Weak Coin Flipping with bias 1/10

Atul Singh Arora
Supervisor: Jérémie Roland

2017 End

Abstract

We will convert a bias 1/10 game described by Mochon in the Time Independent Point Game (TIPG) framework, which was introduced by Mochon and Kitaev to facilitate construction of protocols in a more convenient language, into an explicit protocol defined in the language of Quantum Mechaincs (QM). This breaks the long standing barrier of bias 1/6. Two new features needed to be incorporated to cross this limit. First, all protocols till the 1/6 limit could be time ordered but to cross the barrier, roughly speaking, one has to let the 'future' points influence the 'past' points. This is accomplished by means of what is called a catalyst state but it wasn't clear how to describe them using QM. Second, beyond 1/6 the unitaries that implement the basic step are defined on a large enough space so as to introduce newer unfamiliar degrees of freedom which in turn non-trivially affect the bias. This, at least partially, explains why it wasn't straightforward to cross the barrier.

Contents

1 Introduction and Prior Art

1.1 Problem statement and SDP

[state as P_B^* =max[tr($\rho\Pi_B$)] wrt constraints that one player follows the protocol honestly; similarly P_A^* and then bound max[P_A^*, P_B^*].]

1.2 Dual SDP and The TDPG Framework

[add the constraints in the maximization itself and then switch variables to get Z_A and similarly Z_B. Now plot the eigenspectrum of Z_A, Z_Bs with probability weights given by $|\psi\rangle$.]

[Introduce the idea of operator monotone functions and then convert the problem into that of points and transitions; condition for validity of a transition is given by the constraint equation; don't derive this]

1.3 The TIPG Framework

[The catalyst state thing; this I think I will have to prove because I use this]

2 TDPG $\to$ Explicit Protocol, Framework (TEF)

2.1 Motivation and Conventions

The objective of this section would be to convert a TDPG into an explicit protocol. This is far from easy and in Mochons (and in the improved version due to Aharonov et. al.) it remains the only non-constructive part of the entire analysis. Consequently we would only try to remove all trivially generalisable steps and reduce the non-trivial step into its simplest form. The non-trivial step would then be analysed separately.

Intuitively, the most natural way of constructing Zs and a $|\psi\rangle$ given an arbitrary frame (think of TDPG as a sequence of frames) is to construct an entangled state that encodes the label and define Zs to contain the coordinates corresponding to the labels. Let us make this idea more precise.

Definition (Canonical Form)**.** The set $\{|\psi\rangle, Z^A, Z^B\}$ is said to be in the Canonical Form with respect to a set of points in a frame of a TDPG if (see image #1) $|\psi\rangle = \sum_i \sqrt{P_i}\,|ii\rangle \otimes |.\rangle$, $Z^A = (\sum x_i |i\rangle\langle i| + \text{junk}) \otimes |.\rangle\langle .|$ and $Z^B = (\sum y_i |i\rangle\langle i| + \text{junk}) \otimes |.\rangle\langle .|$ where $|.\rangle$ represent extra uncoupled registers which might be present and junk refers to superfluous coordinates indexed by i as $\tilde{x}_i |i\rangle\langle i|$ for i s.t. $P_i = 0$ (similarly for $\tilde{y}_i$).

I would call a Canonical Form *clean* if no junk is present. It is easy to see that the 'label' $|ii\rangle$ will correspond to a point with coordinates x_i, y_i and weight P_i in the frame. It is tempting to imagine that we systematically construct, from each frame of a TDPG, a canonical form of $|\psi\rangle$ *s* and Zs. The unitaries can be deduced from the evolution of $|\psi\rangle$. This approach has two problems, (1) it doesn't manifestly mean that the unitaries would be decomposable into moves by Alice and Bob who communicate only through the messaging register and (2) the constraints imposed consecutive Zs, of the form $Z_{n-1} \otimes \mathbb{I} \geq U_n^\dagger (Z_n \otimes \mathbb{I}) U_n$, are not satisfied in general. This construction will ensure these issues are dissolved.

The framework will output variables in the reverse time convention indexed as, for example, $|\psi_{(i)}\rangle$, $Z_{(i)}$, $U_{(i)}$. The variables at the i^{th} step of the protocol (which follows the forward time convention) would be given by $|\psi_i\rangle = |\psi_{(N-i)}\rangle$, $Z_i = Z_{(N-i)}$ and $U_i = U^\dagger_{(N-i)}$. Note that the results so obtain extend naturally to the case where U_i may not be unitary and contains projections.

Self Praise: Basic Moves Work Out of the Box

Recall the three basic moves of a TDPG were given by

1. Raise: $p_1[x, y] \to p_1[x', y]$ s.t. $x' \geq x$.
2. Merge: $p_1[x_1, y] + p_2[x_2, y] \to p_1 + p_2 \left[\frac{p_1 x_1 + p_2 x_2}{p_1 + p_2}, y\right]$
3. Split: $(p_1 + p_2)\left[\left(\frac{p_1 w_1 + p_2 w_2}{p_1 + p_2}\right)^{-1}, y\right] \to p_1[x_1, y] + p_2[x_2, y]$ where $w_1 = 1/p_1$ and $w_2 = 1/p_2$.

We would show that our framework nearly singles out the explicit Unitaries that implement these moves which in turn (when generalised to n points) are enough to construct the former best known protocol from its TDPG. Note, however, that these moves do not exhaust the set of moves and more advanced moves will be constructed to go beyond the 1/6 limit.

2.2 The Framework

Intuition

Imagine a clean canonical description is given. Let the labels on the points one wants to transform be indexed by $\{g_i\}$ and let us also assume that one wishes to apply an x-transition (meaning Alice will perform the non-trivial step). Let the labels of the points that one wishes to leave untouched be given by $\{k_i\}$. We can write the state as

$$|\psi_{(1)}\rangle = \left(\sum_i \sqrt{p_{g_i}}\,|g_i g_i\rangle_{AB} + \sum_i \sqrt{p_{k_i}}\,|k_i k_i\rangle_{AB}\right) \otimes |m\rangle_M .$$

We want Bob to send his part of $|g_i\rangle$ states to Alice through the message register. One way is that he conditionally swaps to obtain the following

$$|\psi_{(2)}\rangle = \sum_i \sqrt{p_{g_i}}\,|g_i g_i\rangle_{AM} \otimes |m\rangle_B + \sum_i \sqrt{p_{k_i}}\,|k_i k_i\rangle_{AB} \otimes |m\rangle_M .$$

This should at most force all the points to align along the y-axis but no non-trivial constraint should arise (speaking with hindsight). Let $\{h_i\}$ be the labels of the new points after the transformation. We will assume that h and g index orthonormal vectors. Alice can update the probabilities and labels by locally performing a unitary to obtain

$$|\psi_{(3)}\rangle = \sum_i \sqrt{p_{h_i}}\,|h_i h_i\rangle_{AM} \otimes |m\rangle_B + \sum_i \sqrt{p_{k_i}}\,|k_i k_i\rangle_{AB} \otimes |m\rangle_M .$$

It is precisely this step which we analyse in special cases. Bob must now accept this by 'unswapping' to get

$$\left|\psi_{(4)}\right\rangle = \left(\sum_i \sqrt{p_{h_i}} \left|h_i h_i\right\rangle_{AB} + \sum_i \sqrt{p_{k_i}} \left|k_i k_i\right\rangle_{AB}\right) \otimes \left|m\right\rangle_M$$

which leaves Bob's Z in essentially the standard form (we will see). Remember that in the actual protocol the sequence will get reversed as described above.

Formal Description and Proofs

1. **First frame.**

$$\begin{aligned}
\left|\psi_{(1)}\right\rangle &= \left(\sum_i \sqrt{p_{g_i}} \left|g_i g_i\right\rangle_{AB} + \sum_i \sqrt{p_{k_i}} \left|k_i k_i\right\rangle_{AB}\right) \otimes \left|m\right\rangle_M \\
Z^A_{(1)} &= \sum_i x_{g_i} \left|g_i\right\rangle \left\langle g_i\right|_A + \sum_i x_{k_i} \left|k_i\right\rangle \left\langle k_i\right|_A \\
Z^B_{(1)} &= \sum_i y_{g_i} \left|g_i\right\rangle \left\langle g_i\right|_B + \sum_i y_{k_i} \left|k_i\right\rangle \left\langle k_i\right|_B .
\end{aligned}$$

Proof. Follows from the assumption of starting with a Clean Canonical Form. □

2. **Bob sends to Alice.** With $y \geq \max\{y_{g_i}\}$ the following is a valid choice

$$\begin{aligned}
\left|\psi_{(2)}\right\rangle &= \sum_i \sqrt{p_{g_i}} \left|g_i g_i\right\rangle_{AM} \otimes \left|m\right\rangle_B + \sum_i \sqrt{p_{k_i}} \left|k_i k_i\right\rangle_{AB} \otimes \left|m\right\rangle_M \\
U^{(1)} &= U^{\mathrm{SWP}\{\vec{g},m\}}_{BM} \\
Z^A_{(2)} &= Z^A_{(1)} \\
Z^B_{(2)} &= y\mathbb{I}^{\{\vec{g},m\}}_B + \sum_i y_{k_i} \left|k_i\right\rangle \left\langle k_i\right|_B .
\end{aligned}$$

Proof. We have to prove: (1) $\left|\psi_{(2)}\right\rangle = U^{(2)} \left|\psi_{(1)}\right\rangle$ and (2) $U^{(1)\dagger}\left(Z^B_{(2)} \otimes \mathbb{I}_M\right) U^{(1)} \geq \left(Z^B_{(1)} \otimes \mathbb{I}_M\right)$.
(1) It follows trivially from the defining action of $U^{(1)}$.
(2) For convenience, let momentarily $U = U^{(1)}$ and note that $U^\dagger = U$ so that we can write

$$\begin{aligned}
U\left(Z^B_{(2)} \otimes \mathbb{I}_M\right) U &= y\left(U\left(\mathbb{I}^{\{\vec{g},m\}}_B \otimes \mathbb{I}^{\{\vec{g},m\}}_M\right) U + U \underbrace{\left(\mathbb{I}^{\{\vec{g},m\}}_B \otimes \mathbb{I}^{\{\vec{k},\vec{h}\}}_M\right)}_{\text{outside } U\text{'s action space}} U\right) + U \underbrace{\left(\sum y_{k_i} \left|k_i\right\rangle \left\langle k_i\right| \otimes \mathbb{I}\right)}_{\text{outside } U\text{'s action space}} U \\
&= Z_{(2)} \otimes \mathbb{I}_M \geq Z_{(1)} \otimes \mathbb{I}_M
\end{aligned}$$

so long as $y \geq y_{g_i}$ which is guaranteed by the choice of y. □

3. **Alice's non-trivial step.** We claim that the following is a valid choice,

$$\begin{aligned}
\left|\psi_{(3)}\right\rangle &= \sum_i \sqrt{p_{h_i}} \left|h_i h_i\right\rangle_{AM} \otimes \left|m\right\rangle_B + \sum_i \sqrt{p_{k_i}} \left|k_i k_i\right\rangle_{AB} \otimes \left|m\right\rangle_M \\
E^{(2)} U^{(2)} &= E^{(2)} \left(\left|w\right\rangle \left\langle v\right| + \text{other terms acting on span}\{\left|h_i h_i\right\rangle, \left|g_i g_i\right\rangle\}\right)_{AM} \\
Z^A_{(3)} &= \sum_i x_{h_i} \left|h_i\right\rangle \left\langle h_i\right| + \sum_i x_{k_i} \left|k_i\right\rangle \left\langle k_i\right| \\
Z^B_{(3)} &= Z^B_{(2)}
\end{aligned}$$

where

$$\left|v\right\rangle = \frac{\sum_i \sqrt{p_{g_i}} \left|g_i g_i\right\rangle}{\sqrt{\sum_i p_{g_i}}},\ \left|w\right\rangle = \frac{\sum_i \sqrt{p_{h_i}} \left|h_i h_i\right\rangle}{\sqrt{\sum_i p_{h_i}}},\ E^{(2)} = \left(\sum \left|h_i\right\rangle \left\langle h_i\right|_A + \sum \left|k_i\right\rangle \left\langle k_i\right|_A\right) \otimes \mathbb{I}_M$$

subject to the condition

$$\sum x_{h_i} \left|h_i h_i\right\rangle \left\langle h_i h_i\right| \geq \sum x_{g_i} E^{(2)} U^{(2)} \left|g_i g_i\right\rangle \left\langle g_i g_i\right| U^{(2)\dagger} E^{(2)}$$

and of course the conservation of probability, viz. $\sum p_{g_i} = \sum p_{h_i}$.

Proof. We must show that (1) $E^{(2)}\left|\psi_{(3)}\right\rangle = U^{(2)}\left|\psi_{(2)}\right\rangle$ and (2) $Z^A_{(3)} \otimes \mathbb{I}_M \geq E^{(2)}U^{(2)}\left(Z^A_{(2)} \otimes \mathbb{I}_M\right)U^{(2)\dagger}E^{(2)}$
(1) Observing $E^{(2)}\left|\psi_{(3)}\right\rangle = \left|\psi_{(3)}\right\rangle$ the statement holds almost trivially by construction of $U^{(2)}$.
(2) Consider the space $\mathcal{H} = \text{span}\left\{\left|g_1 g_1\right\rangle, \left|g_2 g_2\right\rangle \ldots, \left|h_1 h_1\right\rangle, \left|h_2, h_2\right\rangle \ldots\right\}$. We will separate all expressions (they are nearly diagonal) into the $\mathcal{H}$ space (which gets non-diagonal) and the rest. We start with the RHS,

$$Z^A_{(2)} \otimes \mathbb{I}_M = \underbrace{\sum x_{g_i} \left|g_i g_i\right\rangle \left\langle g_i g_i\right|}_{\mathrm{I}} + \sum x_{g_i} \left|g_i\right\rangle \left\langle g_i\right| \otimes \left(\mathbb{I} - \left|g_i\right\rangle \left\langle g_i\right|\right) + \sum x_{k_i} \left|k_i\right\rangle \left\langle k_i\right| \otimes \mathbb{I},$$

where only term I is in the operator space spanned by $\mathcal{H}$. Note that all the terms are still diagonal. Next consider the LHS, without the Us,

$$Z^A_{(3)} \otimes \mathbb{I}_M = \underbrace{\sum x_{h_i} \left|h_i h_i\right\rangle \left\langle h_i h_i\right|}_{\mathrm{I}} + \sum x_{h_i} \left|h_i\right\rangle \left\langle h_i\right| \otimes \left(\mathbb{I} - \left|h_i\right\rangle \left\langle h_i\right|\right) + \sum x_{k_i} \left|k_i\right\rangle \left\langle k_i\right| \otimes \mathbb{I},$$

which also has only term I in the $\mathcal{H}$ operator space. Consequently, only on these will U have a non-trivial action. Let us first evaluate the non-$\mathcal{H}$ part where we only need to apply the projector. The result after separating equations where possible is

$$\begin{aligned} \sum x_{h_i} \left|h_i\right\rangle \left\langle h_i\right| \otimes \left(\mathbb{I} - \left|h_i\right\rangle \left\langle h_i\right|\right) &\geq 0 \\ \sum (x_{k_i} - x_{k_i}) \left|k_i\right\rangle \left\langle k_i\right| \otimes \mathbb{I} &\geq 0 \end{aligned}$$

which essentially only implies

$$x_{h_i} \geq 0.$$

Finally the non-trivial part yields

$$\sum x_{h_i} \left|h_i h_i\right\rangle \left\langle h_i h_i\right| \geq \sum x_{g_i} EU \left|g_i g_i\right\rangle \left\langle g_i g_i\right| U^\dagger E$$

which completes the proof. □

4. **Bob accepts Alice's change.** The following is valid.

$$\begin{aligned} \left|\psi_{(4)}\right\rangle &= \left(\sum_i \sqrt{p_{h_i}} \left|h_i h_i\right\rangle_{AB} + \sum_i \sqrt{p_{k_i}} \left|k_i k_i\right\rangle_{AB}\right) \otimes \left|m\right\rangle_M \\ E^{(3)}U^{(3)} &= E^{(3)}U^{\mathrm{SWP}\{\vec{h},m\}}_{BM} \\ Z^A_{(4)} &= Z^A_{(3)} \\ Z^B_{(4)} &= y\sum_i \left|h_i\right\rangle \left\langle h_i\right| + \sum_i y_{k_i} \left|k_i\right\rangle \left\langle k_i\right|_B \end{aligned}$$

where $E^{(3)} = \left(\sum \left|h_i\right\rangle \left\langle h_i\right| + \sum \left|k_i\right\rangle \left\langle k_i\right|\right)_B \otimes \mathbb{I}_M$.

Proof. We have to prove: (1) $E^{(3)}\left|\psi_{(4)}\right\rangle = U^{(3)}\left|\psi_{(3)}\right\rangle$ and (2) $Z^B_{(4)} \otimes \mathbb{I}_M \geq E^{(3)}U^{(3)}\left(Z^B_{(3)} \otimes \mathbb{I}_M\right)U^{(3)\dagger}E^{(3)}$.
(1) This can be proven again, by a direct application of EU (defined to be $E^{(3)}U^{(3)}$ for the proof).
(2) Note that

$$\begin{aligned} EU\left(\mathbb{I}^{\{\vec{g},m\}}_B \otimes \mathbb{I}^{\{\vec{h},\vec{g},\vec{k},m\}}_M\right)U^\dagger E &= EU\left(\mathbb{I}^{\{m\}}_B \otimes \mathbb{I}^{\{\vec{h},\vec{g},\vec{k},m\}}_M\right)U^\dagger E + E\left(\mathbb{I}^{\{\vec{g}\}}_B \otimes \mathbb{I}^{\{\vec{h},\vec{g},\vec{k},m\}}_M\right)E \\ &= EU\left(\mathbb{I}^{\{m\}}_B \otimes \mathbb{I}^{\{\vec{h},m\}}_M\right)U^\dagger E \\ &= \sum \left|h_i\right\rangle \left\langle h_i\right| \otimes \mathbb{I}^{\{m\}}_M. \end{aligned}$$

Since the other term in $Z^B_3 \otimes \mathbb{I}$ is anyway in the non-action space of U it follows that

$$EU(Z^B_3 \otimes \mathbb{I})U^\dagger E = y\sum \left|h_i\right\rangle \left\langle h_i\right| \otimes \mathbb{I}^{\{m\}}_M + \sum y_{k_i} \left|k_i\right\rangle \left\langle k_i\right| \otimes \mathbb{I}_M.$$

It only remains to show that $Z^B_{(4)} \otimes \mathbb{I}_M \geq E^{(3)}U^{(3)}\left(Z^B_{(3)} \otimes \mathbb{I}_M\right)U^{(3)\dagger}E^{(3)}$ which it obviously is because $y\sum \left|h_i\right\rangle \left\langle h_i\right| \otimes \mathbb{I}_M \geq y\sum \left|h_i\right\rangle \left\langle h_i\right| \otimes \mathbb{I}^{\{m\}}_M$ and the y_{k_i} term is common. □

2.3 Important Special Case: The Blinkered Unitary

So far we have not specified the non-trivial $U^{(2)}$ (which I will call U from now) beyond requiring it to have a certain action on the honest state. We now define an important class of U, we call the Blinkered Unitary, as

$$U = |w\rangle\langle v| + |v\rangle\langle w| + \sum |v_i\rangle\langle v_i| + \sum |w_i\rangle\langle w_i| + \mathbb{I}^{\text{outside } \mathcal{H}}$$

and can even drop the last term as we are restricting our analysis to the $\mathcal{H}$ operator space, where $|v\rangle, \{|v_i\rangle\}$ form a complete orthonormal basis and so do $|w\rangle, \{|w_i\rangle\}$ wrt span$\{|g_i g_i\rangle\}$ and span$\{|v_i v_i\rangle\}$ respectively. The blinkered unitary can be used to implement the two non-trivial operations of the set of basic moves.

- Merge: $g_1, g_2 \to h_1$
  We can construct from the very definitions

$$|v\rangle = \frac{\sqrt{p_{g_1}}|g_1g_1\rangle + \sqrt{p_{g_2}}|g_2g_2\rangle}{N}, \; |v_1\rangle = \frac{\sqrt{p_{g_2}}|g_1g_1\rangle - \sqrt{p_{g_1}}|g_2g_2\rangle}{N}, \; |w\rangle = |h_1h_1\rangle$$

  with $N = \sqrt{p_{g_1} + p_{g_2}}$ and even

$$U = |w\rangle\langle v| + |v\rangle\langle w| + |v_1\rangle\langle v_1| \, (= U^\dagger).$$

  I would need

$$EU|g_1g_1\rangle = \frac{\sqrt{p_{g_1}}|w\rangle}{N}, \; EU|g_2g_2\rangle = \frac{\sqrt{p_{g_2}}|w\rangle}{N}$$

  because the constraint was (plugging in the m and n)

$$x_h |h_1h_1\rangle\langle h_1h_1| \geq \sum x_{g_i} EU|g_ig_i\rangle\langle g_ig_i| U^\dagger E$$

  which becomes

$$x_h \geq \frac{p_{g_1}x_{g_1} + p_{g_2}x_{g_2}}{N^{22}}.$$

  This is precisely the merge condition Mochon derives. This can be readily generalised to an $m \to 1$ point merge condition by simply constructing appropriate vectors (which we leave for the appendix).

- Split: $g_1 \to h_1, h_2$

$$|v\rangle = |g_1g_1\rangle, \; |w\rangle = \frac{\sqrt{p_{h_1}}|h_1h_1\rangle + \sqrt{p_{h_2}}|h_2h_2\rangle}{N}, \; |w_1\rangle = \frac{\sqrt{p_{h_2}}|h_1h_1\rangle - \sqrt{p_{h_1}}|h_2h_2\rangle}{N}$$

  with $N = \sqrt{p_{h_1} + p_{h_2}}$ and

$$U = |v\rangle\langle w| + |w\rangle\langle v| + |w_1\rangle\langle w_1| = U^\dagger.$$

  We evaluate $EU|g_1g_1\rangle = |w\rangle$ which upon being plugged into the constraint yields

$$x_{h_1}|h_1h_1\rangle\langle h_1h_1| + x_{h_2}|h_2h_2\rangle\langle h_2h_2| - x_{g_1}|w\rangle\langle w| \geq 0.$$

  This yields the matrix equation

$$\begin{bmatrix} x_{h_1} & \\ & x_{h_2} \end{bmatrix} - \frac{x_{g_1}}{N^2}\begin{bmatrix} p_{h_1} & \sqrt{p_{h_1}p_{h_2}} \\ \sqrt{p_{h_1}p_{h_2}} & p_{h_2} \end{bmatrix} \geq 0$$

$$\mathbb{I} \geq \frac{x_{g_1}}{N^2}\begin{bmatrix} \frac{p_{h_1}}{x_{h_1}} & \sqrt{\frac{p_{h_1}}{x_{h_1}}\frac{p_{h_2}}{x_{h_2}}} \\ \sqrt{\frac{p_{h_1}}{x_{h_1}}\frac{p_{h_2}}{x_{h_2}}} & \frac{p_{h_2}}{x_{h_2}} \end{bmatrix} \quad \left(\begin{array}{l}\text{using the } F - M \geq 0 \\ \implies \mathbb{I} - \sqrt{F}^{-1}M\sqrt{F}^{-1} \geq 0\end{array}\right)$$

$$\frac{x_{g_1}}{N^2}\left(\frac{p_{h_1}}{x_{h_1}} + \frac{p_{h_2}}{x_{h_2}}\right) \leq 1 \quad \left(\begin{array}{l}\text{using the } |\psi\rangle\langle\psi| \text{ trick} \\ \text{and demanding } 1 \geq \langle\psi|\psi\rangle\end{array}\right).$$

  The last statement is the same constraint for a split as the one derived by Mochon. This also readily generalises to the case of $1 \to N$ splits which again we defer to the appendix.

- General $m \to n$: $g_1, g_2 \ldots g_m \to h_1, h_2 \ldots h_n$
  It is not too hard to show that in general one obtains the constraint

$$\frac{1}{\langle x_g \rangle} \geq \left\langle \frac{1}{x_h} \right\rangle$$

  using the appropriate blinkered unitary (which also we show in the appendix).

This class of unitary is enough to convert the 1/6 game into an explicit protocol. However, for games given by Mochon that go beyond 1/6 this class falls short. One way of seeing this is that the general $m \to n$ blinkered transition effectively behaves like an $m \to 1$ merge followed by a $1 \to n$ split, which are a set of moves that are insufficient to break the 1/6 limit (at least using Mochon's games).

3 Games and Protocols

[Introduce the simplest game of raising and obtaining the trivial bias. Then the game with one split and two merges that I think gives the SR protocol like bias. Perhaps show that this can be explicitly converted into a protocol using the aforesaid framework. Then, following Mochon, introduce the infinite limit like procedure, maybe only state you get only 1/6 in the limit using the constraints for a merge and split.]

3.1 Mochon's Approach

3.1.1 Assignments

Recall that a function

$$\sum_{z\in\{x_1,x_2,\dots x_n\}} p(z)[z]$$

is valid if

$$\sum_{z\in\{x_1,x_2,\dots x_n\}} \left(\frac{-1}{\lambda+z}\right) p(z) \geq 0, \quad \sum_{z\in\{x_1\dots x_n\}} p(z) = 0$$

for all $\lambda > 0$ where $x_i \geq 0$. Checking if a generic assignment for p satisfies these infinite constraints in not always easy. Mochon had used a constructive approach here and we will build on to it. Let us state these results with some precision (proven in the appendix, well most) where as above n numbers are assumed to be represented by x_i and each $x_i \geq 0$.

Lemma (Mochon's Denominator)**.** $\sum_{i=1}^n \frac{1}{\prod_{j\neq i}(x_j - x_i)} = 0$ *for* $n \geq 2$.

Lemma (Mochon's f-assignment Lemma)**.** $\sum_{i=1}^n \frac{f(x_i)}{\prod_{j\neq i}(x_j - x_i)} = 0$ *where* $f(x_i)$ *is a polynomial of order* $k \leq n-2$.

Definition (Mochon's TIPG assignment)**.** Given a set of n points $0 < x_1 < x_2 \cdots < x_n$, a polynomial $f(x)$ with order k at most $n-2$ and $f(-\lambda) \geq 0$ for all $\lambda \geq 0$, the probability weights for a TIPG assignment is $p(x_i) = -\frac{f(x_i)}{\prod_{j\neq i}(x_j - x_i)}$.

Mochon was able to show that 'Mochon's TIPG assignment' makes for a valid function (in the TIPG formalism), given by

$$\sum_{i=1}^n p(x_i)[x_i, y]$$

where the notion of validity has been extended to a pair of points. As we will see soon, the power of this construction lies in the fact that we can easily construct polynomials that have roots at arbitrary locations. This allows us to create interesting repeating structures called ladders (due to Mochon) which we can terminate using these polynomials to obtain a game with a finite set of points. These ladders play a pivotal role in achieving smaller biases and the ability to obtain finite ladders is essential for being able to obtain a physical process that would yield the said bias.

We now build a little on Mochon's notation and results.

Definition (Mochon's TDPG assignment)**.** Given Mochon's TIPG assignment, let $\{i\}$ be the set of indices for which $p(x_i) < 0$ and $\{k\}$ be the remaining indices with respect to $\{1, 2, \dots n\}$. The TDPG assignment (in accordance with the notation used in TEF) is given as

$$\begin{aligned}
\{x_{g_1}, x_{g_2} \dots\} &= \{x_i\} \\
\{p_{g_1}, p_{g_2} \dots\} &= \{-p(x_i)\} \\
\{x_{h_1}, x_{h_2} \dots\} &= \{x_k\} \\
\{p_{h_1}, p_{h_2} \dots\} &= \{p(x_k)\}.
\end{aligned}$$

With these in place we make some observations about how initial and final averages behave under such an assignment.

Proposition. $N_h^2 = N_g^2$ *where* $N_g^2 = \sum p_{g_i}$ *and* $N_h^2 = \sum p_{h_i}$ *for Mochon's TDPG assignment.*

Proof. We have to show that $N_h^2 - N_g^2 = \sum p_{h_i} - \sum p_{g_i} = 0$ which is the same as showing $\sum_{i=1}^n p(x_i) = 0$ which holds because we just showed that $\sum_{i=1}^n f(x_i)/\prod_{j\neq i}(x_j - x_i) = 0$ (Mochon's f-assignment Lemma). □

Proposition. $\langle x_h \rangle - \langle x_g \rangle = 0$ *for a Mochon's TDPG assignment with* $k \leq n-3$ *where* $\langle x_h \rangle = \frac{1}{N_h^2}\sum p_{h_i} x_{h_i}$ *and* $\langle x_g \rangle = \frac{1}{N_g^2}\sum p_{g_i} x_{g_i}$.

Proof. This is a direct consequence of Mochon's f-assignment lemma. Let h be the $n-3$ order polynomial defined by Mochon's TDPG assignment so that $\langle x_h \rangle - \langle x_g \rangle \propto \sum p_{h_i} x_i - \sum p_{g_i} x_{g_i} = \sum_{i=1}^n p(x_i) x_i = \sum_{i=1}^n \frac{h(x_i) x_i}{\prod_{j\neq i}(x_j - x_i)} = \sum_{i=1}^n \frac{f(x_i)}{\prod_{j\neq i}(x_j - x_i)} = 0$ because f is an $n-2$ order polynomial. □

Lemma. $\sum_{i=1}^{n}\frac{x_i^{n-1}}{\prod_{j\neq i}(x_j-x_i)}=(-1)^{n-1}$ *for* $n\geq 2$.

Proposition. $\langle x_h\rangle-\langle x_g\rangle=\frac{1}{N_h^2}=\frac{1}{N_g^2}$ *for a Mochon's TDPG assignment with* $k=n-2$ *and coefficient of* x^{n-2} ± 1 *in* $f(x)$. *As above here* $\langle x_h\rangle=\frac{1}{N_h^2}\sum p_{h_i}x_{h_i}$ *and* $\langle x_g\rangle=\frac{1}{N_g^2}\sum p_{g_i}x_{g_i}$.

We will see that typically $N_h=N_g$ are quite large and the average only slightly increase, if at all. We are now in a position to discuss Mochon's games.

3.1.2 Typical Game Structure

We assume an equally spaced n-point lattice given by $x_j=x_0+j\delta x$ where $\delta x=\delta y$ is small and x_0 would essentially give a bound on P_B^* which will be determined by following the constraints; similarly $y_j=y_0+j\delta y$ and we also define $\Gamma_{k+1}=y_{n-k}=x_{n-k}$. Let $P(x_j)$ be the probability weight associated with the point $[x_j,0]$ s.t.

$$\sum_{i=1}^{n}P(x_j)=\frac{1}{2},\sum_{j=1}^{n}\frac{P(x_j)}{x_j}=\frac{1}{2}.$$

Similarly with the point $[0,y_j]$ we associate $P(y_j)$ where $y_j=x_j$ as we will also assume that $x_0=y_0$. These choices explicitly impose symmetry between Alice and Bob which in turn entails that we have to do only half the analysis.

3.2 Bias 1/6

3.2.1 Game

With reference to figure 1b (note the coordinate is supposed to be x_{j-1} not x_{j+1}) we need to satisfy $P(x_{j-1})+P_1(y_j)=P_1(x_{j-1})$ which is probability conservation and $P_1(y_j)y_j\leq P_1(x_{j-1})y_{j-2}$ which is the merge condition. Both of these are automatically satisfied if we make a Mochon's denominator based assignment as follows

$$\begin{aligned}0&\leftrightarrow x_{g_1}\\ y_j&\leftrightarrow x_{g_2}\\ y_{j-2}&\leftrightarrow x_{h_1}\end{aligned}$$

$$\begin{aligned}P(x_{j-1})\leftrightarrow p_{g_1}&=\frac{c(x_{j-1})}{y_jy_{j-2}}\\ P_1(y_j)\leftrightarrow p_{g_2}&=\frac{c(x_{j-1})}{(y_j-y_{j-2})(y_j)}=\frac{c(x_{j-1})}{2y_j\delta y}\\ P_1(x_{j-1})\leftrightarrow p_{h_1}&=\frac{c(x_{j-1})}{(y_j-y_{j-2})(y_{j-2})}=\frac{c(x_{j-1})}{2y_{j-2}\delta y}\end{aligned}$$

where the function $c(x_{j-1})$ must be chosen so that $P_1(y_j)=P_1(x_j)$ which entails

$$\frac{c(x_{j-1})}{\cancel{2}y_j\cancel{\delta y}}=\frac{c(x_j)}{\cancel{2}y_{j-1}\cancel{\delta y}}$$

and that in turn is solved by $c(x_j)=\frac{c_0\delta x}{x_j}$ where we used $x_j=y_j$, $\delta x=\delta y$ and added a δx to make $\sum P(x_j)$ into an integral. Plugging this back we have

$$P_1(x_j)=\frac{c_0}{2x_jx_{j-1}},\ P(x_j)=\frac{c_0\delta x}{x_{j-1}x_jx_{j+1}}$$

Since they involve a sum we will do this in the limit $\delta x\to 0$ and $\Gamma\to\infty$ to avoid dealing with summing a series.

$$\sum_{j=0}^{n}P(x_j)=\frac{1}{2}\to x_0^2\int_{x_0}^{\Gamma}\frac{dx}{x^3}=\frac{x_0^2}{(-2)}\left[\frac{1}{\Gamma^2}-\frac{1}{x_0^2}\right]=\frac{1}{2}$$

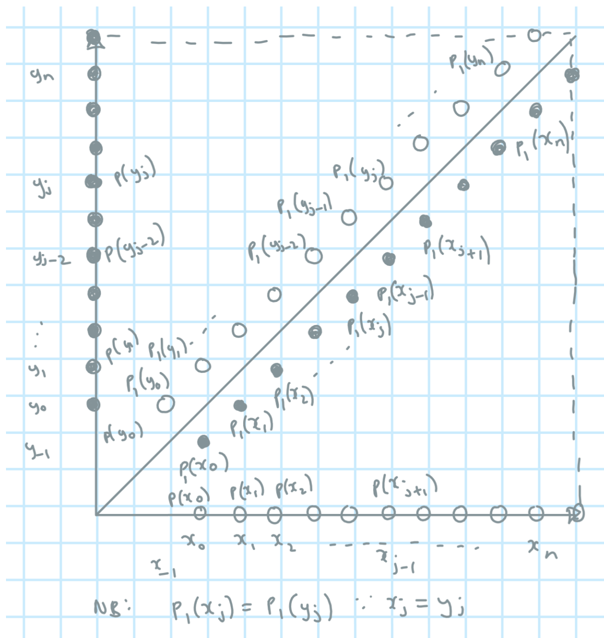
which works as a consistency check because we already deduced that $P_1(x_0)=\frac{1}{2}$. The next condition will yield x_0

$$\sum_{j=0}^{n}\frac{P(x_j)}{x_j}=\frac{1}{2}\to x_0^2\int_{x_0}^{\Gamma}\frac{dx}{x^4}=\frac{x_0^2}{(-3)}\left[\frac{1}{\Gamma^3}-\frac{1}{x_0^3}\right]=\frac{1}{3x_0}=\frac{1}{2}$$

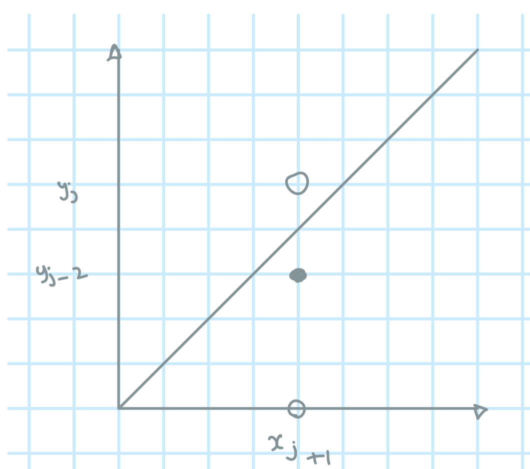
which means

$$x_0=\frac{2}{3}\implies\epsilon=\frac{1}{6}.$$

Of course a more careful analysis must be done to show these things exactly. Aside from the integration step one must also set $c_0(x)=(\Gamma_{n+1}-x)$ in order to terminate the ladder which turns the terminating step on the ladder into a raise. At the moment, however, we satisfy ourselves with this and move on to the more interesting 1/10 game.



(a) The non-trivial step of the protocol



(b) 2 → 1 move illustrated

Figure 1: Building a TDPG/TIPG using merge moves

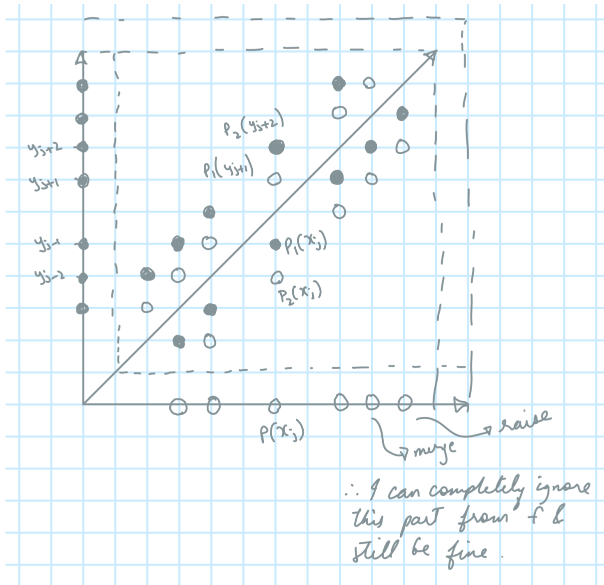


Figure 2: Corrected 1/10 game: The $--+-+$ $3 \to 2$ move based TIPG

3.2.2 Protocol

One can now trivially construct the explicit protocol using the Blinkered Unitaries and the TEF introduced above. One starts with a split, then a raise by Alice and Bob, followed by a merge by Bob, then a merge by Alice and so on until only two points remain. Bob can also start as the description is symmetric. These two can then be raised to the same location and merged. The coordinates of these points will tend to $[\frac{2}{3}, \frac{2}{3}]$ as calculated above. The only creative part left would be the choice of labels that make the description neater from the point of view of the explicit protocol (perhaps show this construction).

3.3 Bias 1/10 Game

With respect to figure 2 we use Mochon's assignment with $f(y_i) = (y_{-2} - y_i)(\Gamma_1 - y_i)(\Gamma_2 - y_i)$ as

$$\frac{f(y_j)c'(x_j)}{\prod_{k\neq j}(y_k - y_j)}.$$

Following the scheme as described above the probabilities become

$$P_2(y_{j+2}) = \frac{-f(y_{j+2})c(x_j)}{4.3(\delta y)^2 y_{j+2}}$$
$$P_1(y_{j+1}) = \frac{-f(y_{j+1})c(x_j)}{3.2(\delta y)^2 y_{j+1}}$$
$$P_1(x_j) = \frac{-f(y_{j-1})c(x_j)}{3.2(\delta y)^2 y_{j-1}}$$
$$P_2(x_j) = \frac{-f(y_{j-2})c(x_j)}{4.3(\delta y)^2 y_{j-2}}$$
$$P(x_j) = \frac{f(0)c(x_j)\delta y}{y_{j+2}y_{j+1}y_{j-1}y_{j-2}}$$

where I added the minus sign to account for the fact that f will be negative for coordinates between y_{-2} and Γ_1. Imposing the symmetry constraints $P_1(y_j) = P_1(x_j)$ I get

$$\frac{f(y_j)c(x_{j-1})}{\cancel{3.2(\delta y)^2}y_j} = \frac{f(y_{j-1})c(x_j)}{\cancel{3.2(\delta y)^2}y_{j-1}}$$

which means

$$c(x_j) = \frac{c_0 f(x_j)}{x_j}$$

where c_0 is a constant. This also entails that $P_2(y_j) = P_2(x_j)$, the second symmetry constraint. Finally we can evaluate

$$P(x_j) = \frac{f(0)f(x_j)\delta x}{x_{j+2}x_{j+1}x_j x_{j-1}x_{j-2}} = \frac{c_0 x_0(x_0 - x_j)}{x_j^5}\delta x + \mathcal{O}(\delta x^2)$$

which means that

$$\sum P(x_j) = \frac{1}{2} = \sum \frac{P(x_j)}{x_j} \to \int_{x_0}^{\Gamma} \frac{(x_0 - x)dx}{x^5} = \int_{x_0}^{\Gamma} \frac{(x_0 - x)dx}{x^6}.$$

We can evaluate this as

$$\begin{aligned}
x_0 \int_{x_0}^{\Gamma} \left(\frac{1}{x^5} - \frac{1}{x^6}\right) dx &= \int_{x_0}^{\Gamma} \left(\frac{1}{x^4} - \frac{1}{x^5}\right) dx \\
\left[\frac{1}{4x_0^3} - \frac{1}{5x_0^4}\right] &= \left[\frac{1}{3x_0^3} - \frac{1}{4x_0^4}\right] \\
\left[\frac{1}{4} - \frac{1}{3}\right] &= \left[\frac{1}{5} - \frac{1}{4}\right] \frac{1}{x_0} \\
x_0 \frac{3-4}{\cancel{4}.3} &= \frac{4-5}{5.\cancel{4}} \\
x_0 = \frac{3}{5} &\implies \epsilon = \frac{3}{5} - \frac{1}{2} = \frac{1}{10}.
\end{aligned}$$

3.4 Bias 1/10 Protocol

3.4.1 The $3 \to 2$ Move on the Axis

In this section we will introduce as many parameters as possible within TEF to implement the largest class of $3 \to 2$ moves. However, we will use our insight to choose an appropriate basis so that the parameters are small which in turn simplifies the analysis.

Recall that

$$|v\rangle = \frac{\sqrt{p_{g_1}}\,|g_{11}\rangle + \sqrt{p_{g_2}}\,|g_{22}\rangle + \sqrt{p_{g_3}}\,|g_{33}\rangle}{N_g}$$

and let

$$\begin{aligned}
|v_1\rangle &= \frac{\sqrt{p_{g_3}}\,|g_{22}\rangle - \sqrt{p_{g_2}}\,|g_{33}\rangle}{N_{v_1}} \\
|v_2\rangle &= \frac{-\frac{(p_{g_2}+p_{g_3})}{\sqrt{p_{g_1}}}\,|g_{11}\rangle + \sqrt{p_{g_2}}\,|g_{22}\rangle + \sqrt{p_{g_3}}\,|g_{33}\rangle}{N_{v_2}}
\end{aligned}$$

where $N_{v_1}^2 = p_{g_3} + p_{g_2}$ and $N_{v_2}^2 = \frac{(p_{g_2}+p_{g_3})^2}{p_{g_1}} + p_{g_2} + p_{g_3}$. Recall that

$$\begin{aligned}
|w\rangle &= \frac{\sqrt{p_{h_1}}\,|h_{11}\rangle + \sqrt{p_{h_2}}\,|h_{22}\rangle}{N_h} \\
|w_1\rangle &= \frac{\sqrt{p_{h_2}}\,|h_{11}\rangle - \sqrt{p_{h_1}}\,|h_{22}\rangle}{N_h}.
\end{aligned}$$

Now we define

$$\begin{aligned}
|v_1'\rangle &= \cos\theta\,|v_1\rangle + \sin\theta\,|v_2\rangle \\
|v_2'\rangle &= \sin\theta\,|v_1\rangle - \cos\theta\,|v_2\rangle
\end{aligned}$$

where we know (from hindsight) that $\cos\theta \approx 1$. The full unitary (which is manifestly unitary) we define to be

$$U = |w\rangle\langle v| + (\alpha\,|v_1'\rangle + \beta\,|w_1\rangle)\langle v_1'| + |v_2'\rangle\langle v_2'| + (\beta\,|v_1'\rangle - \alpha\,|w_1\rangle)\langle w_1| + |v\rangle\langle w|$$

where $|\alpha|^2 + |\beta|^2 = 1$ for $\alpha, \beta \in \mathbb{C}$. There is some freedom in choosing U in the sense that $\alpha\,|v\rangle + \beta\,|w_1\rangle$ would also work (then $|v\rangle\langle w| \to |v_1\rangle\langle w|$) because these don't influence the constraint equation. That is what we evaluate now. We will need terms of the form $EU\,|g_{ii}\rangle$ with $E = \mathbb{I}^{\{h_{ii}\}} + \mathbb{I}^{\{k_{ii}\}}$ but we neglect the $|k_{ii}\rangle$ part as it doesn't participate in the dynamics. This entails that on the $\{|g_{ii}\rangle\}$ space

$$\begin{aligned}
E_h U E_g &= |w\rangle\langle v| + \beta\,|w_1\rangle\langle v_1'| \\
&= |w\rangle\langle v| + \beta\,|w_1\rangle(\cos\theta\,\langle v_1| + \sin\theta\,\langle v_2|).
\end{aligned}$$

Consequently I have

$$\begin{aligned}
E_h U\,|g_{11}\rangle &= \frac{\sqrt{p_{g_1}}}{N_g}\,|w\rangle + \left[\cos\theta.0 - \sin\theta\frac{p_{g_2}+p_{g_3}}{\sqrt{p_{g_1}}N_{v_2}}\right]\beta\,|w_1\rangle \\
E_h U\,|g_{22}\rangle &= \frac{\sqrt{p_{g_2}}}{N_g}\,|w\rangle + \left[\cos\theta\frac{\sqrt{p_{g_3}}}{N_{v_1}} + \sin\theta\frac{\sqrt{p_{g_2}}}{N_{v_2}}\right]\beta\,|w_1\rangle \\
E_h U\,|g_{33}\rangle &= \frac{\sqrt{p_{g_3}}}{N_g}\,|w\rangle + \left[-\cos\theta\frac{\sqrt{p_{g_2}}}{N_{v_1}} + \sin\theta\frac{\sqrt{p_{g_3}}}{N_{v_2}}\right]\beta\,|w_1\rangle.
\end{aligned}$$

Recall that the constraint equation was

$$\sum x_{h_i} \left|h_{ii}\right\rangle \left\langle h_{ii}\right| - \sum x_{g_i} E_h U \left|g_{ii}\right\rangle \left\langle g_{ii}\right| U^\dagger E_h \geq 0$$

where the first sum becomes

$$\begin{bmatrix} \langle x_h \rangle & \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1} - x_{h_2}) \\ \text{h.c.} & \frac{p_{h_2}x_{h_1}+p_{h_1}x_{h_2}}{N_h^2} \end{bmatrix}$$

in the $\left|w\right\rangle, \left|w_1\right\rangle$ basis. Since we plan to use the $3 \to 2$ move with one point on the axis, we take $x_{g_1} = 0$. Consequently we need only evaluate

$$x_{g_2} E_h U \left|g_{22}\right\rangle \left\langle g_{22}\right| U^\dagger E_h \dot{=} x_{g_2} \begin{bmatrix} \frac{p_{g_2}}{N_g^2} & \beta\left(\cos\theta\frac{\sqrt{p_{g_3}p_{g_2}}}{N_g N_{v_1}} + \sin\theta\frac{p_{g_2}}{N_g N_{v_2}}\right) \\ \text{h.c.} & \left(\cos\frac{\sqrt{p_{g_3}}}{N_{v_1}} + \sin\theta\frac{\sqrt{p_{g_2}}}{N_{v_2}}\right)^2 |\beta|^2 \end{bmatrix}$$

$$x_{g_3} E_h U \left|g_{33}\right\rangle \left\langle g_{33}\right| U^\dagger E_h \dot{=} x_{g_3} \begin{bmatrix} \frac{p_{g_3}}{N_g^2} & \beta\left(-\cos\theta\frac{\sqrt{p_{g_2}p_{g_3}}}{N_g N_{v_1}} + \sin\theta\frac{p_{g_3}}{N_g N_{v_2}}\right) \\ \text{h.c.} & \left(-\cos\frac{\sqrt{p_{g_2}}}{N_{v_1}} + \sin\frac{\sqrt{p_{g_3}}}{N_{v_2}}\right)^2 |\beta|^2 \end{bmatrix}$$

which means that the constraint equation becomes

$$\begin{bmatrix} \langle x_h \rangle - \langle x_g \rangle & \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1} - x_{h_2}) - \beta\cos\theta\frac{\sqrt{p_{g_2}p_{g_3}}}{N_g N_{v_1}}(x_{g_2} - x_{g_3}) - \beta\sin\theta\left\langle x_g\right\rangle\frac{N_g}{N_{v_2}} \\ \text{h.c.} & \frac{p_{h_2}x_{h_1}+p_{h_1}x_{h_2}}{N_h^2} - |\beta|^2\left[\frac{\cos^2\theta}{N_{v_1}^2}(p_{g_3}x_{g_2} + p_{g_2}x_{g_3}) + \frac{\sin^2\theta}{\left(N_{v_2}^2/N_g^2\right)}\left\langle x_g\right\rangle + \frac{2\cos\theta\sin\theta\sqrt{p_{g_3}p_{g_2}}}{N_{v_1}N_{v_2}}(x_{g_2} - x_{g_3})\right] \end{bmatrix} \geq 0.$$

We already showed that Mochon's transition is average non-decreasing viz. $\langle x_h \rangle - \langle x_g \rangle \geq 0$. We will set the off-diagonal elements of the matrix above to zero and show that the second diagonal element, the second eigenvalue therefore, is positive.

Setting the off-diagonal to zero one can obtain θ by solving the quadratic in terms of β although the expression will not be particularly pretty. To establish existence and positivity we need to simplify our expressions.

So far everything was exact even though the choice of basis and techniques were chosen based on experience. Now we claim that $\theta \approx \mathcal{O}(\delta y)$ at most (where $\delta y = \delta x$ is the lattice spacing) and since δy will be taken to be small we can take the small θ limit and to first order in θ the constraints become

$$\frac{\frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1} - x_{h_2}) - \beta\frac{\sqrt{p_{g_2}p_{g_3}}}{N_g N_{v_1}}(x_{g_2} - x_{g_3})}{\beta\left\langle x_g\right\rangle\frac{N_g}{N_{v_2}}} = \theta + \mathcal{O}(\delta y^2)$$

and

$$\frac{p_{h_2}x_{h_1} + p_{h_1}x_{h_2}}{N_h^2} - |\beta|^2\left[\frac{p_{g_3}x_{g_2} + p_{g_2}x_{g_3}}{N_{v_1}^2} + 2\theta\frac{\sqrt{p_{g_3}p_{g_2}}}{N_{v_1}N_{v_2}}(x_{g_2} - x_{g_3})\right] + \mathcal{O}(\delta y^2) \geq 0.$$

If our claim is wrong when we evaluate θ we will get zero order terms but as we show in the following section $\theta = 0.\delta y + \mathcal{O}(\delta y^2)$ in fact.

3.4.2 Validity of the $3 \to 2$ Move

With respect to figure 2 we have

$$P_2(y_{j+2}) = p_{h_2} = \frac{-f(y_{j+2})}{4.3\delta y^2 y_{j+2}}$$

$$P_1(y_{j+1}) = p_{g_3} = \frac{-f(y_{j+1})}{3.2\delta y^2 y_{j+1}}$$

$$P_1(x_j) = p_{h_1} = \frac{-f(y_{j-1})}{3.2\delta y^2 y_{j-1}}$$

$$P_2(x_j) = p_{g_2} = \frac{-f(y_{j-2})}{4.3\delta y^2 y_{j-2}}$$

$$P(x_j) = p_{g_1} = \frac{f(0)\delta y}{y_{j+2}y_{j+1}y_{j-1}y_{j-2}}$$

where we assumed $f(0) > 0$ and $f(y) < 0$ for $y > y_0'$, $y_0' = y_0 + \delta y$. We also scaled by δy to make $P(x_j)$ into a nice density. So far everything is exact. We will now convert all expressions to first order in δy. To this end we note

$$f(y_{j+m}) = f(y_j) + \frac{\partial f}{\partial y} m\delta y + \mathcal{O}(\delta y^2)$$

$$\frac{1}{y_{j+m}} = (y_j + m\delta y)^{-1} = \frac{1}{y_j}\left(1 + m\frac{\delta y}{y_j}\right)^{-1} = \frac{1}{y_j} - m\frac{\delta y}{y_j^2} + \mathcal{O}(\delta y^2)$$

where $\frac{\partial f}{\partial y}$ refers to $\frac{\partial f(y)}{\partial y}|_{y_j}$. We define and evaluate

$$
\begin{aligned}
P_k^m &= \frac{-f(y_{j+m})}{k\delta y^2 y_{j+m}} \\
&= \frac{1}{k\delta y^2}\left[-f(y_j) - \frac{\partial f}{\partial y} m\delta y + \mathcal{O}(\delta y^2)\right]\left[\frac{1}{y_j} - m\frac{\delta y}{y_j^2} + \mathcal{O}(\delta y^2)\right] \\
&= \frac{1}{k\delta y^2}\left[-\frac{f}{y_j} - m\frac{\delta y}{y_j}\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right) + \mathcal{O}(\delta y^2)\right] \\
&= \frac{1}{k y_j \delta y^2}\left[-f - m\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right) + \mathcal{O}(\delta y^2)\right]
\end{aligned}
$$

where f means $f(y_j)$. In this notation

$$
\begin{aligned}
p_{h_2} &= P_{12}^2,\ p_{h_1} = P_6^{-1} \\
p_{g_2} &= P_{12}^{-2},\ p_{g_3} = P_6^1.
\end{aligned}
$$

With an eye at the off-diagonal condition we evaluate

$$
P_{k_1}^{m_1} P_{k_2}^{m_2} = \frac{1}{k_1 k_2}\left(\frac{1}{y_j\delta y^2}\right)^2\left[f^2 + f\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right)(m_1 + m_2) + \mathcal{O}(\delta y^2)\right]
$$

and

$$
P_{k_1}^{m_1} + P_{k_2}^{m_2} = \frac{1}{y_j\delta y^2}\left[-\left(\frac{1}{k_1} + \frac{1}{k_2}\right) f - \left(\frac{m_1}{k_1} + \frac{m_2}{k_2}\right)\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right) + \mathcal{O}(\delta y^2)\right].
$$

We now evaluate

$$
\begin{aligned}
\sqrt{p_{h_1} p_{h_2}} &= \sqrt{P_{12}^2 P_6^{-1}} = \frac{1}{y_j\delta y^2}\sqrt{\frac{1}{12.6}\left[f^2 + f\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right) + \mathcal{O}(\delta y^2)\right]} \\
N_h^2 &= P_{12}^2 + P_6^{-1} = \frac{1}{y_j\delta y^2}\left[-\left(\frac{1}{12} + \frac{1}{6}\right) f - \cancelto{0}{\left(\frac{2}{12} - \frac{1}{6}\right)}\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right) + \mathcal{O}(\delta y^2)\right] \\
&= \frac{1}{4 y_j\delta y^2}\left[-f + \mathcal{O}(\delta y^2)\right]
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\sqrt{p_{g_2} p_{g_3}} &= \sqrt{P_{12}^{-2} P_6^{1}} = \frac{1}{y_j\delta y^2}\sqrt{\frac{1}{12.6}\left[f^2 - f\delta y\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right) + \mathcal{O}(\delta y^2)\right]} \\
N_g^2 &= P_{12}^{-2} + P_6^1 + p_{g_1} = \frac{1}{4 y_j\delta y^2}\left[-f + \mathcal{O}(\delta y^2)\right] + \left[\frac{f(0)\delta y}{y_j^4} + \mathcal{O}(\delta y^2)\right] \\
&= \frac{1}{4 y_j\delta y^2}\left[-f + \mathcal{O}(\delta y^2)\right] \\
N_{v_1}^2 &= \frac{1}{4 y_j\delta y^2}\left[-f + \mathcal{O}(\delta y^2)\right]
\end{aligned}
$$

where even though it seems like we have neglected p_{g_1} when we take the ratios the meaning of keeping first order in δy would become precise. We can actually take $\beta = 1$ and obtain

$$
\begin{aligned}
\theta &= \frac{4\sqrt{\frac{1}{12.6}}(-3\delta y)\left[f.(\cancel{1} + \frac{\delta y}{2f}\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right)) - f.(\cancel{1} - \frac{\delta y}{2f}\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right)) + \mathcal{O}(\delta y^2)\right]}{\langle x_g\rangle \frac{N_g}{N_{v_2}}} \\
&= 0 + \mathcal{O}(\delta y^2).
\end{aligned}
$$

This shows that to first order the off-diagonal term is zero for $\theta = 0$.

Now we will show that the second diagonal element is positive to first order in δy. Using the fact that $\theta = \mathcal{O}(\delta y^2)$ we have

$$
\frac{p_{h_2} x_{h_1} + p_{h_1} x_{h_2}}{N_h^2} - \frac{p_{g_3} x_{g_2} + p_{g_2} x_{g_3}}{N_{v_1}^2} + \mathcal{O}(\delta y^2) \geq 0
$$

as the positivity condition. This becomes

$$
\begin{aligned}
=&\frac{P_{12}^2 y_{j-1}+P_6^{-1}y_{j+2}}{N_h^2}-\frac{P_6^1 y_{j-2}+P_{12}^{-2}y_{j+1}}{N_{v_1}^2}+\mathcal{O}(\delta y^2)\\
=&\left(\frac{4y_j\delta y^2}{-f}\right)\frac{1}{y_j\delta y^2}\\
&\left\{\frac{1}{12}\left[-f-2\delta y\gamma\right](y_j-\delta y)+\frac{1}{6}\left[-f+\gamma\delta y\right](y_j+2\delta y)-\left(\frac{1}{6}\left[-f-\delta y\gamma\right](y_j-2\delta y)+\frac{1}{12}\left[-f+2\delta y\gamma\right](y_j+\delta y\right)\right\}+\mathcal{O}(\delta y^2)\\
=&\frac{-2}{3f}\left\{\frac{1}{2}(\cancel{-fy}+f\delta y-2y\delta y\gamma)+(\cancel{-fy}-2f\delta y+y\delta y\gamma)-\left((\cancel{-fy}+2f\delta y-y\delta y\gamma)+\frac{1}{2}\left(\cancel{-fy}-f\delta y+2y\delta y\gamma\right)\right)\right\}+\mathcal{O}(\delta y^2)\\
=&\frac{-2}{3f}\left\{(f\delta y-2y\delta y\gamma)+2(-2f\delta y+y\delta y\gamma)\right\}+\mathcal{O}(\delta y^2)\\
=&\frac{-2}{3f}\left\{-3f\delta y\right\}+\mathcal{O}(\delta y^2)=2\delta y+\mathcal{O}(\delta y^2)\geq 0
\end{aligned}
$$

where $\gamma=\left(\frac{\partial f}{\partial y}-\frac{f}{y_j}\right)$ and we suppressed the index j in y_j for simplicity. This establishes the validity of the $3\to 2$ transition for a closely spaced lattice.

Remarks: Note that only the proof of validity was done to first order in δy. The unitary itself is known exactly (θ can be obtained by solving the quadratic).
Using $f(y)=(y_0'-y)(\Gamma_1-y)(\Gamma_2-y)$ we can implement the last two moves (TODO: argue this properly and write the unitaries). The only remaining task is implementing the $2\to 2$ move in the last step because we implicitly assumed here that $\sqrt{p_{g_2}}\neq 0$ (else the vectors which we assumed are orthonormal, cease to be so).

3.4.3 The $2\to 2$ Move on the Axis and its validity

We claim that the $2\to 2$ move can be implemented using

$$U=\left|w\right\rangle\left\langle v\right|+(\alpha\left|v\right\rangle+\beta\left|w_1\right\rangle)\left\langle v_1\right|+\left|v\right\rangle\left\langle w\right|+(\beta\left|v\right\rangle-\alpha\left|w_1\right\rangle)\left\langle w_1\right|$$

where as before $|\alpha|^2+|\beta|^2=1$,

$$
\begin{aligned}
\left|v\right\rangle&=\frac{1}{N_g}\left(\sqrt{p_{g_1}}\left|g_{11}\right\rangle+\sqrt{p_{g_2}}\left|g_{22}\right\rangle\right),\\
\left|w\right\rangle&=\frac{1}{N_h}\left(\sqrt{p_{h_1}}\left|h_{11}\right\rangle+\sqrt{p_{h_2}}\left|h_{22}\right\rangle\right),\\
\left|v_1\right\rangle&=\frac{1}{N_g}\left(\sqrt{p_{g_2}}\left|g_{11}\right\rangle-\sqrt{p_{g_1}}\left|g_{22}\right\rangle\right),
\end{aligned}
$$

and

$$\left|w_1\right\rangle=\frac{1}{N_h}\left(\sqrt{p_{h_2}}\left|h_{11}\right\rangle-\sqrt{p_{h_1}}\left|h_{22}\right\rangle\right).$$

We evaluate the constraint equation using

$$
\begin{aligned}
E_hU\left|g_{11}\right\rangle&=\frac{\sqrt{p_{g_1}}\left|w\right\rangle+\beta e^{-i\phi_g}e^{i\phi_h}\sqrt{p_{g_2}}\left|w_1\right\rangle}{N_g}\\
E_hU\left|g_{22}\right\rangle&=\frac{\sqrt{p_{g_2}}\left|w\right\rangle-\beta e^{-i\phi_g}e^{i\phi_h}\sqrt{p_{g_1}}\left|w_1\right\rangle}{N_g}
\end{aligned}
$$

and

$$E_hU\left|g_{11}\right\rangle\left\langle g_{11}\right|U^\dagger E_h=\frac{1}{N_g^2}\begin{array}{c|cc} & \langle w| & \langle w_1| \\ \hline |w\rangle & p_{g_1} & \beta e^{i(\phi_h-\phi_g)}\sqrt{p_{g_2}p_{g_1}} \\ |w_1\rangle & \text{h.c.} & |\beta|^2 p_{g_2}\end{array}$$

(similarly for $L\left|g_{22}\right\rangle\left\langle g_{22}\right|L^\dagger$) as

$$\begin{bmatrix}\langle x_h\rangle-\langle x_g\rangle & \frac{1}{N_g^2}\left[\sqrt{p_{h_1}p_{h_2}}(x_{h_1}-x_{h_2})-\beta\sqrt{p_{g_1}p_{g_2}}(x_{g_1}-x_{g_2})\right]\\ \text{h.c.} & \frac{1}{N_g^2}\left[p_{h_2}x_{h_1}+p_{h_1}x_{h_2}-|\beta|^2\left(p_{g_2}x_{g_1}+p_{g_1}x_{g_2}\right)\right]\end{bmatrix}\geq 0$$

where we absorbed the phase freedom in β, a free parameter, which will be fixed shortly. We will use the same strategy as above and take the first diagonal element to be zero. Our burden would be to first show that

$$\sqrt{\frac{p_{h_1}p_{h_2}}{p_{g_1}p_{g_2}}}\frac{(x_{h_1}-x_{h_2})}{(x_{g_1}-x_{g_2})}=\beta\leq 1$$

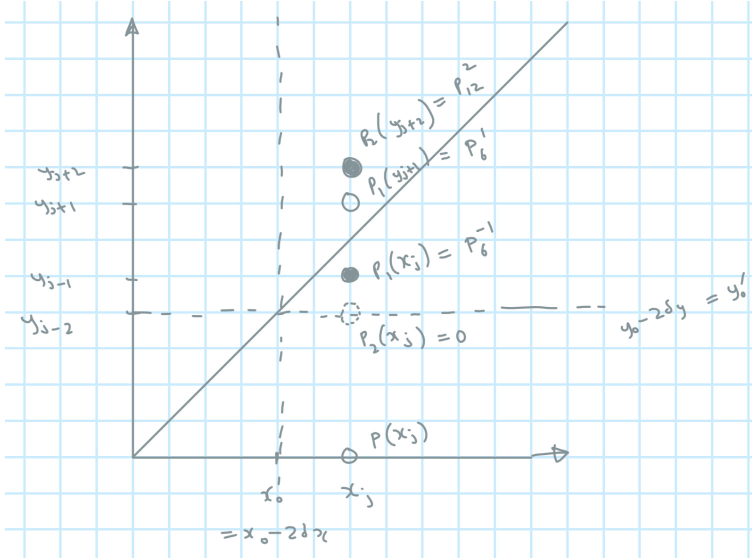


Figure 3: Final Move: The special $2 \to 2$ Transition

and subsequently

$$\frac{1}{N_g^2}\left[p_{h_2}x_{h_1} + p_{h_1}x_{h_2} - |\beta|^2\left(p_{g_2}x_{g_1} + p_{g_1}x_{g_2}\right)\right] \geq 0.$$

What makes this situation special (compared to the $3 \to 2$ merge) is that $f(y_{j-2}) = 0$ which we use to write

$$f(y_{j+k}) = \left.\frac{\partial f}{\partial y}\right|_{y_{j-2}} (k+2)\delta y = -(k+2)\alpha\delta y$$

where

$$\alpha = -\left.\frac{\partial f}{\partial y}\right|_{y_{j-2}} = (\Gamma_1 - y_{j-2})(\Gamma_2 - y_{j-2}).$$

Using the axis situation as depicted in figure 3 we note that

$$\begin{aligned}
p_{h_1} &= P_1(x_j) = \frac{-f(y_{j-1})}{3.2\delta y^2 y_{j-1}} = \frac{\alpha + \mathcal{O}(\delta y)}{6\delta y y_j}\\
p_{h_2} &= P_2(y_{j+2}) = \frac{-f(y_{j+2})}{4.3\delta y^2 y_{j+2}} = \frac{\alpha + \mathcal{O}(\delta y)}{3\delta y y_j}\\
x_{h_1} &= y_{j-1},\, x_{h_2} = y_{j+2}
\end{aligned}$$

while

$$\begin{aligned}
p_{g_1} &= P(x_j) = \frac{f(0)\delta y}{y_{j+2}y_{j+1}y_{j-1}y_{j-2}} = \frac{f(0)\delta y + \mathcal{O}(\delta y^2)}{y_j^4}\\
p_{g_2} &= P_1(y_{j+1}) = \frac{-f(y_{j+1})}{3.2\delta y^2 y_{j+1}} = \frac{\alpha + \mathcal{O}(\delta y)}{2\delta y y_j}\\
x_{g_1} &= 0,\, x_{g_2} = y_{j+1}.
\end{aligned}$$

This entails

$$\begin{aligned}
\beta &= \sqrt{\frac{p_{h_1}p_{h_2}}{p_{g_1}p_{g_2}}\frac{(x_{h_1} - x_{h_2})}{(x_{g_1} - x_{g_2})}} = \sqrt{\frac{\alpha^{\cancel{2}} + \mathcal{O}(\delta y)}{\overset{3}{\cancel{6.3}}\delta y^2 y_j^{\cancel{2}}}\,\frac{\cancel{2\delta y}\cancel{y_j^4}y_j}{\cancel{\delta y}(f(0)\cancel{\alpha} + \mathcal{O}(\delta y))}\,\frac{\cancel{(3\delta y)^2}}{\cancel{y_j^2 + \mathcal{O}(\delta y)}}}\\
&= \sqrt{\frac{y_0'\alpha + \mathcal{O}(\delta y)}{f(0)}} = \sqrt{\frac{(\Gamma_1 - y_{j-2})(\Gamma_2 - y_{j-2}) + \mathcal{O}(\delta y)}{\Gamma_1\Gamma_2}} \leq 1
\end{aligned}$$

where we used $f(0) = y_0'\Gamma_1\Gamma_2$ and assumed δy is small compared Γs (which is the case) for the inequality in the last step to hold.

The second condition can be proven similarly

$$\begin{aligned}
\frac{1}{N_g^2}\left[p_{h_2}x_{h_1}+p_{h_1}x_{h_2}-|\beta|^2\left(p_{g_2}x_{g_1}+p_{g_1}x_{g_2}\right)\right] &\geq \frac{1}{N_g^2}\left[p_{h_2}x_{h_1}+p_{h_1}x_{h_2}-p_{g_2}x_{g_1}\right] \\
&= \frac{1}{N_g^2}\left[\frac{\alpha+\mathcal{O}(\delta y)}{3\delta y y_j}y_{j-1}+\frac{\alpha+\mathcal{O}(\delta y)}{6\delta y y_j}y_{j+2}-\frac{f(0)\delta y+\mathcal{O}(\delta y^2)}{y_j^4}y_{j+1}\right] \\
&= \frac{1}{3\delta y N_g^2}\left[(\alpha+\mathcal{O}(\delta y))\left(\frac{3}{2}\right)-\underbrace{\frac{f(0)\delta y^2+\mathcal{O}(\delta y^3)}{y_j^3}}_{\in\mathcal{O}(\delta y^2)}\right] \\
&= \frac{1}{2\delta y N_g^2}\left[(\Gamma_1-y_{j-2})(\Gamma_2-y_{j-2})+\mathcal{O}(\delta y)\right]\geq 0
\end{aligned}$$

where the last step holds for δy small enough.

4 Conclusion

5 Appendix

5.1 Blinkered $m \to n$ Transition

Recall that the unitary I had described was of the form $U = |w\rangle\langle v| + |v\rangle\langle w| + \sum |v_i\rangle\langle v_i| + \sum |w_i\rangle\langle w_i|$. It is evident that having a scheme for generating these $|v_i\rangle$ and $|w_i\rangle$ will be useful as we explore more complicated transitions. More precisely, I will need to complete a set containing one vector into a complete orthonormal basis. Let me do this first and then return to the analysis of a $3 \to 2$ merge.

Completing an Orthonormal Basis

Consider an orthonormal complete set of basis vectors $\{|g_i\rangle\}$, and a vector $|v\rangle = \frac{\sum_i \sqrt{p_i}|g_i\rangle}{\sqrt{\sum_i p_i}}$. I will describe a scheme for constructing vectors $|v_i\rangle$ s.t. $\{|v\rangle, \{|v_i\rangle\}\}$ is a complete orthonormal set of basis vectors. Formally, I can do this inductively. Instead, I will do this by examples for that makes it intuitive and demonstrates the generalisable argument right away. The first I define to be

$$|v_1\rangle = \frac{\sqrt{p_1}\,|g_1\rangle - \frac{p_1}{\sqrt{p_2}}\,|g_2\rangle}{\sqrt{p_1 + \frac{p_1^2}{p_2}}} \left(= \frac{\sqrt{p_1}\,|g_1\rangle - \sqrt{p_2}\,|g_2\rangle}{\sqrt{p_1+p_2}}, \text{ the familiar one} \right)$$

which is manifestly normalised and orthogonal to $|v\rangle$, i.e. $\langle v|v_1\rangle = p_1 - p_1 = 0$. The next vector is

$$|v_2\rangle = \frac{\sqrt{p_1}\,|g_1\rangle + \sqrt{p_2}\,|g_2\rangle - \frac{(p_1+p_2)}{\sqrt{p_3}}\,|g_3\rangle}{\sqrt{p_1 + p_2 + \frac{(p_1+p_2)^2}{p_3}}}$$

which is again manifestly normalised and orthogonal to $|v_1\rangle$ because $\langle v_2|v_1\rangle = \langle v|v_1\rangle$. $\langle v|v_2\rangle = p_1 + p_2 - (p_1 + p_2) = 0$. Similarly one can construct the $(k+1)^{\text{th}}$ basis vector as

$$|v_k\rangle = \frac{\sum_{i=1}^k \sqrt{p_k}\,|g_k\rangle - \frac{\sum_{i=1}^k p_k}{\sqrt{p_{k+1}}}\,|g_{k+1}\rangle}{N_k}$$

where the $N_k = \sqrt{\sum_{i=1}^k p_k + \frac{(\sum_{i=1}^k p_k)^2}{p_{k+1}}}$ and obtain the full set.

The Analysis

Back to the analysis. Recall that the constraint equation was

$$\underbrace{\sum x_{h_i}\,|h_{ii}\rangle\langle h_{ii}|}_{\text{I}} + \underbrace{x\mathbb{I}^{\{g_{ii}\}}}_{\text{II}} \geq \underbrace{\sum x_{g_i} U\,|g_{ii}\rangle\langle g_{ii}|\,U^\dagger}_{\text{III}}$$

where I have introduced the notation $|h_{ii}\rangle = |h_i h_i\rangle$ in the interest of efficiency. The $g_1, g_2, g_3 \to h_1, h_2$ transition requires me to know

$$U = |v\rangle\langle w| + |w\rangle\langle v| + |v_1\rangle\langle v_1| + |v_2\rangle\langle v_2| + |w_1\rangle\langle w_1|.$$

Using the procedure above I can evaluate the vectors of interest

$$|v\rangle = \frac{\sqrt{p_{g_1}}\,|g_{11}\rangle + \sqrt{p_{g_2}}\,|g_{22}\rangle + \sqrt{p_{g_3}}\,|g_{33}\rangle}{N_g}$$

$$|v_1\rangle = \frac{\sqrt{p_{g_1}}\,|g_{11}\rangle - \frac{p_{g_1}}{\sqrt{p_{g_2}}}\,|g_{22}\rangle}{N_{g_1}}$$

$$|v_2\rangle = \frac{\sqrt{p_{g_1}}\,|g_{11}\rangle + \sqrt{p_{g_2}}\,|g_{22}\rangle - \frac{(p_{g_1}+p_{g_2})}{\sqrt{p_{g_3}}}\,|g_{33}\rangle}{N_{g_2}}$$

$$|w\rangle = \frac{\sqrt{p_{h_1}}\,|h_{11}\rangle + \sqrt{p_{h_2}}\,|h_{22}\rangle}{N_h}$$

$$|w_1\rangle = \frac{\sqrt{p_{h_2}}\,|h_{11}\rangle - \sqrt{p_{h_1}}\,|h_{22}\rangle}{N_h}$$

where N_g, N_{g_1}, N_{g_2}, N_h are normalisations. In fact I want to express the constraints in this basis. To evaluate the first term I use the above to find

$$
\begin{aligned}
|h_{11}\rangle &= \frac{\sqrt{p_{h_1}}\,|w\rangle + \sqrt{p_{h_2}}\,|w_1\rangle}{N_h} \\
|h_{22}\rangle &= \frac{\sqrt{p_{h_2}}\,|w\rangle - \sqrt{p_{h_1}}\,|w_1\rangle}{N_h}
\end{aligned}
$$

which leads to

$$
\begin{aligned}
\mathrm{I} &= x_{h_1}\,|h_{11}\rangle\langle h_{11}| + x_{h_2}\,|h_{22}\rangle\langle h_{22}| \\
&= \frac{x_{h_1}}{N_h^2}\left[\begin{array}{c|cc} & \langle w| & \langle w_1| \\ \hline |w\rangle & p_{h_1} & \sqrt{p_{h_1}p_{h_2}} \\ |w_1\rangle & \sqrt{p_{h_1}p_{h_2}} & p_{h_2} \end{array}\right] + \frac{x_{h_2}}{N_h^2}\left[\begin{array}{c|cc} & \langle w| & \langle w_1| \\ \hline |w\rangle & p_{h_2} & -\sqrt{p_{h_1}p_{h_2}} \\ |w_1\rangle & -\sqrt{p_{h_1}p_{h_2}} & p_{h_1} \end{array}\right] \\
&= \frac{1}{N_h^2}\left[\begin{array}{c|cc} & \langle w| & \langle w_1| \\ \hline |w\rangle & p_{h_1}x_{h_1} + p_{h_2}x_{h_2} & \sqrt{p_{h_1}p_{h_2}}(x_{h_1} - x_{h_2}) \\ |w_1\rangle & \sqrt{p_{h_1}p_{h_2}}(x_{h_1} - x_{h_2}) & p_{h_2}x_{h_1} + p_{h_1}x_{h_2} \end{array}\right].
\end{aligned}
$$

(Remark: I had made a mistake in this term which was causing the matrix to sometimes become negative; after correction, the matrix seems to be positive for Mochon's f-function based construction) Evaluation of II is nearly trivial for identity can be expressed in any basis and that yields

$$
\begin{aligned}
\mathrm{II} &= x(|v\rangle\langle v| + |v_1\rangle\langle v_1| + |v_2\rangle\langle v_2|) \\
&= \left[\begin{array}{c|ccc} & \langle v| & \langle v_1| & \langle v_2| \\ \hline |v\rangle & x & & \\ |v_1\rangle & & x & \\ |v_2\rangle & & & x \end{array}\right].
\end{aligned}
$$

For the last term

$$
\mathrm{III} = \underbrace{x_{g_1}U\,|g_{11}\rangle\langle g_{11}|\,U^\dagger}_{\text{(i)}} + \underbrace{x_{g_2}U\,|g_{22}\rangle\langle g_{22}|\,U^\dagger}_{\text{(ii)}} + \underbrace{x_{g_3}U\,|g_{33}\rangle\langle g_{33}|\,U^\dagger}_{\text{(iii)}}
$$

I evaluate

$$
\begin{aligned}
U\,|g_{11}\rangle &= \frac{\sqrt{p_{g_1}}}{N_g}\,|w\rangle + \frac{\sqrt{p_{g_1}}}{N_{g_1}}\,|v_1\rangle + \frac{\sqrt{p_{g_1}}}{N_{g_2}}\,|v_2\rangle \\
U\,|g_{22}\rangle &= \frac{\sqrt{p_{g_2}}}{N_g}\,|w\rangle + \frac{\left(-\frac{p_{g_1}}{\sqrt{p_{g_2}}}\right)}{N_{g_1}}\,|v_1\rangle + \frac{\sqrt{p_{g_2}}}{N_{g_2}}\,|v_2\rangle \\
U\,|g_{33}\rangle &= \frac{\sqrt{p_{g_3}}}{N_g}\,|w\rangle + 0\,|v_1\rangle + \frac{\left(-\frac{p_{g_1}+g_{g_2}}{\sqrt{p_{g_3}}}\right)}{N_{g_2}}\,|v_2\rangle\,.
\end{aligned}
$$

I must now find each sub term, starting with the most regular

$$
\text{(i)} = x_{g_1}p_{g_1}\left[\begin{array}{c|ccc} & \langle v_1| & \langle v_2| & \langle w| \\ \hline |v_1\rangle & \frac{1}{N_{g_1}^2} & \frac{1}{N_{g_1}N_{g_2}} & \frac{1}{N_{g_1}N_g} \\ |v_2\rangle & \frac{1}{N_{g_2}N_{g_1}} & \frac{1}{N_{g_2}^2} & \frac{1}{N_{g_2}N_g} \\ |w\rangle & \frac{1}{N_gN_{g_1}} & \frac{1}{N_gN_{g_2}} & \frac{1}{N_g^2} \end{array}\right].
$$

For the second term, I re-write $U\,|g_{22}\rangle = \sqrt{p_{g_2}}\left(\frac{1}{N_g}\,|w\rangle - \frac{1}{N'_{g_1}}\,|v_1\rangle + \frac{1}{N_{g_2}}\,|v_2\rangle\right)$ where I have defined

$$
N'_{g_1} = \frac{p_{g_2}}{p_{g_1}}N_{g_1}
$$

to obtain

$$
(ii) = x_{g_2}p_{g_2}\left[\begin{array}{c|ccc} & \langle v_1| & \langle v_2| & \langle w| \\ \hline |v_1\rangle & \frac{1}{N'^2_{g_1}} & -\frac{1}{N'_{g_1}N_{g_2}} & -\frac{1}{N'_{g_1}N_g} \\ |v_2\rangle & -\frac{1}{N_{g_2}N'_{g_1}} & \frac{1}{N_{g_2}^2} & \frac{1}{N_{g_2}N_g} \\ |w\rangle & -\frac{1}{N_gN'_{g_1}} & \frac{1}{N_gN_{g_2}} & \frac{1}{N_g^2} \end{array}\right]
$$

and finally $U\left|g_{33}\right\rangle = \sqrt{p_{g_3}}\left(\frac{1}{N_g}\left|w\right\rangle + 0\left|v_1\right\rangle - \frac{1}{N'_{g_2}}\left|v_2\right\rangle\right)$ with

$$N'_{g_2} = \frac{p_{g_3}}{p_{g_1}+p_{g_2}}$$

to get

$$(\mathrm{iii}) = x_{g_3}p_{g_3}\left[\begin{array}{c|ccc} & \langle v_1| & \langle v_2| & \langle w| \\ \hline |v_1\rangle & & & \\ |v_2\rangle & & \frac{1}{N'^2_{g_2}} & -\frac{1}{N'_{g_2}N_g} \\ |w\rangle & & -\frac{1}{N_g N'_{g_2}} & \frac{1}{N_g^2} \end{array}\right].$$

Now I can combine all of these into a single matrix and try to obtain some simpler constraints.

$$M \overset{\text{def}}{=} \left[\begin{array}{c|ccccc} & \langle v| & \langle v_1| & \langle v_2| & \langle w| & \langle w_1| \\ \hline |v\rangle & x & & & & \\ |v_1\rangle & & x-\frac{x_{g_1}p_{g_1}}{N_{8_1}^2}-\frac{x_{g_2}p_{g_2}}{N'^2_{g_1}} & -\frac{x_{g_1}p_{g_1}}{N_{g_1}N_{g_2}}+\frac{x_{g_2}p_{g_2}}{N'_{g_1}N_{g_2}} & -\frac{x_{g_1}p_{g_1}}{N_{g_1}N_g}+\frac{x_{g_2}p_{g_2}}{N'_{g_1}N_g} & \\ |v_2\rangle & & -\frac{x_{g_1}p_{g_1}}{N_{g_2}N_{g_1}}+\frac{x_{g_2}p_{g_2}}{N_{g_2}N'_{g_1}} & x-\frac{x_{g_1}p_{g_1}}{N_{g_2}^2}-\frac{x_{g_2}p_{g_2}}{N_{g_2}^2}-\frac{x_{g_3}p_{g_3}}{N'^2_{g_2}} & -\frac{x_{g_1}p_{g_1}}{N_{g_2}N_g}-\frac{x_{g_2}p_{g_2}}{N_{g_2}N_g}+\frac{x_{g_3}p_{g_3}}{N'_{g_2}N_g} & \\ |w\rangle & & -\frac{x_{g_1}p_{g_1}}{N_gN_{g_1}}+\frac{x_{g_2}p_{g_2}}{N_gN'_{g_1}} & -\frac{x_{g_1}p_{g_1}}{N_gN_{g_2}}-\frac{x_{g_2}p_{g_2}}{N_gN_{g_2}}+\frac{x_{g_3}p_{g_3}}{N_gN'_{g_2}} & \frac{p_{h_1}x_{h_1}+p_{h_2}x_{h_2}}{N_h^2}-\frac{1}{N_g^2}\sum_i x_{g_i}p_{g_i} & \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1}-x_{h_2}) \\ |w_1\rangle & & & & \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1}-x_{h_2}) & \frac{p_{h_2}x_{h_1}+p_{h_1}x_{h_2}}{N_h^2} \end{array}\right] \geq 0.$$

Despite this appearing to be a complicated expression, I can conclude that it will always be so that larger the x looser will be the constraint. To show this and to simplify this calculation, note that M can be split into a scalar condition, $x \geq 0$ (from the $|v\rangle\langle v|$ part) and a sub-matrix which I choose to write as

$$\begin{array}{c|c|c} & \langle v_1| \quad \langle v_2| & \langle w| \quad \langle w_1| \\ \hline \begin{matrix}|v_1\rangle \\ |v_2\rangle\end{matrix} & C & B^T \\ \hline \begin{matrix}|w\rangle \\ |w_1\rangle\end{matrix} & B & A \end{array} \geq 0.$$

Now since $\begin{bmatrix} C & B^T \\ B & A \end{bmatrix} \geq 0 \iff \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \geq 0 \iff C \geq 0,\ A - BC^{-1}B^T \geq 0,\ (\mathbb{I}-CC^{-1})B^T = 0$ using Shur's Complement condition for positivity where C^{-1} is supposed to be the generalised inverse. Since x is in our hands, we can take it to be sufficiently large so that $C > 0$ and thereby make sure that $\mathbb{I} - CC^{-1} = 0$. Evidently then, the only condition of interest is

$$A - BC^{-1}B^T \geq 0.$$

I can do even better than this actually. Note that if $C > 0$ then $C^{-1} > 0$ and that the second term is of the form

$$\underbrace{\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}}_{B}\underbrace{\begin{bmatrix} \alpha & \gamma \\ \gamma & \beta \end{bmatrix}}_{C^{-1}}\underbrace{\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}}_{B^T} = \begin{bmatrix} \begin{bmatrix} a & b \end{bmatrix}\begin{bmatrix} \alpha & \gamma \\ \gamma & \beta \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} & 0 \\ 0 & 0 \end{bmatrix} \geq 0$$

because $C^{-1} > 0$. I can therefore write my constraint equation as

$$A \geq BC^{-1}B^T \geq 0$$

and note that $A \geq 0$ is a necessary condition. This also becomes a sufficient condition in the limit that $x \to \infty$ because $C^{-1} \to 0$ in that case. We have thereby reduced the analysis to simply checking if

$$\begin{bmatrix} \frac{p_{h_1}x_{h_1}+p_{h_2}x_{h_2}}{N_h^2} - \frac{1}{N_g^2}\sum_i x_{g_i}p_{g_i} & \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1}-x_{h_2}) \\ \frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1}-x_{h_2}) & \frac{p_{h_2}x_{h_1}+p_{h_1}x_{h_2}}{N_h^2} \end{bmatrix} \geq 0.$$

This being a 2×2 matrix can be checked for positivity by the trace and determinant method. Another possibility is the use of Schur's Complement conditions again. Here, however, I intend to use a more general technique (similar to the one used in the split analysis). Let me introduce

$$\langle x_g \rangle \overset{\text{def}}{=} \frac{1}{N_g^2}\sum_i x_{g_i}p_{g_i},\ \left\langle \frac{1}{x_h} \right\rangle \overset{\text{def}}{=} \frac{1}{N_h^2}\sum_i \frac{p_{h_i}}{x_{h_i}}$$

and recall/note that term (I) and one element from term (III) constitute matrix A, which can also be written as

$$\begin{aligned} A &= x_{h_1}\left|h_{11}\right\rangle\left\langle h_{11}\right| + x_{h_2}\left|h_{22}\right\rangle\left\langle h_{22}\right| - \langle x_g\rangle\left|w\right\rangle\left\langle w\right| \\ &= \begin{array}{c|cc} & \langle h_{11}| & \langle h_{22}| \\ \hline |h_{11}\rangle & x_{h_1} & \\ |h_{22}\rangle & & x_{h_2} \end{array} - \langle x_g\rangle\left|w\right\rangle\left\langle w\right| \end{aligned}$$

Note that this now has the exact same form as that of the split constraint with $x_{g_1} \to \langle x_g \rangle$. I use the same $F - M \geq 0 \iff \mathbb{I} - \sqrt{F}^{-1} M \sqrt{F}^{-1} \geq 0$ for $F > 0$ technique to obtain $\mathbb{I} \geq \langle x_g \rangle \left| w'' \right\rangle \left\langle w'' \right|$ where $\left| w'' \right\rangle = \frac{\sqrt{\frac{p_{h_1}}{x_{h_1}}} |h_{11}\rangle + \sqrt{\frac{p_{h_2}}{x_{h_2}}} |h_{22}\rangle}{N_h}$. Normalising this one gets $|w'\rangle = \frac{|w''\rangle}{\sqrt{\left\langle \frac{1}{x_h} \right\rangle}}$ which entails $\mathbb{I} \geq \langle x_g \rangle \left\langle \frac{1}{x_h} \right\rangle |w'\rangle \langle w'|$ and that leads us to the final condition

$$\frac{1}{\langle x_g \rangle} \geq \left\langle \frac{1}{x_h} \right\rangle.$$

In fact all the techniques used in reaching this result can be extended to the $m \to n$ transition case as well and so the aforesaid result should hold in general.

Mochon's Assignments

Lemma (Mochon's Denominator)**.** $\sum_{i=1}^{n} \frac{1}{\prod_{j \neq i}(x_j - x_i)} = 0$ *for* $n \geq 2$.

Proof. I will prove this by induction (following Mochon's proof, just optimised for dummies instead of space). For $n = 2$

$$\frac{1}{(x_2 - x_1)} + \frac{1}{(x_1 - x_2)} = 0.$$

Now I show that if the result holds for $n - 1$ and it would also hold for n which would complete the inductive proof. I start with noting that

$$\frac{1}{(x_n - x_i)(x_1 - x_i)} = \frac{1}{x_n - x_1} \left[\frac{1}{x_1 - x_i} - \frac{1}{x_n - x_i} \right].$$

This is useful because it helps breaking the product into a sum. My strategy would be to pull off one common term so that I can apply the result to the remaining $n - 1$ terms. The expression of interest is

$$\sum_{i=1}^{n} \frac{1}{\prod_{j \neq i}(x_j - x_i)} = \frac{1}{\prod_{j \neq 1}(x_j - x_1)} + \sum_{i=2}^{n-1} \frac{1}{\prod_{j \neq i}(x_j - x_i)} + \frac{1}{\prod_{j \neq n}(x_j - x_n)}$$

where notice that the ith term in the sum (of the second term) can be written as

$$\frac{1}{(x_n - x_i)(x_1 - x_i) \prod_{j \neq i,1,n}(x_j - x_i)} = \frac{1}{x_n - x_1} \left[\frac{1}{\prod_{j \neq i,n}(x_j - x_i)} - \frac{1}{\prod_{j \neq 1,i}(x_j - x_i)} \right].$$

The first term can be written as

$$\frac{1}{(x_n - x_1) \prod_{j \neq 1,n}(x_j - x_1)}$$

while the last can be written as

$$\frac{-1}{(x_n - x_1) \prod_{j \neq n,1}(x_j - x_n)}.$$

Putting all these together, I get

$$\begin{aligned}
\sum_{i=1}^{n} \frac{1}{\prod_{j \neq i}(x_j - x_i)} &= \frac{1}{(x_n - x_1)} \left[\underbrace{\frac{1}{\prod_{j \neq 1,n}(x_j - x_1)} + \sum_{i=2}^{n-1} \frac{1}{\prod_{j \neq i,n}(x_j - x_i)}} - \underbrace{\sum_{i=2}^{n-1} \frac{1}{\prod_{j \neq 1,i}(x_j - x_i)} - \frac{1}{\prod_{j \neq 1,n}(x_j - x_n)}} \right] \\
&= \frac{1}{(x_n - x_1)} \left[\sum_{i=1}^{n-1} \frac{1}{\prod_{j \neq i,n}(x_j - x_i)} - \sum_{i=2}^{n} \frac{1}{\prod_{j \neq 1,i}(x_j - x_i)} \right]
\end{aligned}$$

where both sums disappear if the result holds for $n - 1$. This completes the proof. □

Lemma (Mochon's f-assignment Lemma)**.** $\sum_{i=1}^{n} \frac{f(x_i)}{\prod_{j \neq i}(x_j - x_i)} = 0$ *where* $f(x_i)$ *is of order* $k \leq n - 2$.

Proof. Again I do this by induction on k. For $k = 0$ the result holds by the previous result. I assume it holds for order $k - 1$ and show using this that it will also hold for order k (this proof is also Mochon's). Let $g(x_i)$ be a polynomial of order $k - 1$ s.t.

$$\sum_{i=1}^{n} \frac{f(x_i)}{\prod_{j \neq i}(x_j - x_i)} = \sum_{i=1}^{n} \frac{(x_1 - x_i)(x_2 - x_i) \ldots (x_k - x_i) - g(x_i)}{\prod_{j \neq i}(x_j - x_i)}.$$

Notice that the first part of the sum will disappear for all $1 \leq i \leq k$ because of the numerator. Consequently I can write the aforesaid as

$$\begin{aligned}
&= \sum_{i=k+1}^{n} \frac{(x_1 - x_i)(x_2 - x_i)\dots(x_k - x_i)}{\prod_{j\neq i}(x_j - x_i)} - \sum_{i=1}^{n} \frac{g(x_i)}{\prod_{j\neq i}(x_j - x_i)} \\
&= \sum_{i=k+1}^{n} \frac{1}{\prod_{j\neq i,1,2\dots,k}(x_j - x_i)} \\
&= 0
\end{aligned}$$

where in the first step, the second term becomes zero by assuming the result holds for $k-1$ and in the second step the sum disappears because of the previous result (Mochon's Denominator). Note that $k \leq n-2$ for the aforesaid argument to work because otherwise the last step would become invalid. $\square$

Lemma. $\sum_{i=1}^{n} \frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)} = (-1)^{n-1}$ *for* $n \geq 2$*.*

Proof. Let me call $d(n) = \sum_{i=1}^{n} \frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)}$ because I will do this inductively. I can then write

$$d(2) = \frac{x_1}{x_2 - x_1} + \frac{x_2}{x_1 - x_2} = \frac{x_1(x_1 - x_2) + x_2(x_2 - x_1)}{(x_2 - x_1)(x_1 - x_2)} = -1.$$

I assume the result holds for $d(n)$ and write

$$\begin{aligned}
d(n+1) &= \sum_{i=1}^{n+1} \frac{x_i^n}{\prod_{j\neq i}(x_j - x_i)} \\
&= \sum_{i=1}^{n+1} \frac{-(x_{n+1} - x_i)(x_i^{n-1}) + x_{n+1}x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)} \\
&= -\sum_{i=1}^{n+1}(x_{n+1} - x_i)\frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)} + x_{n+1} \underbrace{\sum_{i=1}^{n+1} \frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)}}_{=0 \text{ (Mochon's Denominator)}} \\
&= -\sum_{i=1}^{n} \cancel{\frac{(x_{n+1} - x_i)}{(x_{n+1} - x_i)}} \frac{x_i^{n-1}}{\prod_{j\neq i,n+1}(x_j - x_i)} + \cancelto{0}{(x_{n+1} - x_{n+1})} \frac{x_{n+1}^{n-1}}{\prod_{j\neq n+1}(x_j - x_{n+1})} \\
&= -d(n).
\end{aligned}$$

$\square$

Proposition. $\langle x_h \rangle - \langle x_g \rangle = \frac{1}{N_h^2} = \frac{1}{N_g^2}$ *for a Mochon's TDPG assignment with* $k = n-2$ *and coefficient of* x^{n-2} ± 1 *in* $f(x)$*. As above here* $\langle x_h \rangle = \frac{1}{N_h^2}\sum p_{h_i}x_{h_i}$ *and* $\langle x_g \rangle = \frac{1}{N_g^2}\sum p_{g_i}x_{g_i}$*.*

Proof. Note, to start with, that the coefficient of x^{n-2} being ± 1 is not an artificial requirement because for killing $n-2$ points $f(x)$ will have the form

$$f(x) = (x_{k_1} - x)(x_{k_2} - x)\dots(x_{k_{n-2}} - x) = (-1)^{n-2}x^{n-2} + \tilde{f}(x)$$

where $\tilde{f}$ is a polynomial of order $n-2$. Observe that

$$\begin{aligned}
N_h^2\left(\langle x_h \rangle - \langle x_g \rangle\right) &= \sum_{i=1}^{n} p(x_i)x_i = -\sum_{i=1}^{n} \frac{x_i f(x_i)}{\prod_{j\neq i}(x_j - x_i)} \\
&= -\sum_{i=1}^{n} \frac{x_i(-1)^{n-2}x_i^{n-2}}{\prod_{j\neq i}(x_j - x_i)} - \sum_{i} \frac{\tilde{f}(x_i)}{\prod_{j\neq i}(x_j - x_i)} \\
&= -(-1)^{n-2}\sum_{i=1}^{n} \frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)} \\
&= 1
\end{aligned}$$

where the second term in the second step vanishes because of Mochon's f-assignment Lemma and the last step follows from the previous result. $\square$