditional probabilities involving repeated measurements
  - Different from temporal inequalities, also involving repeated measurements (Legget and Garg 1985, Brukner et al. 2004).
- Our inequalities are linear in conditional probabilities but degree-6 in entries of the density operator.
  - Quadratic inequalities studied before (Uffink 2002).
Conclusion

- Introduced a model to discuss Simpson’s Paradox in the quantum setting
- Found non-classical reversal (violation of a Bell-type inequality)
- Proved limit on the non-classical reversal (Tsirelson-type inequality)
- Constructed instances approaching the limit
- Our inequalities are linear in conditional probabilities involving repeated measurements
  - Different from temporal inequalities, also involving repeated measurements (Legget and Garg 1985, Brukner et al. 2004).
- Our inequalities are linear in conditional probabilities but degree-6 in entries of the density operator.
  - Quadratic inequalities studied before (Uffink 2002).
So what?

Remember my tagline

*A “quantum drug” can be life-saving for both men and women, but deadly for all.*

Quantum decision-making? Quantum correlation v.s. causality?
So what?

Remember my tagline

A “quantum drug” can be life-saving for both men and women, but deadly for all.

Quantum decision-making? Quantum correlation v.s. causality?
Critics said

- “It is not surprising.” The same old thing of non-commutativity of operators.
- “It is not telling me anything new about physics.”
- “The analogy using quantum drug is misleading.”
- ....
Critics said

- “It is not surprising.” The same old thing of non-commutativity of operators.
- “It is not telling me anything new about physics.”
- “The analogy using quantum drug is misleading.”
- ....
Critics said

- “It is not surprising.” The same old thing of non-commutativity of operators.
- “It is not telling me anything new about physics.”
- “The analogy using quantum drug is misleading.”
- ....
Critics said

- “It is not surprising.” The same old thing of non-commutativity of operators.
- “It is not telling me anything new about physics.”
- “The analogy using quantum drug is misleading.”
- ....
My response to critics

caveat emptor
My response to critics

caveat emptor
Connection with quantum cognition?

**Disjunctive Effect** (Tversky and Shafir 1992) in human decision-making can be seen as a violation of the convexity property

- Option A is preferred over Option B when either outcome of an unknown W is revealed.
- But when W is not revealed, B is preferred over A.

**Quantum Cognition** (Bruza et al. 2009) an emerging field to use quantum probabilities to describe human decision-making.
Disjunctive Effect (Tversky and Shafir 1992) in human decision-making can be seen as a violation of the convexity property

- Option A is preferred over Option B when either outcome of an unknown $W$ is revealed.
- But when $W$ is not revealed, $B$ is preferred over $A$.

Quantum Cognition (Bruza et al. 2009) an emerging field to use quantum probabilities to describe human decision-making.
Connection with quantum cognition?

Disjunctive Effect (Tversky and Shafir 1992) in human decision-making can be seen as a violation of the convexity property

- Option A is preferred over Option B when either outcome of an unknown $W$ is revealed.
- But when $W$ is not revealed, $B$ is preferred over $A$.

Quantum Cognition (Bruza et al. 2009) an emerging field to use quantum probabilities to describe human decision-making.
Connection with quantum cognition?

**Disjunctive Effect** (Tversky and Shafir 1992) in human decision-making can be seen as a violation of the convexity property

- Option $A$ is preferred over Option $B$ when *either* outcome of an unknown $W$ is revealed.
- But when $W$ is *not revealed*, $B$ is preferred over $A$.

**Quantum Cognition** (Bruza et al. 2009) an emerging field to use quantum probabilities to describe human decision-making.
Connection with quantum cognition?

Disjunctive Effect (Tversky and Shafir 1992) in human decision-making can be seen as a violation of the convexity property

- Option A is preferred over Option B when either outcome of an unknown W is revealed.
- But when W is not revealed, B is preferred over A.

Quantum Cognition (Bruza et al. 2009) an emerging field to use quantum probabilities to describe human decision-making.