Title: Network nonlocality and large linear programs

Speakers: Victor Gitton

Series: Quantum Foundations

Date: November 23, 2023 - 11:00 AM

URL: https://pirsa.org/23110074

Abstract: Network nonlocality, and more specifically, triangle network nonlocality, is a basic feature of modern causal modelling when going beyond Bell scenarios. However, despite the apparent simplicity of the problems one may formulate, relatively little is known due to the hardness of certifying nonlocality in networks. In this talk, I will describe a motivating example of a quantum triangle distribution, the Elegant Joint Measurement due to Nicolas Gisin, that is strongly believed to be nonlocal even in the presence of experimental noise. I will then present the ongoing effort to produce a computer-assisted proof of nonlocality for this distribution, thereby developing a toolkit to tackle general nonlocality problems. This effort is based on the inflation technique for causal inference, but taken to higher levels than what was generally considered tractable. This is made possible by a number of optimization techniques, involving symmetry reductions, branch-and-bound optimization, and most importantly, the use of a Frank-Wolfe algorithm to bypass the need to call a standard linear program solver.

Zoom link https://pitp.zoom.us/j/97499052021?pwd=R1EyU2pmc1hFSzJ1UEpJQ1h0RnQzdz09

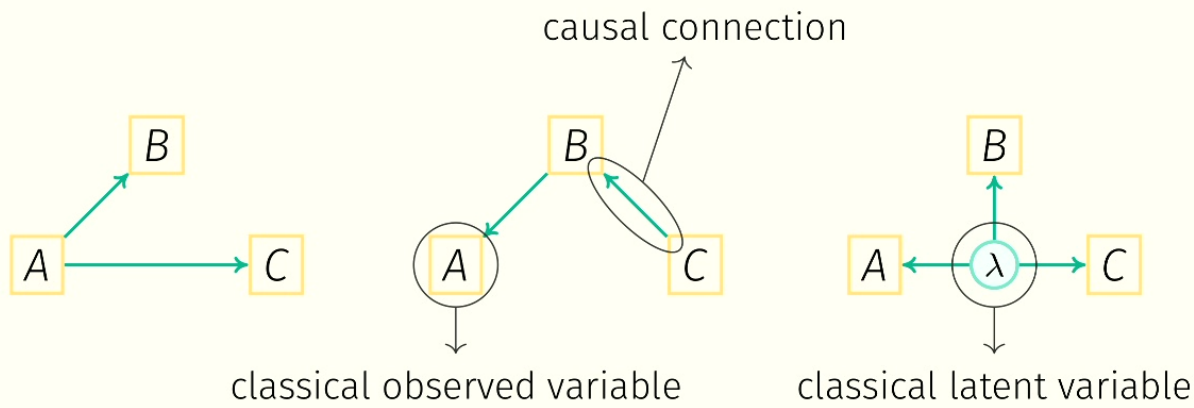
Network nonlocality and large linear programs

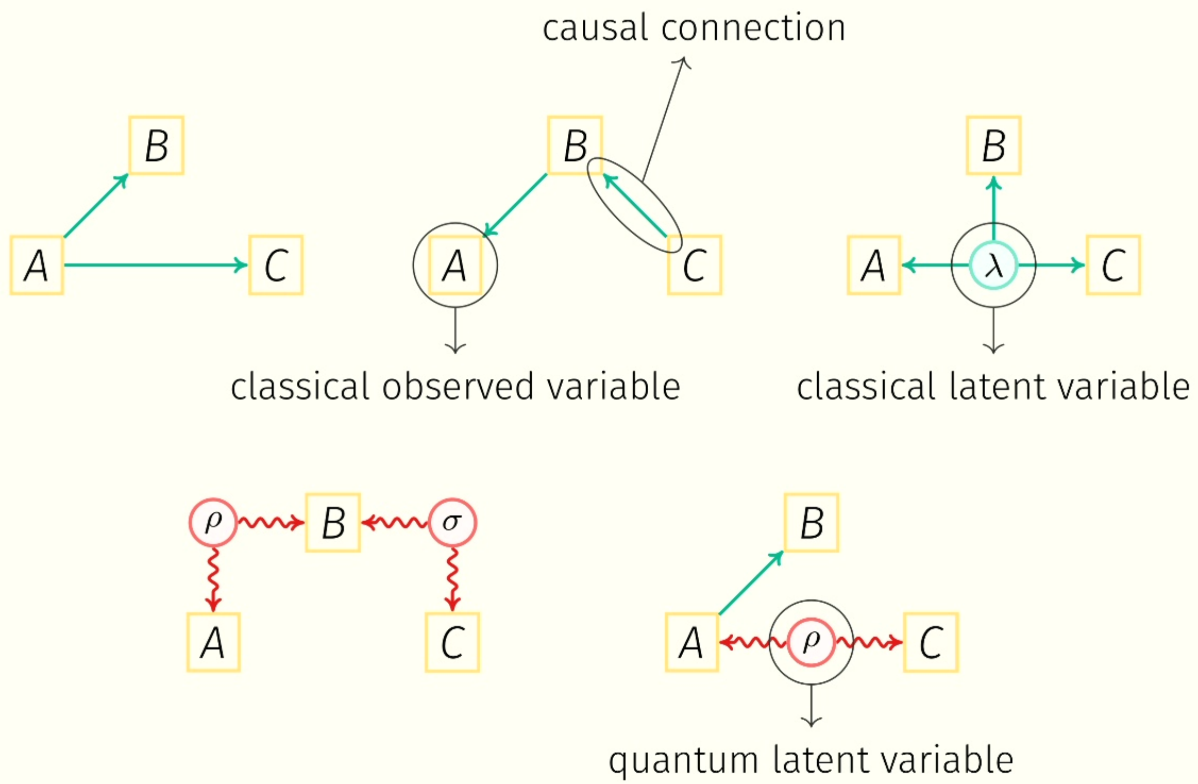
Victor Gitton

November 23rd, 2023

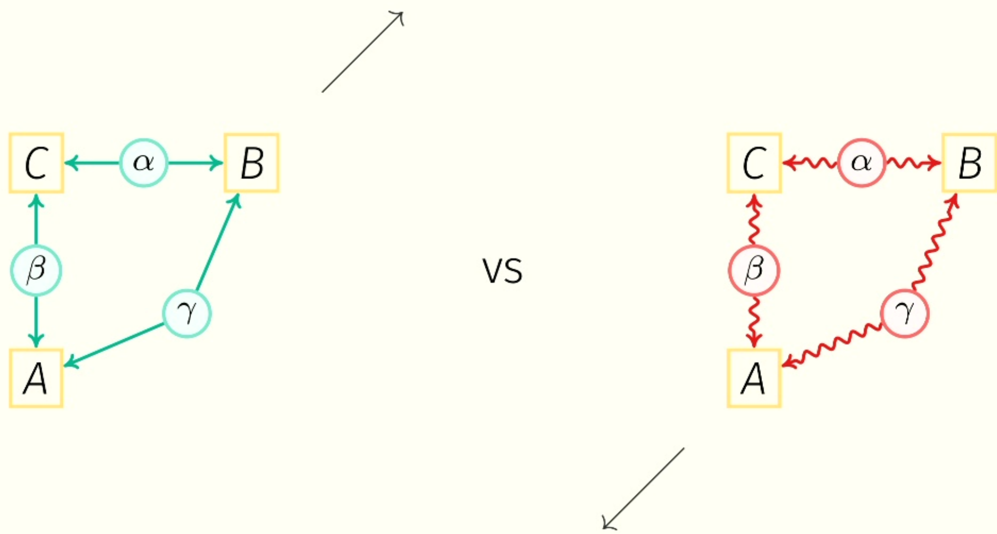
Institute for Theoretical Physics, ETH Zürich

Part I — Nonlocality


causal connection
B
A
C
B
A
C
B
A
λ
C
classical observed variable
classical latent variable


causal connection
B
A
C
B
A
C
classical observed variable
B
A
λ
C
classical latent variable
ρ
B
σ
A
C
B
A
ρ
C
quantum latent variable

$$\mathcal{L} = \left\{ p(abc) = \int_{[0,1]^{\times 3}} \mathrm{d}\alpha\, \mathrm{d}\beta\, \mathrm{d}\gamma\, p_A(a|\beta\gamma) p_B(b|\gamma\alpha) p_C(c|\alpha\beta) \right\}$$



$$\mathcal{Q} = \left\{ p(abc) = \mathrm{Tr}\left[\left(E^a_{\beta_L \gamma_R} \otimes F^b_{\gamma_L \alpha_R} \otimes G^c_{\alpha_L \beta_R} \right) \left(\rho_{\alpha_L \alpha_R} \otimes \sigma_{\beta_L \beta_R} \otimes \tau_{\gamma_L \gamma_R} \right) \right] \right\}$$

Experimental nonclassicality in a causal network without assuming freedom of choice

Emanuele Polino, Davide Poderini, Giovanni Rodari, Iris Agresti, Alessia Suprano, Gonzalo Carvacho, Elie Wolfe, Askery Canabarro, George Moreno, Giorgio Milani, Robert W. Spekkens, Rafael Chaves & Fabio Sciarrino

Nature Communications **14**, Article number: 909 (2023) | Cite this article

Genuine Quantum Nonlocality in the Triangle Network

Marc-Olivier Renou, Elisa Bäumer, Sadra Boreiri, Nicolas Brunner, Nicolas Gisin, and Salman Beigi
Phys. Rev. Lett. **123**, 140401 – Published 30 September 2019

Can we get noise-robust, "genuine" triangle nonlocality?

The elegant joint measurement

$p_{\mathrm{ejm}}(abc)$ such that, for $k \neq l \neq m \neq k \in \{0, 1, 2, 3\}$,

$$p_{\mathrm{ejm}}(kkk) = \frac{25}{256},$$
$$p_{\mathrm{ejm}}(kkl) = p_{\mathrm{ejm}}(klk) = p_{\mathrm{ejm}}(lkk) = \frac{1}{256},$$
$$p_{\mathrm{ejm}}(klm) = \frac{5}{256}$$

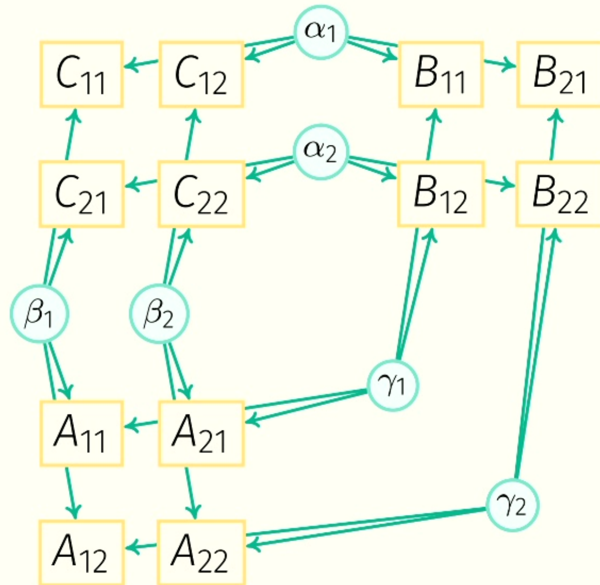
$p_{\mathrm{ejm}} \in \mathcal{Q}$ and strong indications that $p_{\mathrm{ejm}} \notin \mathcal{L}$, even with noise

Inflation

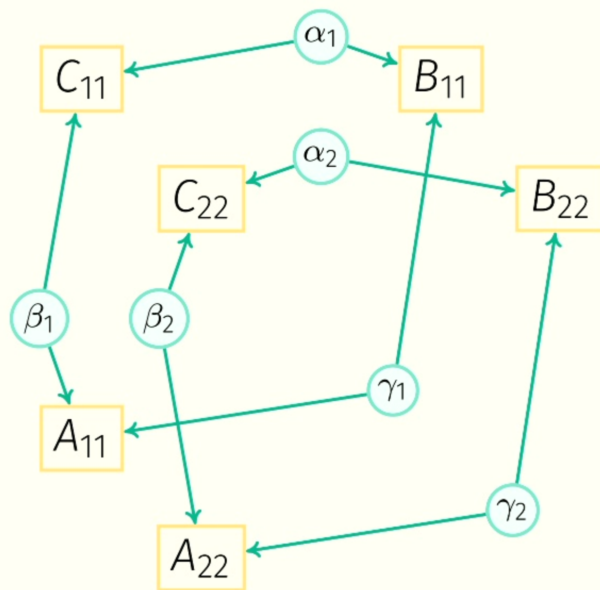
- Inflation gives outer approximations of network correlations
- For classical and non-signaling networks, get a linear program relaxation
- Also exists for quantum networks, get a semidefinite program relaxation
- Best technique for proofs of nonlocality in networks!

Protocol: pick an inflation graph, then formulate a relaxation

2×2×2 inflation graph:



2×2×2 inflation graph:



$$p(abc) \in \mathcal{L} \implies \exists p_A, p_B, p_C \implies \exists Q_{\mathrm{h}} \text{ (honest)} \implies \exists Q \text{ (relaxed)}$$

$$Q\left(\begin{array}{cc|cc} c_{11} & & b_{11} & \\ & c_{22} & & b_{22} \\ \hline a_{11} & & & \\ & a_{22} & & \end{array}\right) = p(a_{11}b_{11}c_{11})\,p(a_{22}b_{22}c_{22}),$$

$$Q\left(\begin{array}{cc|cc} c_{11} & c_{12} & b_{11} & b_{21} \\ c_{21} & c_{22} & b_{12} & b_{22} \\ \hline a_{11} & a_{21} & & \\ a_{12} & a_{22} & & \end{array}\right) = Q\left(\begin{array}{cc|cc} c_{21} & c_{22} & b_{12} & b_{22} \\ c_{11} & c_{12} & b_{11} & b_{21} \\ \hline a_{11} & a_{21} & & \\ a_{12} & a_{22} & & \end{array}\right) \text{ etc}$$

$$\exists \text{ nonnegative, normalized } Q \text{ s.t. } \begin{cases} M(Q) = p^{\otimes 2}, \\ \forall g \in G: \; g(Q) = Q. \end{cases}$$

$$\iff \exists \text{ nonnegative, normalized } Q \text{ s.t. } M \circ \mathcal{G}(Q) = p^{\otimes 2}$$

$$\text{where } \mathcal{G}(\cdot) = \frac{1}{|G|} \sum_{g \in G} g(\cdot)$$

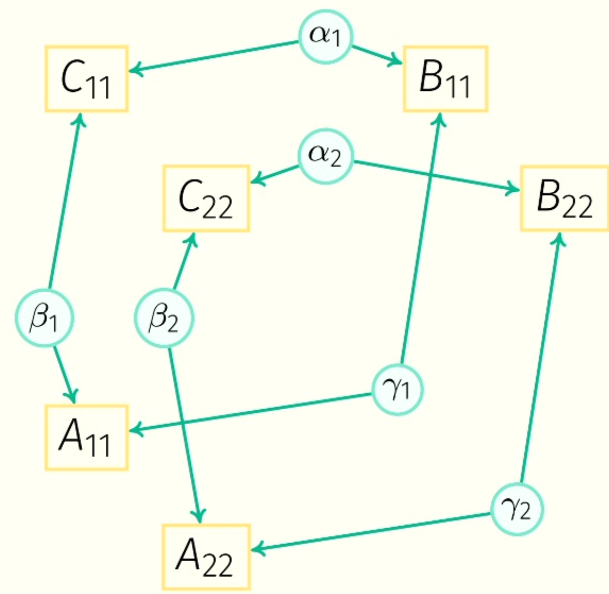
$$\iff p^{\otimes 2} \in \mathrm{conv}\left\{ M \circ \mathcal{G}\left(Q_e^{\mathrm{det}}\right) \right\}_{e \in \mathcal{E}} \text{ where } e \simeq \begin{array}{cc|cc} c_{11} & c_{12} & b_{11} & b_{21} \\ c_{21} & c_{22} & b_{12} & b_{22} \\ \hline a_{11} & a_{21} & & \\ a_{12} & a_{22} & & \end{array}$$

$$p(abc) \in \mathcal{L} \implies \exists p_A, p_B, p_C \implies \exists Q_{\mathrm{h}} \text{ (honest)} \implies \exists Q \text{ (relaxed)}$$

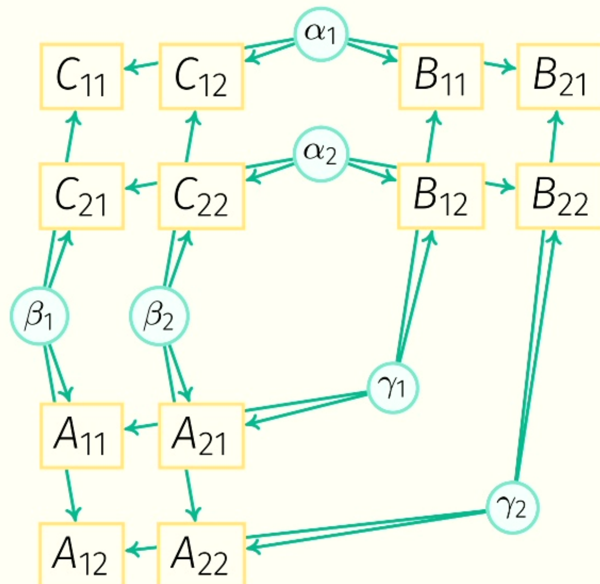
$$Q\left(\begin{array}{cc|cc} c_{11} & & b_{11} & \\ & c_{22} & & b_{22} \\ \hline a_{11} & & & \\ & a_{22} & & \end{array}\right) = p(a_{11}b_{11}c_{11})p(a_{22}b_{22}c_{22}),$$

$$Q\left(\begin{array}{cc|cc} c_{11} & c_{12} & b_{11} & b_{21} \\ c_{21} & c_{22} & b_{12} & b_{22} \\ \hline a_{11} & a_{21} & & \\ a_{12} & a_{22} & & \end{array}\right) = Q\left(\begin{array}{cc|cc} c_{21} & c_{22} & b_{12} & b_{22} \\ c_{11} & c_{12} & b_{11} & b_{21} \\ \hline a_{11} & a_{21} & & \\ a_{12} & a_{22} & & \end{array}\right) \text{ etc}$$

2×2×2 inflation graph:



2×2×2 inflation graph:



Polytope membership problem:

$$x \in \mathrm{conv}\{y_e\}_e \implies \text{inconclusive}$$
$$x \notin \mathrm{conv}\{y_e\}_e \implies \text{nonlocal, } p \notin \mathcal{L}$$

with $x, y_e \in \mathbb{R}^n$ and $|\mathcal{E}| = m$. For the previous 2×2×2 problem:

$$n = 4^6 \approx 4 \cdot 10^3, \qquad m = 4^{12} \approx 16 \cdot 10^6$$

For a 2×2×4 problem that we solved:

$$n \approx 4 \cdot 10^5, \qquad m = 4^{20} \approx 10^{12}$$

How to deal with this?

$$\exists \text{ nonnegative, normalized } Q \text{ s.t. } \begin{cases} M(Q) = p^{\otimes 2}, \\ \forall g \in G: \ g(Q) = Q. \end{cases}$$

$$\iff \exists \text{ nonnegative, normalized } Q \text{ s.t. } M \circ \mathcal{G}(Q) = p^{\otimes 2}$$

$$\text{where } \mathcal{G}(\cdot) = \frac{1}{|G|} \sum_{g \in G} g(\cdot)$$

$$\iff p^{\otimes 2} \in \text{conv}\left\{ M \circ \mathcal{G}\left(Q_e^{\text{det}}\right) \right\}_{e \in \mathcal{E}} \text{ where } e \simeq \begin{array}{cc|cc} c_{11} & c_{12} & b_{11} & b_{21} \\ c_{21} & c_{22} & b_{12} & b_{22} \\ \hline a_{11} & a_{21} & & \\ a_{12} & a_{22} & & \end{array}$$

Symmetry reduction

Column symmetry:

Find G_{col} such that $y_{g_{\text{col}}(e)} = y_e$.

Quotient $\mathcal{E} \setminus G_{\text{col}}$ to get $\tilde{m} \approx m/|G_{\text{col}}|$.

Row symmetry:

Find G_{row} such that $(y_e)_{g_{\text{row}}(i)} = (y_e)_i$ and $(x)_{g_{\text{row}}(i)} = (x)_i$.

Quotient $\{1, \ldots, n\} \setminus G_{\text{row}}$ to get $\tilde{n} \approx n/|G_{\text{row}}|$.

$$\text{EJM, } 2\times2\times2 : \quad \begin{array}{c} n \approx 4 \cdot 10^3 \\ m \approx 16 \cdot 10^6 \end{array} \quad \longrightarrow \quad \begin{array}{c} \tilde{n} = 33 \\ \tilde{m} \approx 15 \cdot 10^3 \end{array}$$

Column symmetry:

Find G_{col} such that $y_{g_{\text{col}}(e)} = y_e$.

Quotient $\mathcal{E} \setminus G_{\text{col}}$ to get $\tilde{m} \approx m/|G_{\text{col}}|$.

Row symmetry:

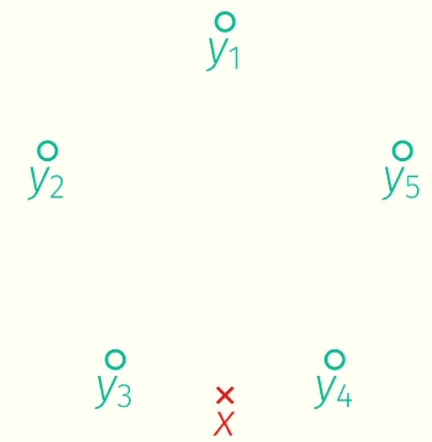
Find G_{row} such that $(y_e)_{g_{\text{row}}(i)} = (y_e)_i$ and $(x)_{g_{\text{row}}(i)} = (x)_i$.

Quotient $\{1, \ldots, n\} \setminus G_{\text{row}}$ to get $\tilde{n} \approx n/|G_{\text{row}}|$.

$$\text{EJM, } 2\times2\times2 : \quad \begin{matrix} n \approx 4 \cdot 10^3 \\ m \approx 16 \cdot 10^6 \end{matrix} \quad \longrightarrow \quad \begin{matrix} \tilde{n} = 33 \\ \tilde{m} \approx 15 \cdot 10^3 \end{matrix}$$

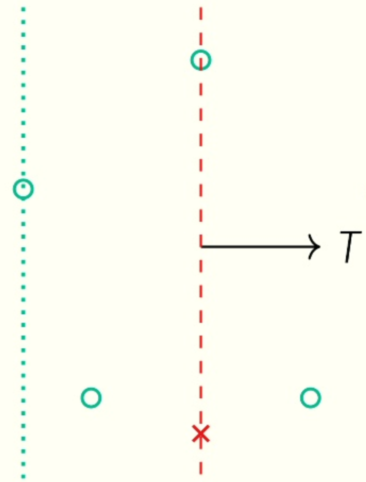
$$\text{EJM, } 2\times2\times4 : \quad \begin{matrix} n \approx 4 \cdot 10^5 \\ m \approx 10^{12} \end{matrix} \quad \longrightarrow \quad \begin{matrix} \tilde{n} \approx 4 \cdot 10^3 \\ \tilde{m} \approx 2 \cdot 10^8 \end{matrix}$$

If $x \notin \mathrm{conv}\{y_e\}_e$:



If $x \notin \mathrm{conv}\{y_e\}_e$:

Pick $T \in \mathbb{R}^n$



$$T \cdot x \geq \min_e T \cdot y_e \implies ?$$

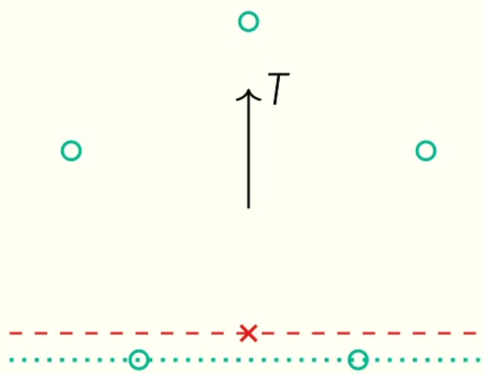
If $x \notin \mathrm{conv}\{y_e\}_e$:

Pick $T \in \mathbb{R}^n$

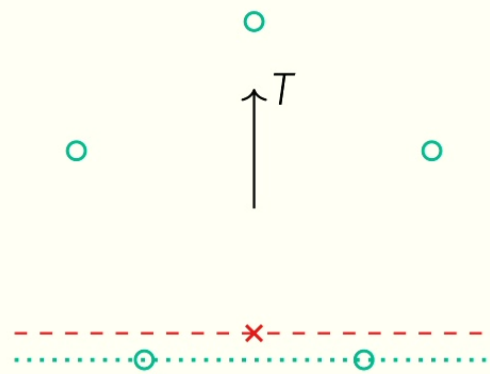
$$T \cdot x < \min_e T \cdot y_e \implies x \notin \mathrm{conv}\{y_e\}_e \implies \text{nonlocal, } p \notin \mathcal{L}$$

If $x \in \mathrm{conv}\{y_e\}_e$:

For all $T \in \mathbb{R}^n$



$$T \cdot x \geq \min_e T \cdot y_e \implies ?$$



Bruteforce algorithm: randomly sample T. What else can we do?

Frank-Wolfe

Proposed by Marguerite Frank and Philip Wolfe in 1956.

Exists in various flavours: vanilla FW (a.k.a. Gilbert's algorithm), Pairwise FW, fully corrective FW, etc

arXiv:2211.14103 [pdf, other] math.OC

Conditional Gradient Methods

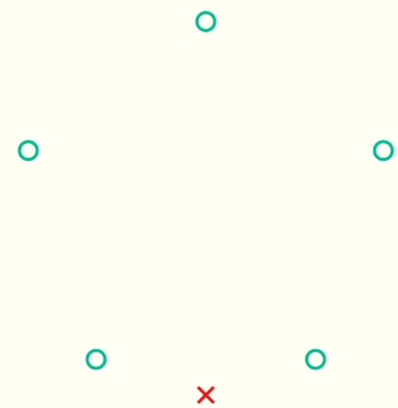
Authors: Gábor Braun, Alejandro Carderera, Cyrille W. Combettes, Hamed Hassani, Amin Karbasi, Aryan Mokhtari, Sebastian Pokutta

Improved local models and new Bell inequalities via Frank-Wolfe algorithms

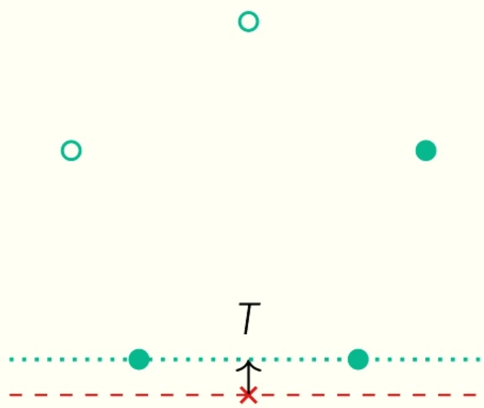
Sébastien Designolle, Gabriele Iommazzo, Mathieu Besançon, Sebastian Knebel, Patrick Gelß, and Sebastian Pokutta

Phys. Rev. Research **5**, 043059 – Published 18 October 2023

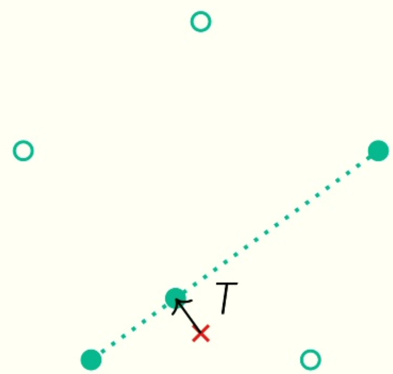
If $x \notin \mathrm{conv}\{y_e\}_e$:



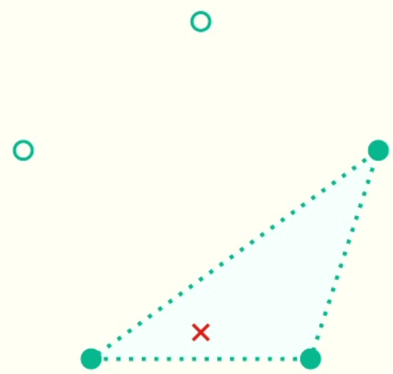
If $x \notin \text{conv}\{y_e\}_e$:



If $x \in \mathrm{conv}\{y_e\}_e$:



If $x \in \text{conv}\{y_e\}_e$:



$\implies x \in \text{conv}\{y_e\}_e \implies$ inconclusive

Linear minimization oracle:

$$\mathrm{LMO}(T) = \left(\min_e T \cdot y_e, \operatorname{argmin}_e T \cdot y_e \right)$$

Algorithm: Fully-corrective Frank-Wolfe

$\mathcal{S} \leftarrow \{y_e\}$ `// choose arbitrary initial extremal point`

repeat

  $s^* \leftarrow \operatorname{argmin}_{s \in \mathrm{conv}(\mathcal{S})} \|x - s\|_2$

  $T \leftarrow s^* - x$

  $(\alpha, e) \leftarrow \mathrm{LMO}(T)$

  if $T \cdot x < \alpha$ **then**

    return *nonlocal,* $p \notin \mathcal{L}$

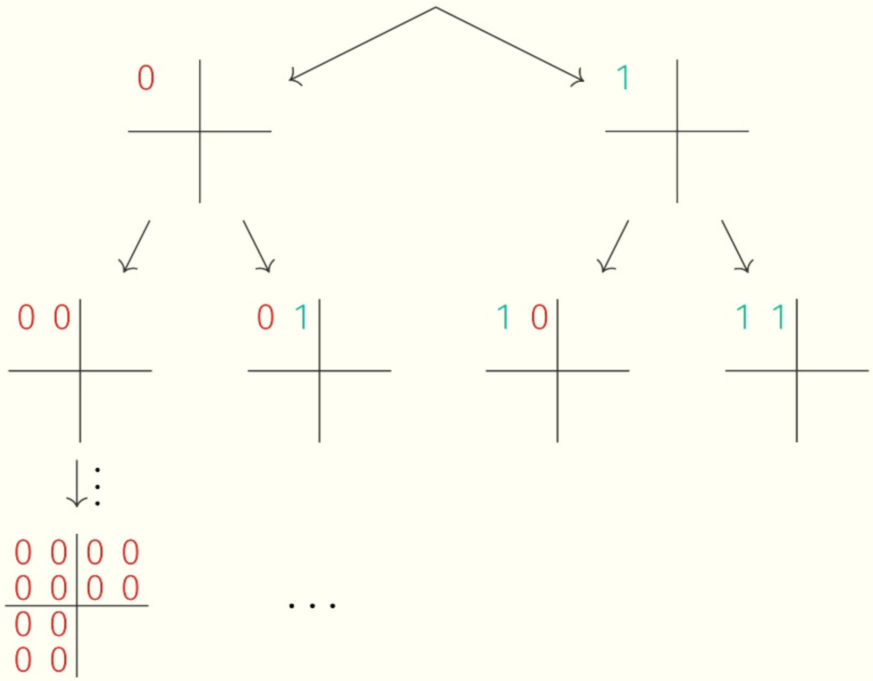
  else

    $\mathcal{S} \leftarrow \mathcal{S} \cup \{y_e\}$

until $\|T\|_2 < \varepsilon$

return *inconclusive*

Need $\min_e T \cdot y_e$. Use a tree search:

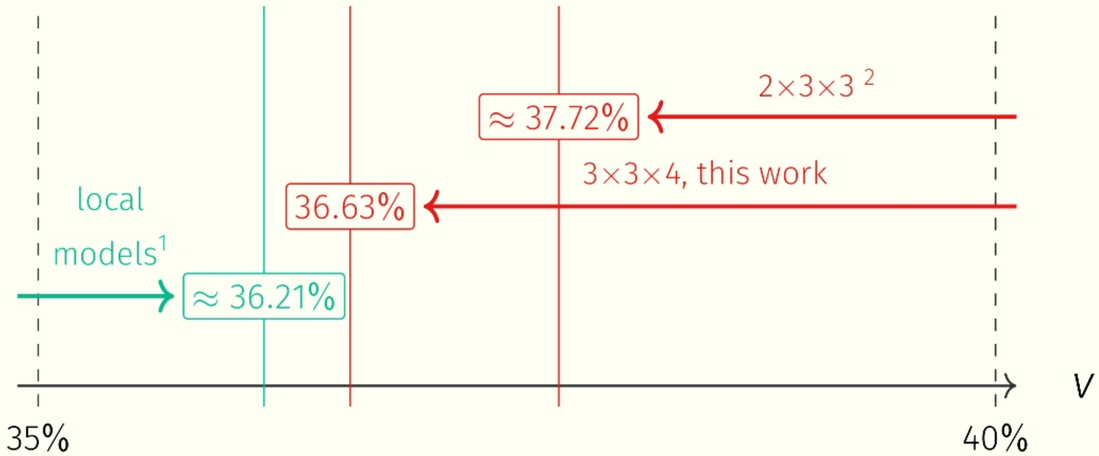


This allows some optimizations: early symmetry reduction, parallelization and branch-and-bound.

Let $p_{\text{srb}} = \frac{1}{2}([000] + [111])$ and $p^{(v)}_{\text{srb}} = v p_{\text{srb}} + (1 - v)\frac{1}{8}$.

[1] N. Gisin, JD. Bancal, Y. Cai, P. Remy, A. Tavakoli, E. Zambrini Cruzeiro, S. Popescu, N. Brunner, Nat. Commun. 11, 2378 (2020)

Let $p_{\text{srb}} = \frac{1}{2}([000] + [111])$ and $p^{(v)}_{\text{srb}} = v p_{\text{srb}} + (1 - v)\frac{1}{8}$.

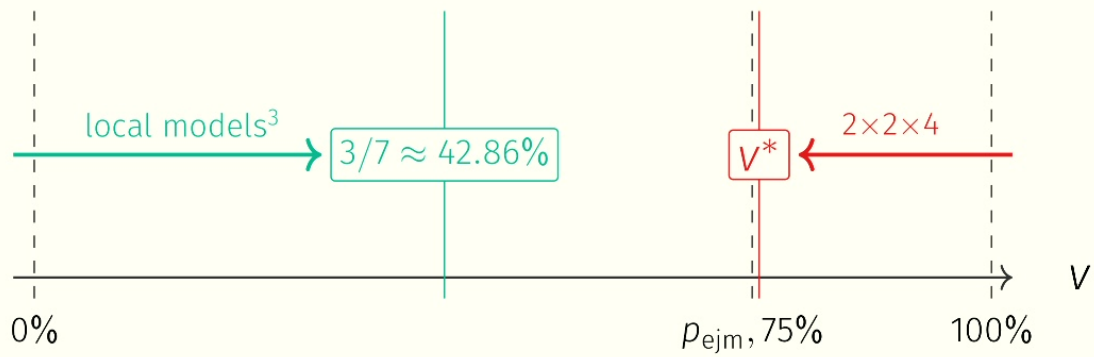


[1] N. Gisin, JD. Bancal, Y. Cai, P. Remy, A. Tavakoli, E. Zambrini Cruzeiro, S. Popescu, N. Brunner, Nat. Commun. 11, 2378 (2020)
[2] A. Pozas-Kerstjens, A. Girardin, T. Kriváchy, A. Tavakoli, N. Gisin, arXiv:2305.03745

Noise model:

$$p_{\text{ns}} = \frac{1}{8}\sum_{a}[aaa] + \frac{1}{48}\sum_{a\neq b\neq c\neq a}[abc]$$

$$p_{\text{ns}}^{(v)} = v p_{\text{ns}} + (1-v)\frac{1}{64}$$



where $v^* = \frac{775}{1024} \pm \frac{1}{1024} \approx 75.7\%$

[3] E. Bäumer, T. Kriváchy, VG, N. Gisin, R. Renner, in preparation

- Use inflation level 2×3×3 to prove the nonlocality of the EJM
- Look at nonlocality in other scenarios: bilocal network with inputs, classical triangle network with symmetric causal mechanisms, . . .
- Try to improve critical nonlocal visibility of the singlet state?
- Extend the framework to non-fanout inflation to better characterize non-signaling networks
- Generalize to other types of polynomial optimization problems?
- Apply Frank-Wolfe to quantum inflation SDPs?

In practice, end up storing $\approx n$ points, which is $\ll m$.
Only the LMO sees the m extremal points.

Algorithm: Fully-corrective Frank-Wolfe

$\mathcal{S} \leftarrow \{y_e\}$ `// choose arbitrary initial extremal point`

repeat

$s^* \leftarrow \underset{s \in \mathrm{conv}(\mathcal{S})}{\mathrm{argmin}} \|x - s\|_2$

$T \leftarrow s^* - x$

$(\alpha, e) \leftarrow \mathrm{LMO}(T)$

if $T \cdot x < \alpha$ **then**

return *nonlocal*, $p \notin \mathcal{L}$

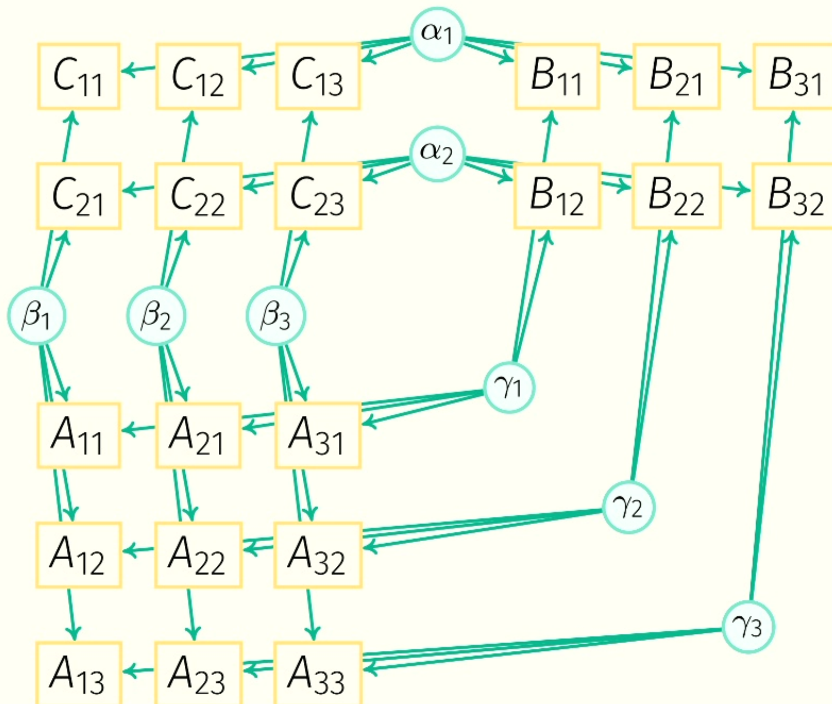
else

$\mathcal{S} \leftarrow \mathcal{S} \cup \{y_e\}$

until $\|T\|_2 < \varepsilon$

return *inconclusive*

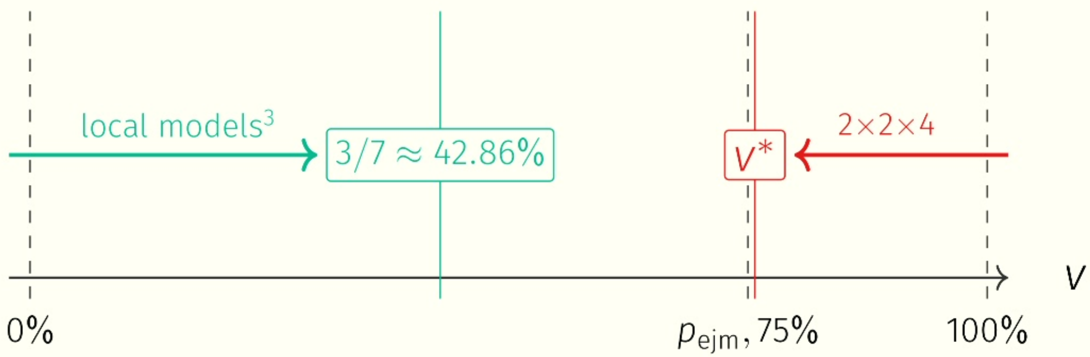
2×3×3 inflation graph:



Noise model:

$$p_{\text{ns}} = \frac{1}{8}\sum_{a}[aaa] + \frac{1}{48}\sum_{a\neq b\neq c\neq a}[abc]$$

$$p_{\text{ns}}^{(v)} = v p_{\text{ns}} + (1-v)\frac{1}{64}$$



where $v^* = \frac{775}{1024} \pm \frac{1}{1024} \approx 75.7\%$

[3] E. Bäumer, T. Kriváchy, VG, N. Gisin, R. Renner, in preparation

In practice, end up storing $\approx n$ points, which is $\ll m$.
Only the LMO sees the m extremal points.

Algorithm: Fully-corrective Frank-Wolfe

$\mathcal{S} \leftarrow \{y_e\}$ `// choose arbitrary initial extremal point`

repeat

$$s^* \leftarrow \operatorname*{argmin}_{s \in \mathrm{conv}(\mathcal{S})} \|x - s\|_2$$

$T \leftarrow s^* - x$

$(\alpha, e) \leftarrow \mathrm{LMO}(T)$

if $T \cdot x < \alpha$ **then**

return *nonlocal,* $p \notin \mathcal{L}$

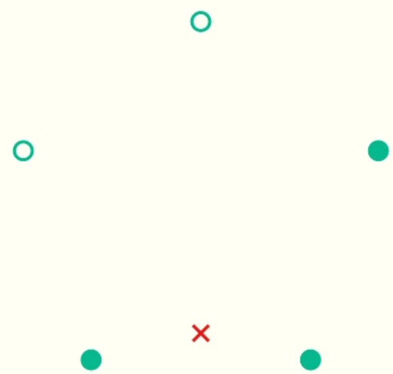
else

$\mathcal{S} \leftarrow \mathcal{S} \cup \{y_e\}$

until $\|T\|_2 < \varepsilon$

return *inconclusive*

If $x \in \mathrm{conv}\{y_e\}_e$:



Linear minimization oracle:

$$\mathrm{LMO}(T) = \left(\min_{e} T \cdot y_e, \operatorname{argmin}_{e} T \cdot y_e\right)$$

Algorithm: Fully-corrective Frank-Wolfe

$\mathcal{S} \leftarrow \{y_e\}$ `// choose arbitrary initial extremal point`

repeat

$s^* \leftarrow \operatorname{argmin}_{s \in \mathrm{conv}(\mathcal{S})} \|x - s\|_2$

$T \leftarrow s^* - x$

$(\alpha, e) \leftarrow \mathrm{LMO}(T)$

if $T \cdot x < \alpha$ **then**

return *nonlocal,* $p \notin \mathcal{L}$

else

$\mathcal{S} \leftarrow \mathcal{S} \cup \{y_e\}$

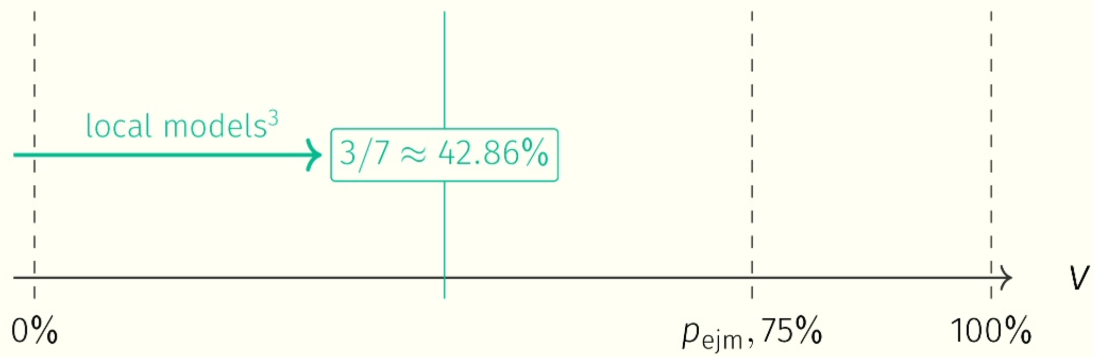
until $\|T\|_2 < \varepsilon$

return *inconclusive*

Noise model:

$$p_{\text{ns}} = \frac{1}{8}\sum_{a}[aaa] + \frac{1}{48}\sum_{a\neq b\neq c\neq a}[abc]$$

$$p_{\text{ns}}^{(v)} = v p_{\text{ns}} + (1-v)\frac{1}{64}$$



[3] E. Bäumer, T. Kriváchy, VG, N. Gisin, R. Renner, in preparation

In practice, end up storing $\approx n$ points, which is $\ll m$.
Only the LMO sees the m extremal points.

Algorithm: Fully-corrective Frank-Wolfe

$\mathcal{S} \leftarrow \{y_e\}$ `// choose arbitrary initial extremal point`

repeat

$s^* \leftarrow \underset{s \in \mathrm{conv}(\mathcal{S})}{\text{argmin}} \|x - s\|_2$

$T \leftarrow s^* - x$

$(\alpha, e) \leftarrow \mathrm{LMO}(T)$

if $T \cdot x < \alpha$ **then**

return *nonlocal,* $p \notin \mathcal{L}$

else

$\mathcal{S} \leftarrow \mathcal{S} \cup \{y_e\}$

until $\|T\|_2 < \varepsilon$

return *inconclusive*

```
    end if
  end if
  x_{t+1} = x_t − γ_t a_t                                ▷ update the weights
end for
```

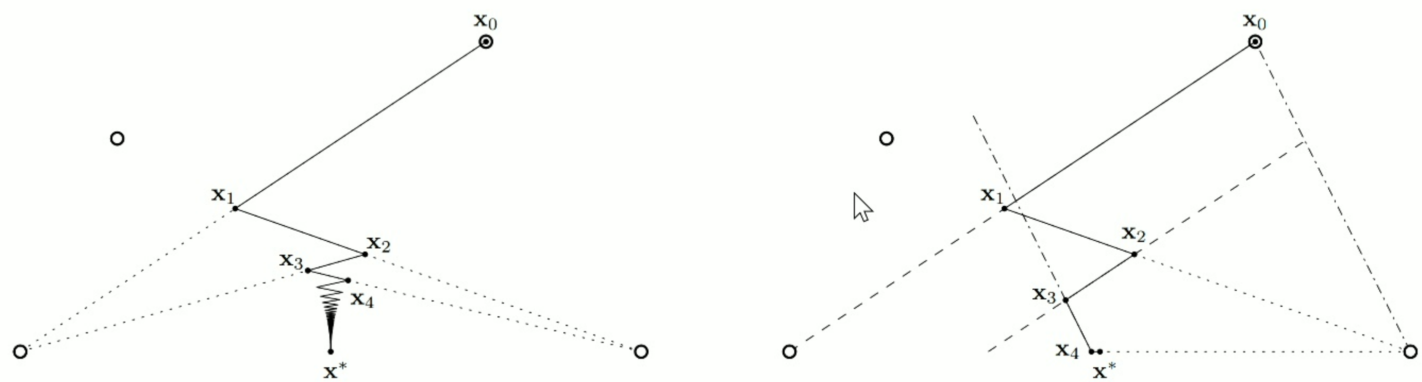


Figure 3. For a quadratic function $\frac{1}{2}\|\mathbf{x} - \mathbf{x}^*\|_2^2$, illustration of the zig-zagging effect suffered by the standard Frank-Wolfe algorithm (left), that is, Algorithm 1 in the main text, and of the countermeasure implemented by its pairwise variant (right), that is, Algorithm 2. Importantly, the pairwise steps in the latter are parallel to the lines joining the worst and best cached vertices with respect to the gradient $\mathbf{x}_t - \mathbf{x}^*$ at the current iterate $\mathbf{x}_t$. Note that the second pairwise step going from $\mathbf{x}_3$ to $\mathbf{x}_4$ (parallel to the dashed-dotted line) drops the unfavourable initial vertex.

In practice, end up storing $\approx n$ points, which is $\ll m$.
Only the LMO sees the m extremal points.

Algorithm: Fully-corrective Frank-Wolfe

$\mathcal{S} \leftarrow \{y_e\}$ `// choose arbitrary initial extremal point`

repeat

- $s^* \leftarrow \operatorname{argmin}_{s \in \operatorname{conv}(\mathcal{S})} \|x - s\|_2$
- $T \leftarrow s^* - x$
- $(\alpha, e) \leftarrow \mathrm{LMO}(T)$
- **if** $T \cdot x < \alpha$ **then**
  - **return** *nonlocal*, $p \notin \mathcal{L}$
- **else**
  - $\mathcal{S} \leftarrow \mathcal{S} \cup \{y_e\}$

until $\|T\|_2 < \varepsilon$

return *inconclusive*