Quantum Weak Coin Flipping

Atul Singh Arora*, Jérémie Roland, Stephan Weis†

Université libre de Bruxelles

November 5, 2018
(v2, November 2019)

Abstract

We investigate weak coin flipping, a fundamental cryptographic primitive where two distrustful parties need to remotely establish a shared random bit. A cheating player can try to bias the output bit towards a preferred value. For weak coin flipping the players have known opposite preferred values. A weak coin-flipping protocol has a bias ϵ if neither player can force the outcome towards their preferred value with probability more than $\frac{1}{2}+\epsilon$. While it is known that all classical protocols have $\epsilon = \frac{1}{2}$, Mochon showed in 2007 [1] that quantumly weak coin flipping can be achieved with arbitrarily small bias (near perfect) but the best known explicit protocol has bias 1/6 (also due to Mochon, 2005 [2]). We propose a framework to construct new explicit protocols achieving biases below 1/6. In particular, we construct explicit unitaries for protocols with bias approaching 1/10. To go below, we introduce what we call the Elliptic Monotone Align (EMA) algorithm which, together with the framework, allows us to numerically construct protocols with arbitrarily small biases.

Contents

*responsible for all mistakes
†currently at the Universidade de Coimbra, Portugal

1 Introduction

We investigate coin flipping, a fundamental cryptographic primitive where two distrustful parties need to remotely generate a shared unbiased random bit. A cheating player can try to bias the output bit towards a preferred value. For weak coin flipping the players have known opposite preferred values. A weak coin flipping (WCF) protocol has a bias ϵ if neither player can force the outcome towards his/her preferred value with probability more than $\frac{1}{2}+\epsilon$. For strong coin-flipping there are no a priori preferred values and the bias is defined similarly. Restricting to classical resources, neither weak nor strong coin flipping is possible under information-theoretic security, as there always exists a player [3] who can force any outcome with probability 1. However, in a quantum world, strong coin flipping protocols with bias strictly less than $\frac{1}{2}$ have been shown and the best known explicit protocol has bias $\frac{1}{4}$ [4]. Nevertheless, Kitaev gave a lower bound of $\frac{1}{\sqrt{2}}-\frac{1}{2}$ for the bias of any quantum strong coin flipping, so an unbiased protocol is not possible.

As for weak coin flipping, the current best known explicit protocol—the Dip Dip Boom protocol—is due to Mochon [2] and has bias 1/6. In a breakthrough result, he even proved the existence of a quantum weak coin-flipping protocol with arbitrarily low bias $\epsilon > 0$, hence showing that near-perfect weak coin flipping is theoretically possible [1]. This fundamental result for quantum cryptography, unfortunately, was proved non-constructively, by elaborate successive reductions (80 pages) of the protocol to different versions of so-called point games, a formalism introduced by Kitaev [5] in order to study coin flipping. Consequently, the structure of the protocol whose existence is proved is lost. A systematic verification of this by independent researchers recently led to a simplified proof [6] (*only* 50 pages) but eleven years later, an explicit weak coin-flipping protocol is still unknown, despite various expert approaches ranging from the distillation of a protocol using the proof of existence to numerical search [7, 8]. Further, weak coin flipping provides, via black-box reductions, optimal protocols for strong coin flipping [9], bit commitment [10] and a variant of oblivious transfer [11] (fundamental cryptographic primitives). It is also used to implement other cryptographic tasks such as leader election [12] and dice rolling [13].

We construct a framework that allows us to convert simple point games (i.e. corresponding to known protocols) into explicit quantum protocols defined in terms of unitaries and projectors. We use the said framework to convert a bias 1/10 point game into its corresponding explicit protocol making it the first improvement of its kind in the last thirteen years since Mochon's Dip Dip Boom protocol (bias 1/6) [2].

Our second contribution, the Elliptic Monotone Align (EMA) algorithm, is a numerical algorithm which can provably find the unitaries required for implementing protocols with arbitrary biases, including the ones with $\epsilon \to 0$.

1.1 State of the Art | Kitaev's Formalisms and Mochon's Games

Let us start with noting two features of weak coin flipping. First, note that we can say, without loss of generality, that if the bit is zero it means Alice won and if the bit is one it means Bob won. Why is that? In weak coin flipping we know both players have known preferences. Alice wants zero and Bob wants one[1] since if they both wanted the same outcome bit, there would be no need to flip a coin. If a player gets what they want, we say they won. Second, we note that there are four situations which can arise in a weak coin flipping scenario of which three are of interest. Let us

[1] Alice wanting a zero and Bob wanting a one is just an uninteresting relabelling.

denote by HH the situation where both Alice and Bob are honest, i.e. follow the protocol. In this situation we want the protocol to be such that both Alice and Bob (a) win with equal probability and (b) are in agreement with each other. In the situation HC where Alice is honest and Bob is cheating, the protocol must protect Alice from a cheating Bob. In this situation, a cheating Bob tries to convince an honest Alice that he has won. His probability of succeeding by using his best cheating strategy is denoted by P_B^* where the star/asterisk refers to a cheating player and the subscript denotes the outcome he desires to enforce on the honest player. The CH situation where Bob is honest and Alice is cheating naturally points us to the corresponding definition of P_A^*. The situation CC where both players are cheating is not of interest to us as nothing can be said which depends on the protocol. This is because nobody is following the protocol.

A trivial example of a weak coin flipping protocol is where Alice flips a coin and reveals the outcome to Bob over the phone. A cheating Alice can simply lie and always win against an honest Bob which means $P_A^* = 1$. On the other hand, a cheating Bob can not do anything to convince Alice that he has won, unless it happens by random chance on the coin flip. This corresponds to $P_B^* = \frac{1}{2}$. The bias of the protocol is $\max[P_A^*, P_B^*] - \frac{1}{2}$ which for this naïve protocol amounts to $\frac{1}{2}$, the worst possible. Manifestly, constructing protocols where one player is protected is nearly trivial. Constructing protocols where neither player is able to cheat (against an honest player) is the real challenge.

Given a WCF protocol it is not a priori clear how the best success probability of a cheating player, denoted by $P_{A/B}^*$, should be computed as the strategy space can be dauntingly large. It turns out that all quantum WCF protocols can be defined using the exchange of a message register interleaved with the players applying the unitaries U_i locally (see Figure 3) until a final measurement, say Π_A denoting Alice won and Π_B denoting Bob won, is made in the end. Computing P_A^* in this case reduces to a semi-definite program (SDP) in ρ: maximise $P_A^* = \mathrm{tr}(\Pi_A \rho)$ given the constraint that the honest player follows the protocol. Similarly for computing P_B^* one can define another SDP. Using SDP duality one can turn this maximization problem over cheating strategies into a minimization problem over dual variables $Z_{A/B}$. Any dual feasible assignment then provides an upper bound on the cheating probabilities $P_{A/B}^*$. SDPs are usually easy to handle but in this case, there are two SDPs, and we must optimise both simultaneously (see Subsection 2.1). Note that here we assume the protocol is known and we are trying to find bounds on P_A^* and P_B^*. However, our goal is to find good protocols. So what we would like is a framework which allows us to do both, construct protocols and find the associated P_A^* and P_B^*. Kitaev gave us such a framework.

He converted this problem about matrices (Zs, ρs and Us) into a problem about points on a plane, which Mochon called Kitaev's Time Dependent Point Game (TDPG) framework. In this framework, one is concerned with a sequence of frames—the positive quadrant of the plane with some points and their probability weights—which must start with a fixed frame and end with a frame that has only one point. The fixed starting frame consists of two points at $[0, 1]$ and $[1, 0]$ with weight $1/2$. The end frame must be a single point, say at $[\beta, \alpha]$, with weight 1. The objective of the protocol designer is to get this end point as close to the origin as possible by transitioning through intermediate frames (see Figure 1) by following certain rules (we describe these shortly). The magic of this formalism, roughly stated, is that if one abides by these rules then corresponding to every such valid sequence of frames, there exists a WCF protocol with $P_A^* = \alpha$, $P_B^* = \beta$ (see Subsection 2.3).

We now describe these rules. Consider a given frame and focus on a set of points that fall on this vertical (or horizontal) line. Let the y coordinate (or x coordinate) of the ith point be given by z_{g_i} and the weight be given by p_{g_i}. Let z_{h_i} and p_{h_i} denote the corresponding quantity in the subsequent frame. Then, the following conditions must hold

1. the probabilities are conserved, viz. $\sum_i p_{g_i} = \sum_i p_{h_i}$

2. for all $\lambda > 0$

$$\sum_i \frac{\lambda z_{g_i}}{\lambda + z_{g_i}} p_{g_i} \leq \sum_i \frac{\lambda z_{h_i}}{\lambda + z_{h_i}} p_{h_i}. \tag{1}$$

Note that from one frame to the next, one can either make a horizontal transition or a vertical transition. By combining these sequentially one can obtain the desired form of the final frame, i.e. a single point. The aforesaid rule and the points in the frames arise from the dual variables $Z_{A/B}$. Just as the state ρ evolves through the protocol, so do the dual variables $Z_{A/B}$. The points and their weights in the TDPG are exactly the eigenvalue pairs of $Z_{A/B}$ with the probability weight assigned to them by the honest state $|\psi\rangle$ at a given point in the protocol ($|\psi\rangle$ and ρ are closely related). The aforementioned rules are related to the dual constraints. Given an explicit WCF protocol and a feasible assignment for the dual variables witnessing a given bias, it is straightforward to construct the TDPG. However, going backwards, constructing the WCF dual from a TDPG is highly non-trivial and no general construction is known.

Our main contribution is precisely to this part. We construct a framework which allows for a ready conversion of simple TDPGs into explicit protocols, and once supplemented with the EMA algorithm, it can convert any TDPG into its corresponding protocol. This is relevant because Mochon's breakthrough result was to define a family of games[2] with bias $\epsilon = \frac{1}{4k+2}$ where k encodes the number of points that are involved in the non-trivial step (for $k = 1$ it reduces to a version of the Dip Dip Boom (bias 1/6) protocol) which means, effectively, we can numerically construct quantum weak coin flipping protocols with arbitrarily small bias (see Section 5 of either [1, 6]).

As this point game formalism is the cornerstone of the analysis, we simplify the rules further and then apply them to construct a simple example game. Later, we convert this example game into an explicit protocol using our framework. If we restrict ourselves to transitions involving only one initial and one final point, the second condition reduces to $z_g \leq z_h$ (we suppressed the subscript). This is called a *raise*. It means that we can always increase the coordinate of a single point. What about going from one initial point to many final points (note that the points before and after must lie along either a horizontal or a vertical line)? The second condition in this case becomes $1/z_g \geq \langle 1/z_h \rangle$, that is the harmonic mean of the final points must be greater than or equal to that of the initial point, where $\langle f(z_h) \rangle := \left(\sum_i f(z_{h_i}) p_{h_i}\right) / \left(\sum_j p_{h_j}\right)$. This is called a *split*. Finally, we can ask: What happens upon merging many points into a single point? The second condition becomes $\langle z_g \rangle \leq z_h$, that is the final position must not be smaller than the average initial position (where $\langle f(z_g) \rangle$ is analogously defined). This is called a *merge*. While these three transitions/moves do not exhaust the set of moves, they are enough to construct games that almost achieve the bias 1/6. Let us construct a simple game as an example. We start with the initial frame and raise the point $[1, 0]$ along the vertical to $[1, 1]$ (see Figure 1). We know this move is allowed as it is just a raise. Next we merge the points $[0, 1]$ with $[1, 1]$ using a horizontal merge. The x-coordinate of the resulting point can at best be $\frac{1}{2}.0 + \frac{1}{2}.1 = \frac{1}{2}$ where we used the fact that both points have weight 1/2. Thus we end up with a single point at $[\frac{1}{2}, 1]$ with all the weight. Kitaev's framework tells us that there must exist a protocol which yields $P_A^* = 1$ while $P_B^* = \frac{1}{2}$. This, however, is the phone protocol that we started our discussion with! It is a neat consistency check but it yields a trivial bias. This is because we did not use the split. If we use a split once, we can, by simply matching the weights, already obtain a game with $P_A^* = P_B^* = \frac{1}{\sqrt{2}}$. Protocols corresponding to this bias were found by various researchers [14, 15, 16] long before this framework was known. In fact, the bias

[2]Mochon describes his games in Kitaev's Time Independent Point Game (TIPG) framework but it is straightforward to go back from a TIPG to a TDPG.

of the said weak coin flipping protocol, $\epsilon = \frac{1}{\sqrt{2}} - \frac{1}{2}$, was exactly the lower bound for *strong* coin flipping. It was an exciting time (we imagine) as the technique used to bound strong coin flipping fails for weak coin flipping. The matter was not resolved for a while. This protocol held the record for being the best known weak coin flipping protocol until Mochon progressively showed that if we use multiple splits wisely at the beginning followed by a raise, one simply needs to use merges thereafter to obtain a game with bias almost 1/6, which corresponds to his Dip Dip Boom protocol. The Dip Dip Boom protocol, is actually a family of protocols which in the limit of infinite rounds of communication yields bias 1/6. Going lower, therefore, is not a straight forward extension and we need to use moves which can not be decomposed into the three basic ones, splits, merges and raises. Our contribution is to find ways of constructing the unitaries corresponding to these moves.

1.2 A Framework | First Contribution

We first describe our framework for converting a TDPG into an explicit protocol. We start by defining a 'canonical form' for any given frame of a TDPG. This allows one to write the WCF dual variables, Zs, and the honest state $|\psi\rangle$ associated with each frame of the TDPG. We define a sequence of quantum operations, unitaries and projections, which allow Alice and Bob to transition from the initial frame to the final frame. It turns out that there is only one non-trivial quantum operation in the sequence which we leave partially specified for the moment. This means that we know that the unitary should send the honest initial state to the honest final state. However the action of the unitary on the orthogonal space, which intuitively is what would bestow on it the cheating prevention/detection capability, is obtained as an interesting constraint. Using the SDP formalism we write the constraints at each step of the sequence on the Zs and show that they are indeed satisfied (see Theorem 46 for a full statement of the following, Subsection 3.2 for its proof and the description of the framework).

Theorem 1 (TEF constraint (simplified))**.** *If a unitary matrix U acting on the space span$\{|g_1\rangle, |g_2\rangle \dots, |h_1\rangle, |h_2\rangle \dots\}$ satisfying the constraints*

$$U|v\rangle = |w\rangle,$$

$$\sum_i x_{h_i} |h_i\rangle\langle h_i| - \sum_i x_{g_i} E_h U |g_i\rangle\langle g_i| U^\dagger E_h \geq 0 \tag{2}$$

can be found for every move/transition (see Definition 47 and Definition 9) of a TDPG then an explicit protocol with the corresponding bias can be obtained using the TDPG–to–Explicit–protocol Framework (TEF), where $\{|g_i\rangle\}, \{|h_i\rangle\}$ are orthonormal vectors and if the transition is horizontal

- *the initial points have x_{g_i} as their x-coordinate and p_{g_i} as their corresponding probability weight,*
- *the final points have, similarly, x_{h_i} as their x-coordinate and p_{h_i} as their corresponding probability weight*
- *E_h is a projection onto the span$\{|h_i\rangle\}$ space,*
- *$|v\rangle = \sum_i \sqrt{p_{g_i}} |g_i\rangle / \sqrt{\sum p_{g_i}}$, $|w\rangle = \sum_i \sqrt{p_{h_i}} |h_i\rangle / \sqrt{\sum p_{h_i}}$*

and if the transition is vertical, the x_{g_i} and x_{h_i} become the y-coordinates y_{g_i} and y_{h_i} with everything else unchanged.

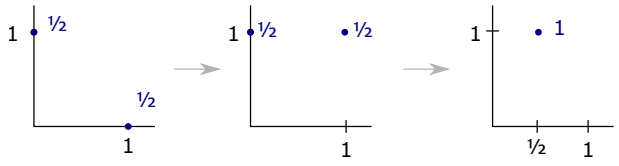


Figure 1: Point game corresponding to the weak coin flipping over the phone protocol.

Note that the TDPG already specifies the coordinates x_{h_i}, x_{g_i} and the probabilities p_{h_i}, p_{g_i} which satisfy, Equation (1), the scalar condition. Our task therefore reduces to finding the correct U which satisfies the aforesaid matrix constraints. It is this general problem that is solved by our EMA algorithm which we describe later.

Given such a unitary U acting on the space span$\{|g_1\rangle, |g_2\rangle \ldots |h_1\rangle, |h_2\rangle \ldots\}$ one can construct a unitary, $U^{(2)}_{AM}$, acting non-trivially on the space span$\{|g_1 g_1\rangle_{AM}, |g_2 g_2\rangle_{AM} \ldots, |h_1 h_1\rangle_{AM}, |h_2 h_2\rangle_{AM}, \ldots\}$ by mapping $|g_i\rangle \to |g_i g_i\rangle_{AM}$, $|h_i\rangle \to |h_i h_i\rangle_{AM}$ and as identity otherwise. We now informally describe how to convert a TDPG into an explicit protocol. It suffices to show what a transition from a given frame to the next frame corresponds to in terms of the protocol. In this discussion, we refer to them as the initial frame and the final frame. Assume that the corresponding non-trivial $U^{(2)}_{AM}$ is known. As we saw, a given transition would either be horizontal or vertical. We assume it is horizontal without loss of generality[3]. We label the points that do not participate in this horizontal transition, i.e. remain unchanged in both frames, by $k_1, k_2 \ldots$ in both frames. The points in the initial frame involved in this transition are labelled $g_1, g_2 \ldots$ and the ones in the final frame are labelled $h_1, h_2 \ldots$. All the points are now labelled. We denote the coordinates of the final points by $x_{h_1}, x_{h_2} \ldots$ and the probability weights by $p_{h_1}, p_{h_2} \ldots$. We similarly define x_{g_i}, p_{g_i} and x_{k_i}, p_{k_i}. The Hilbert space of interest is given by $\mathcal{H} := \text{span}\{|k_1\rangle, |k_2\rangle \ldots, |g_1\rangle, |g_2\rangle \ldots, |h_1\rangle, |h_2\rangle \ldots, |m\rangle\}$ where each vector is assumed orthonormal ($|m\rangle$ is just an idle state in which the message register is assumed to be initially and returned to finally). We assume that Alice's register, Bob's register and the message register each have dimension at least as large as $\dim(\mathcal{H})$. The state (by state in this discussion, we mean the honest state) corresponding to the initial frame is assumed to have the form

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

Bob: Assume Bob has the message register. He applies the conditional swap $U^{\text{SWP}\{\vec{g},m\}}_{BM}$ where $U^{\text{SWP}\{\vec{g},m\}}_{BM}$ swaps conditionally on both registers being in the subspace span$\{|g_1\rangle, |g_2\rangle \ldots, |m\rangle\}$. The state after this operation is

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

He then sends the message register to Alice.

Alice: Alice applies the non-trivial unitary $U^{(2)}_{AM}$ on her local register and the message register. She then measures $\{E^{(2)}, \mathbb{I} - E^{(2)}\}$ where $E^{(2)} := (\sum |h_i\rangle \langle h_i| + \sum |k_i\rangle \langle k_i|)_A \otimes \mathbb{I}_M$. The state at

[3]Mochon's point games have a repeating structure he calls a "ladder". Corresponding to each k he constructs a family of point games parametrised by the number of points in this ladder. The game approaches the bias $\epsilon = (4k+2)^{-1}$ as the number of points is increased (the value is reached in the limit of infinite points). Consequently, we consider a finite set of points in the transition.

this point is

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

If the outcome corresponds to the latter, she declares herself to be the winner. Otherwise she sends the message register back to Bob.

Bob: Bob again applies a conditional swap $U_{BM}^{\mathrm{SWP}\{\vec{h},m\}}$ followed by a measurement corresponding to $\{E^{(3)}, \mathbb{I} - E^{(3)}\}$ where $E^{(3)} := (\sum_i |h_i\rangle \langle h_i| + \sum_i |k_i\rangle \langle k_i|)_B \otimes \mathbb{I}_M$. The final state is

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

If the outcome corresponds to $\mathbb{I} - E^{(3)}$, Bob declares himself the winner.

As the final state is in the same form as the initial state, one can progressively build the sequence corresponding to the complete protocol. Once the entire sequence is known, one must reverse the order of all the operations to obtain the final protocol. Note that the message register is initially decoupled, it then gets entangled, and finally it emerges decoupled again. This simplifies the analysis (and also entails that one need not keep the message register coherent for the duration of the protocol; keeping it coherent for each round individually is sufficient).

Let us try to apply this procedure to our example game (see Figure 1). We label the points in the first frame as g_1 and g_2. The state is given by $\frac{1}{\sqrt{2}} (\left|g_1 g_1\right\rangle_{AB} + \left|g_2 g_2\right\rangle_{AB}) \otimes \left|m\right\rangle_M$. (This should make it clear that the order is reversed here because we want to end with an EPR like state so that when Alice and Bob make a measurement, they agree on a random bit.) We simply claim for the moment that raising does not require Alice and Bob to do anything. This means that we can consider the second frame with the same labels. We now apply the merge transition by using the aforesaid recipe, where Bob applies a swap, sends the message register to Alice, she applies $U_{AM}^{(2)}$ and the projector, returns the message register to Bob and he applies the final swap and measurement. We continue to assume we are given the correct $U_{AM}^{(2)}$ that implements the merge step. The state one obtains after the application of these unitaries turns out to be $\left|h_1 h_2\right\rangle_{AB} \otimes \left|m\right\rangle_M$. (This looks like the state we should start with, completely unentangled. This is intuitively why the actual protocol is a reversed version of what we have.) Our procedure can be applied to any point game, granted the non-trivial unitary $U^{(2)}$ can be found. The central issue is that there is no general recipe known for constructing $U^{(2)}$s.

To address this we can prove that what we call the Blinkered Unitary satisfies the required constraints for both the split and merge moves (see Subsection 3.3). It is defined as

$$U_{\mathrm{blink}} = \left|w\right\rangle \langle v| + \left|v\right\rangle \langle w| + \sum_i \left|v_i\right\rangle \langle v_i| + \sum_i \left|w_i\right\rangle \langle w_i| \tag{3}$$

where $|v\rangle$, $\{|v_i\rangle\}$ and $|w\rangle$, $\{|w_i\rangle\}$ are orthonormal vectors spanning the $\{|g_i\rangle\}$ and $\{|h_i\rangle\}$ space respectively. With these the former best protocol (bias 1/6) can already be derived from its TDPG, in a manner analogous to the one used for the example game. This was not known (to the best of our knowledge), even though the protocol itself was separately known and analysed. We next study the family of bias 1/10 TDPGs and isolate the precise moves required to implement it (see Subsection **??**). Let $n_g \to n_h$ denote a move from n_g initial points to n_h final points. While the bias 1/6 games used a $2 \to 1$ merge as its key move, the bias 1/10 games use a combination of $3 \to 2$ and $2 \to 2$ moves (these can not be produced by a combination of merges and splits, as was pointed out earlier). We give analytic expressions for these unitaries and show that they satisfy

the required constraints (see Subsection 4.4). In particular, we show that for $3 \to 2$ moves with $x_{g_1} < x_{g_2} < x_{g_3}$ and $x_{h_1} < x_{h_2}$

$$U_{3\to 2} = |w\rangle\langle v| + |w_1\rangle\langle v_1'| + |v_2'\rangle\langle v_2'| + |v_1'\rangle\langle w_1| + |v\rangle\langle w| \tag{4}$$

satisfies the required constraints (under some further technical conditions which are satisfied by the 1/10 games of interest), where

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

and

$$|w\rangle = \frac{\sqrt{p_{h_1}}|h_1\rangle + \sqrt{p_{h_2}}|h_2\rangle}{N_h}, |w_1\rangle = \frac{\sqrt{p_{h_2}}|h_1\rangle - \sqrt{p_{h_1}}|h_2\rangle}{N_h}$$

are normalised vectors (this fixes the normalisation factors) which we use to define

$$|v_1'\rangle = \cos\theta|v_1\rangle + \sin\theta|v_2\rangle, |v_2'\rangle = \sin\theta|v_1\rangle - \cos\theta|v_2\rangle$$

where $\cos\theta$ is obtained by solving

$$\begin{aligned}
\frac{\sqrt{p_{h_1}p_{h_2}}}{N_h^2}(x_{h_1} - x_{h_2}) - \cos\theta\frac{\sqrt{p_{g_2}p_{g_3}}}{N_g N_{v_1}}(x_{g_2} - x_{g_3}) \\
- \sin\theta\langle x_g\rangle\frac{N_g}{N_{v_2}} = 0
\end{aligned}$$

and choosing the solution which is closer to 1. Similarly we give an explicit unitary corresponding to the second move, i.e. the $2 \to 2$ move. For the second move, i.e. the $2 \to 2$ move with $x_{g_1} < x_{g_2}$ and $x_{h_1} < x_{h_2}$, we show that

$$U_{2\to 2} = |w\rangle\langle v| + (\alpha|v\rangle + \beta|w_1\rangle)\langle v_1| + |v\rangle\langle w| + (\beta|v\rangle - \alpha|w_1\rangle)\langle w_1|$$

satisfies the required constraints (again, under further technical conditions which are satisfied by the 1/10 games of interest) where

$$\begin{aligned}
|v\rangle &= \frac{1}{N_g}\left(\sqrt{p_{g_1}}|g_1\rangle + \sqrt{p_{g_2}}|g_2\rangle\right), \\
|v_1\rangle &= \frac{1}{N_g}\left(\sqrt{p_{g_2}}|g_1\rangle - \sqrt{p_{g_1}}|g_2\rangle\right)
\end{aligned}$$

and

$$\begin{aligned}
|w\rangle &= \frac{1}{N_h}\left(\sqrt{p_{h_1}}|h_1\rangle + \sqrt{p_{h_2}}|h_2\rangle\right) \\
|w_1\rangle &= \frac{1}{N_h}\left(\sqrt{p_{h_2}}|h_1\rangle - \sqrt{p_{h_1}}|h_2\rangle\right).
\end{aligned}$$

Further, $\alpha, \beta \in \mathbb{R}$ are such that $\alpha^2 + \beta^2 = 1$ and

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

This lets us, in effect, convert Mochon's family of bias 1/10 games into explicit protocols, finally breaking the 1/6 barrier. Mochon's games achieving lower biases correspond to larger unitary matrices. Consequently, this approach based on guessing the correct form of the solution becomes untenable.

1.3 EMA Algorithm | Second Contribution

To go lower than 1/10 we use our Elliptic Monotone Align (EMA) algorithm which we now describe. Note that if we neglect the projector in Equation (2), we can express it as $X_h \geq U X_g U^\dagger$ where X_h, X_g are diagonal matrices with positive entries (justified in Subsection 5.1). Surprisingly, it is possible to show that we can restrict ourselves to orthogonal matrices without loss of generality (see Subsection 5.2). Once we restrict to real numbers, it is easy to see that the set of vectors $\mathcal{E}_{X_h} := \{|u\rangle \,|\, \langle u| X_h |u\rangle = 1\}$ describe the boundary of an ellipsoid as $\sum_i u_i^2/(x_{h_i}^{-1}) = 1$ (note x_{h_i} is fixed here and u_i is the variable). Similarly $\mathcal{E}_{OX_gO^T}$ represents a rotated ellipsoid where O is orthogonal (see Figure 2). Note that larger the x_{h_i} (or x_{g_i}) higher is the curvature of the ellipsoid along the associated direction. It is not hard to see the aforesaid inequality, geometrically, as the $\mathcal{E}_{X_h}$ ellipsoid being contained inside the $\mathcal{E}_{OX_gO^T}$ ellipsoid (the order gets reversed; see Subsection 6.1).

Recall that the orthogonal matrix also has the property $O|v\rangle = |w\rangle$. Imagine that in addition, we have $\langle w| X_h |w\rangle = \langle v| X_g |v\rangle$ which in terms of the point game means that the average is preserved (as was the case for merge). In terms of the ellipsoids, it means that the ellipsoids touch along the $|w\rangle$ direction. More precisely, the point $|c\rangle := |w\rangle / \sqrt{\langle w| X_h |w\rangle}$ belongs to both $\mathcal{E}_{X_h}$ and $\mathcal{E}_{OX_gO^T}$. Since the inequality tells us the smaller h ellipsoid is contained inside the larger g ellipsoid, and we now know that they touch at the point $|c\rangle$, we conclude that their normals evaluated at $|c\rangle$ must be equal. Further, we can conclude that the inner ellipsoid must be more curved than the outer ellipsoid.

Mark the point $|c\rangle$ on the $\mathcal{E}_{OX_gO^T}$ ellipsoid. Now imagine rotating the $\mathcal{E}_{X_g}$ ellipsoid to the $\mathcal{E}_{OX_gO^T}$ ellipsoid. The normal at the marked point must be mapped to the normal of $\mathcal{E}_{X_h}$ at $|c\rangle$. It turns out that to evaluate the normals $|n_h\rangle$ on $\mathcal{E}_{X_h}$ at $|c\rangle$ and $|n_g\rangle$ on $\mathcal{E}_{X_g}$ at the marked point, one only needs to know $X_h, X_g, |v\rangle$ and $|w\rangle$. Complete knowledge of O is not required and yet we can be sure that $O|n_g\rangle = |n_h\rangle$ which means O must have a term $|n_h\rangle\langle n_g|$. In fact, one can even evaluate the curvature from the aforesaid quantities. It so turns out that when this condition is expressed precisely, it becomes an instance of the same problem we started with one less dimension allowing us to iteratively find O, which so far we had only assumed to exist. This, however, only works under our assumption that $\langle w| X_h |w\rangle = \langle v| X_g |v\rangle$. This is not always the case which we address next.

A monotone function f is defined to be a function which has the property "$x \geq y \implies f(x) \geq f(y)$". An operator monotone function[4] is obtained from a generalisation of the aforesaid property to matrices, which in our notation can be expressed as "$X_h \geq OX_gO^T \implies f(X_h) \geq Of(X_g)O^T$". In mathematics, it is known that for a certain class of operator monotone functions f,

[4]Note that the monotone function $f(x) = x^2$ is not an operator monotone. This is a counter-example: $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \geq \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

f^{-1} is also an operator monotone. Using these results in conjunction with results from Aharonov et al. [6] we conclude that one can show that there is always an operator monotone f such that $\langle w| f(X_h) |w\rangle = \langle v| f(X_g) |v\rangle$. (This result also admits a beautiful geometric interpretation. It means that to establish $\mathcal{E}_{X_h}$ is inside $\mathcal{E}_{OX_gO^T}$, which essentially means we look at all different directions and make sure the h ellipsoid is inside the g ellipsoid, we can instead look along a single direction $|w\rangle$ and check that all the different ellipsoids $\mathcal{E}_{f(X_h)}$ are inside the corresponding $\mathcal{E}_{Of(X_g)O^T}$ ellipsoids along just this direction, for every operator monotone f in the class indicated earlier.) Since the orthogonal matrix which solves the initial problem also solves the one mapped by f, we can use our technique on the latter to proceed. This shows how and why our EMA algorithm works. Let us summarise the algorithm into an informal statement.

Definition (EMA Algorithm (informal))**.** Given a transition from a TDPG the algorithm proceeds in three phases.

1. Initialise
   - Tightening procedure: Bring the final points close to zero until the corresponding ellipsoids start to touch.
   - Spectral domain, matrices: Find the spectrum of the matrices which represent the ellipsoid. Evaluate the smallest matrix size n needed to represent the problem using ellipsoids.
   - Bootstrapping: Using the aforesaid, define $\left(X_h^{(n)}, X_g^{(n)}, \left|w^{(n)}\right\rangle, \left|v^{(n)}\right\rangle\right) := \underline{\mathrm{X}}^{(n)}$ where the superscript denotes the size of the matrix and vectors.
2. Iterate (neglecting special cases)

   Input: $\underline{\mathrm{X}}^{(k)}$

   Output: $\underline{\mathrm{X}}^{(k-1)}$, the vector $\left|u_h^{(k)}\right\rangle$ and the orthogonal matrices $\bar{O}_g^{(k)}, \bar{O}_h^{(k)}$

   Procedure:
   - Tightening procedure: Similar to the one above, shrink the outer ellipsoid until it touches the inner ellipsoid.
   - Honest align: Use operator monotone functions to make the ellipsoids touch along the $|w\rangle$ direction.
   - Evaluate the Reverse Weingarten Map: Evaluate the curvatures and the normal (which fixes $\left|u_h^{(k)}\right\rangle$) along the $|w\rangle$ direction.
   - Finite Method: Use the curvatures to specify $\underline{\mathrm{X}}^{(k-1)}$ and find the orthogonal matrices $\bar{O}_g^{(k)}, \bar{O}_h^{(k)}$.
3. Reconstruction

   Evaluate $O^{(n)}$ recursively using $O^{(k)} = \bar{O}_g^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}_h^{(k)}$.

Theorem 2 (Correctness of the EMA Algorithm (informal))**.** *Given a transition of a TDPG, the EMA Algorithm always finds a U such that the constraints in Theorem 1 are satisfied.*

See Subsection 1.3 for the complete algorithm and proof of the theorem; in particular Definition 96 and Theorem 97 for the corresponding formal statements. Results from a preliminary numerical implementation of the EMA algorithm are discussed in Section 8.

Despite the apparent simplicity of the main argument there were many difficulties we had to address in order to prove the aforesaid statement. We had to extend the results about operator monotone functions to be able to use them for performing the tightening step as indicated and for being certain that the solution unitary/orthogonal matrix stays unchanged under these transformations. We also extended some results related to different representations of the aforesaid transitions as these situations arise in the tightening procedure (see Subsection 7.2.1). Finding an easy method for evaluating the curvatures—the reverse Weingarten map—was key as has been noted (see Subsection 6.2). The trickiest part of the algorithm, which we have not mentioned here in the introduction, was handling the cases where one of the tangent directions of an ellipsoid has an infinite curvature. For concreteness, imagine an ellipse which under an operator monotone gets mapped to a line segment. The tip of the line segment, if viewed as a limit of an ellipse, has an infinite curvature. In these cases, our finite analysis breaks down as the normal is no longer well defined. For the moves used by Mochon in his 1/18 game for instance that we tried to numerically solve using this algorithm, this infinite case does not appear. However, to solve the split move using the algorithm (instead of the blinkered unitaries) the infinite case does appear. In either case, our algorithm can handle these infinite cases using what we call the Wiggle-v method (see Subsection 7.2.3 and Subsection 7.3).

The implication is that we can now numerically convert known games with arbitrarily small bias into complete protocols. One remaining question is the effect of noise. In the current analysis two idealising assumptions have been made. First, the EMA algorithm assumes one can solve certain problems with arbitrary precision classically, such as finding the roots of polynomials and diagonalising matrices. Second, in Mochon/Kitaev's point game formalisms, one assumes that the unitaries are known and applied exactly. Neither of these will hold practically, therefore, the effect of noise on the bias of the protocol must be quantified, which we leave as an open problem for further work.

The remaining document starts with stating the previously known formal results and motivating their proofs (Section 2). Thereafter the document is divided into two parts. The first part establishes the TDPG-to-Explicit-protocol Framework, TEF, (Section 3) and then discusses its application to Mochon's games (Section 4) including the one with bias 1/10. The second part starts with laying the groundwork for viewing the problem in terms of ellipsoids (Section 5), summarises some results about ellipsoids (Section 6) and finally discusses the Elliptic Monotone Algorithm itself (Section 7). It ends by stating some observations made through a preliminary numerical implementation of the said algorithm and briefly discussing related open problems (Section 8).

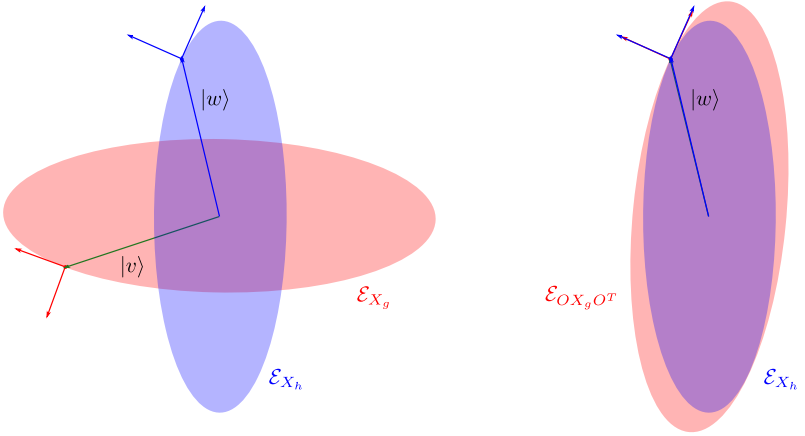


Figure 2: On the left the ellipsoids correspond to the diagonal matrices X_g and X_h. The vectors $|w\rangle$ and $|v\rangle$ indicate only the direction. On the right, the larger ellipsoid is now rotated to corresponding to OX_gO^T. The point of contact is along the vector $|w\rangle = O|v\rangle$.

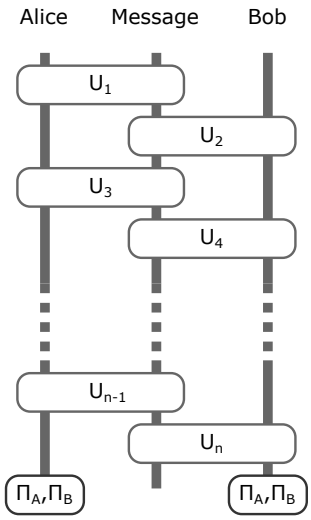


Figure 3: General structure of a Weak Coin Flipping protocol.

2 Prior Art

We start with stating the results up to a certain point from Aharonov et al's [6] paper (which completely formalises Mochon's results [1] and simplifies one of the results proved in its appendix; we build on this improved proof in our work). We will motivate the statements as we go along. It is unlikely that the next section will make perfect sense unless one has already read Aharonov's et al's and/or Mochon's article. The ideas should get clear in the following sections as we use them to construct explicit protocols. Logically, however, we have tried to keep everything consistent (even though the presentation thereof may not be optimal).

We define $\mathbb{R}_{\geq} := [0, \infty)$, $\mathbb{R}_{>} := (0, \infty)$ and similarly $\mathbb{R}_{\leq} := (-\infty, 0], \mathbb{R}_{<} := (-\infty, 0)$. A note about the colour scheme. We use purple for intuitive and non-technical discussions and, in this section, blue for statements/corrections that we have added to the results from Aharonov et al's article.

2.1 WCF protocol as an SDP and its Dual

Any weak coin flipping protocol can be expressed in the following general form (we do not prove this claim here; see [4]).

Definition 3 (WCF protocol with bias ϵ)**.** For n even, an n-message WCF protocol between two players, Alice and Bob, is described by

- three Hilbert spaces with $\mathcal{A}$, $\mathcal{B}$ corresponding to Alice and Bob's private workspaces (Bob does not have any access to $\mathcal{A}$ and Alice to $\mathcal{B}$), and a message space $\mathcal{M}$;
- an initial product state $|\psi_0\rangle = |\psi_{A,0}\rangle \otimes |\psi_{M,0}\rangle \otimes |\psi_{B,0}\rangle \in \mathcal{A} \otimes \mathcal{M} \otimes \mathcal{B}$;
- a set of n unitaries $\{U_1, \dots U_n\}$ acting on $\mathcal{A} \otimes \mathcal{M} \otimes \mathcal{B}$ with $U_i = U_{A,i} \otimes \mathbb{I}_{\mathcal{B}}$ for i odd and $U_i = \mathbb{I}_{\mathcal{A}} \otimes U_{B,i}$ for i even;
- a set of honest states $\{|\psi_i\rangle : i \in [n]\}$ defined by $|\psi_i\rangle = U_i U_{i-1} \dots U_1 |\psi_0\rangle$;
- a set of n projectors $\{E_1, \dots E_n\}$ acting on $\mathcal{A} \otimes \mathcal{M} \otimes \mathcal{B}$ with $E_i = E_{A,i} \otimes \mathbb{I}_{\mathcal{B}}$ for i odd, and $E_i = \mathbb{I}_{\mathcal{A}} \otimes E_{B,i}$ for i even, such that $E_i |\psi_i\rangle = |\psi_i\rangle$;
- two final positive operator valued measure (POVM) $\{\Pi_A^{(0)}, \Pi_A^{(1)}\}$ acting on $\mathcal{A}$ and $\{\Pi_B^{(0)}, \Pi_B^{(1)}\}$ acting on $\mathcal{B}$.

The WCF protocol proceeds as follows:

- In the beginning, Alice holds $|\psi_{A,0}\rangle |\psi_{M,0}\rangle$ and Bob $|\psi_{B,0}\rangle$.
- For $i = 1$ to n:
  - If i is odd, Alice applies U_i and measures the resulting state with the POVM $\{E_i, \mathbb{I} - E_i\}$. On the first outcome, Alice sends the message qubits to Bob; on the second outcome, she ends the protocol by outputting "0", i.e, Alice declares herself to be the winner.
  - If i is even, Bob applies U_i and measures the resulting state with the POVM $\{E_i, \mathbb{I} - E_i\}$. On the first outcome, Bob sends the message qubits to Alice; on the second outcome, he ends the protocol by outputting "1", i.e., Bob declares himself to be the winner.
  - Alice and Bob measure their part of the state with the final POVM and output the outcome of their measurements. Alice wins on outcome "0" and Bob on outcome "1".

The WCF protocol has the following properties:

- Correctness: When both players are honest, Alice and Bob's outcomes are always the same: $\Pi_A^{(0)} \otimes \mathbb{I}_\mathcal{M} \otimes \Pi_B^{(1)} |\psi_n\rangle = \Pi_A^{(1)} \otimes \mathbb{I}_\mathcal{M} \otimes \Pi_B^{(0)} |\psi_n\rangle = 0$.
- Balanced: When both players are honest, they win with probability 1/2: $P_A = \left|\Pi_A^{(0)} \otimes \mathbb{I}_\mathcal{M} \otimes \Pi_B^{(0)} |\psi_n\rangle\right|^2 = \frac{1}{2}$ and $P_B = \left|\Pi_A^{(1)} \otimes \mathbb{I}_\mathcal{M} \otimes \Pi_B^{(1)} |\psi_n\rangle\right|^2 = \frac{1}{2}$.
- ϵ biased: When Alice is honest, the probability that both players agree on Bob winning is $P_B^* \leq \frac{1}{2} + \epsilon$. And conversely, if Bob is honest, the probability that both players agree on Alice winning is $P_A^* \leq \frac{1}{2} + \epsilon$.

To be able to define the bias, we need P_A^* and P_B^* which correspond to the best possible cheating strategy of the opponent. The primal semi-definite program (SDP) formalises this statement.

Theorem 4 (Primal)**.**
$P_B^* = \max Tr((\Pi_A^{(1)} \otimes \mathbb{I}_\mathcal{M})\rho_{AM,n})$ *over all* $\rho_{AM,i}$ *satisfying the constraints*

- $Tr_\mathcal{M}(\rho_{AM,0}) = Tr_{\mathcal{MB}}(|\psi_0\rangle\langle\psi_0|) = |\psi_{A,0}\rangle\langle\psi_{A,0}|$;
- *for* i *odd,* $Tr_\mathcal{M}(\rho_{AM,i}) = Tr_\mathcal{M}(E_i U_i \rho_{AM,i-1} U_i^\dagger E_i)$;
- *for* i *even,* $Tr_\mathcal{M}(\rho_{AM,i}) = Tr_\mathcal{M}(\rho_{AM,i-1})$.

$P_A^* = \max Tr((\mathbb{I}_\mathcal{M} \otimes \Pi_B^{(0)})\rho_{MB,n})$ *over all* $\rho_{BM,i}$ *satisfying the constraints*

- $Tr_\mathcal{M}(\rho_{MB,0}) = Tr_{\mathcal{AM}}(|\psi_0\rangle\langle\psi_0|) = |\psi_{B,0}\rangle\langle\psi_{B,0}|$;
- *for* i *even,* $Tr_\mathcal{M}(\rho_{MB,i}) = Tr_\mathcal{M}(E_i U_i \rho_{MB,i-1} U_i^\dagger E_i)$;
- *for* i *odd,* $Tr_\mathcal{M}(\rho_{MB,i}) = Tr_\mathcal{M}(\rho_{MB,i-1})$.

A feasible solution to an optimisation problem is one which satisfies the constraints but is not necessarily optimal. A feasible solution to the primal problems gives a lower bound on P_A^* and P_B^*. If we consider the duals instead, it is known that, a feasible solution gives an upper bound on P_A^* and P_B^*. This certifies how good the protocol is.

Theorem 5 (Dual)**.**
$P_B^* = \min Tr(Z_{A,0} |\psi_{A,0}\rangle\langle\psi_{A,0}|)$ *over all* $Z_{A,i}$ *under the constraints*

1. $\forall i,\ Z_{A,i} \geq 0$;
2. *for* i *odd,* $Z_{A,i-1} \otimes \mathbb{I}_\mathcal{M} \geq U_{A,i}^\dagger E_{A,i} (Z_{A,i} \otimes \mathbb{I}_\mathcal{M}) E_{A,i} U_{A,i}$;
3. *for* i *even,* $Z_{A,i-1} = Z_{A,i}$;
4. $Z_{A,n} = \Pi_A^{(1)}$.

$P_A^* = \min Tr(Z_{B,0} |\psi_{B,0}\rangle\langle\psi_{B,0}|)$ *over all* $Z_{B,i}$ *under the constraints*

1. $\forall i,\ Z_{B,i} \geq 0$;
2. *for* i *even,* $\mathbb{I}_\mathcal{M} \otimes Z_{B,i-1} \geq U_{B,i}^\dagger E_{B,i} (\mathbb{I}_\mathcal{M} \otimes Z_{B,i}) E_{B,i} U_{B,i}$;
3. *for* i *odd,* $Z_{B,i-1} = Z_{B,i}$;

4. $Z_{B,n} = \prod_B^{(0)}$.

We add one more constraint to the above dual SDPs.

5. $|\psi_{A,0}\rangle$ *is an eigenvector of* $Z_{A,0}$ *with eigenvalue* $\alpha > 0$ *and* $|\psi_{B,0}\rangle$ *is an eigenvector of* $Z_{B,0}$ *with eigenvalue* $\beta > 0$.

Definition 6 (dual feasible points)**.** We call *dual feasible points* any two sets of matrices $\{Z_{A,0}, \ldots, Z_{A,n}\}$ and $\{Z_{B,0}, \ldots, Z_{B,n}\}$ that satisfy the corresponding conditions 1 to 5 as listed in Theorem 5.

It turns out that strong duality holds for the primal problems which means that there is a cheating strategy for Alice and Bob matching the upper bound on P_A^* and P_B^* respectively.

Proposition 7. $P_A^* = \inf \alpha$ *and* $P_B^* = \inf \beta$ *where the infimum is over all dual feasible points and* β, α *are defined in constraint 5 of the definition of the dual feasible points.*

2.2 (Time Dependent) Point Games with EBM transitions

The basic idea here is to remove all inessential information, that is the basis information, from the two aforesaid dual problems. Kitaev's genius was to achieve this by considering, at a given step, the dual variable Z_A, Z_B as observables with $|\psi\rangle$ governing the probability. This combines the evolution of the certificates on cheating probabilities with the evolution of the honest state—the state obtained when both players follow the protocol (nobody cheats). Originally, using a similar manoeuvre, Kitaev settled solvability of the quantum strong coin flipping problem by giving a bound on ϵ. To make this insight precise, first "prob" is defined.

Definition 8 (prob)**.** Consider $Z \geq 0$ and let $\Pi^{[z]}$ represent the projector on the eigenspace of eigenvalue $z \in \mathrm{sp}(Z)$. We have $Z = \sum_z z\Pi^{[z]}$. Let $|\psi\rangle$ be a (not necessarily normalised) vector. We define the function with finite support $\mathrm{prob}[Z, \psi] : [0, \infty) \to [0, \infty)$ as

$$\mathrm{prob}[Z, \psi](z) = \begin{cases} \langle\psi| \Pi^{[z]} |\psi\rangle & \text{if } z \in \mathrm{sp}(Z) \\ 0 & \text{else.} \end{cases}$$

If $Z = Z_A \otimes \mathbb{I}_{\mathcal{M}} \otimes Z_B$, using the same notation, we define the 2−variate function with finite support $\mathrm{prob}[Z_A, Z_B, \psi] : [0, \infty) \times [0, \infty) \to [0, \infty)$ as

$$\mathrm{prob}[Z_A, Z_B, \psi](z_A, z_B) = \begin{cases} \langle\psi| \Pi^{[z_A]} \otimes \mathbb{I}_{\mathcal{M}} \otimes \Pi^{[z_B]} |\psi\rangle & \text{if } (z_A, z_B) \in \mathrm{sp}(Z_A) \times \mathrm{sp}(Z_B), \\ 0 & \text{else.} \end{cases}$$

Think of the aforesaid 2–variate function as assigning a weight on each point of the plane. Going from one such configuration to another is what we would intuitively refer to as a "move" for the moment. Notice that at an odd step i, the dual variable Z_B doesn't change while Z_A does (see Theorem 5). The constraint equation in this step is $Z_{A,i-1} \otimes \mathbb{I}_M \geq U_i^\dagger (Z_{A,i} \otimes \mathbb{I}_M) U_i$. The honest state can be expressed as $|\psi_i\rangle = U_i |\psi_{i-1}\rangle$ but this acts on the complete $\mathcal{A} \otimes \mathcal{M} \otimes \mathcal{B}$ space. Applying the aforesaid method of removing the basis information using the prob method, and appending the fixed $Z_{B,i-1} = Z_{B,i}$, we conclude that $\mathrm{prob}(Z_{A,i-1} \otimes \mathbb{I}_M \otimes Z_{B,i}, |\psi_{i-1}\rangle) \to \mathrm{prob}(Z_{A,i} \otimes \mathbb{I}_M \otimes Z_{B,i}, |\psi_i\rangle)$ should constitute an "allowed move" as it is simply re-expressing the dual SDP in a basis independent form. For the dual, we are assuming the protocol is given to us, i.e. U_i (unitary operations), $\Pi_{A/B}$ (measurements) and $|\psi_0\rangle$ (initial state) are specified, and we have to find the appropriate Zs. However, when we discuss the notion of an "allowed move" we are moving towards a framework which will free us from discussing a specific protocol. This motivates the following definitions.

Definition 9 (EBM line transition)**.** Let $g, h : [0,\infty) \to [0,\infty)$ be two functions with finite supports. The line transition $g \to h$ is EBM if there exist two matrices $0 \leq G \leq H$ and a (not necessarily normalised) vector $|\psi\rangle$ such that $g = \text{prob}[G, \psi]$ and $h = \text{prob}[H, \psi]$.

Definition 10 (EBM transition)**.** Let $p, q : [0,\infty) \times [0,\infty) \to [0,\infty)$ be two functions with finite supports. The transition $p \to q$ is an

- EBM horizontal transition if for all $y \in [0,\infty)$, $p(.,y) \to q(.,y)$ is an EBM line transition, and
- EBM vertical transition if for all $x \in [0,\infty)$, $p(x,.) \to q(x,.)$ is an EBM line transition.

It turns out that when one writes the dual, the order of the constraints gets inverted, i.e. the condition associated with the final measurements and states appears first and the condition associated with the initial state appears in the end. We expect the final state to be like an EPR state and, intuitively, expect two points (in terms of the 2–variate function as described earlier) to be associated with it. This makes it plausible that we will start with two points in the dual when it is expressed in the aforementioned basis independent way. The initial state of the protocol is unentangled. This we expect should correspond to a single point. This helps us accept that we end with a single point in the basis independent expression of the dual. The rules for moving these points must be related to the dual constraints. We already formalised these conditions into EBM transitions. The notation

$$[x_g, y_g](x,y) = \begin{cases} 1 & x_g = x \text{ and } y_g = y \\ 0 & \text{else} \end{cases}$$

will be useful for formalising the complete description into what Mochon dubbed an "Expressible by Matrices" (Time Dependent) point game.

Definition 11 (EBM point game)**.** An EBM point game is a sequence of functions $\{p_0, p_1, \ldots, p_n\}$ with finite support such that

- $p_0 = 1/2[0,1] + 1/2[1,0]$;
- for all even i, $p_i \to p_{i+1}$ is an EBM vertical transition;
- for all odd i, $p_i \to p_{i+1}$ is an EBM horizontal transition;
- $p_n = 1[\beta, \alpha]$ for some $\alpha, \beta \in [0,1]$. We call $[\beta, \alpha]$ the final point of the EBM point game.

Since we started with a WCF protocol, considered its dual and re-expressed it as a TDPG (which is just a basis independent representation), the following should not come as a surprise.

Proposition 12 (WCF $\implies$ EBM point game)**.** *Given a WCF protocol with cheating probabilities P_A^* and P_B^*, along with a positive real number $\delta > 0$, there exists an EBM point game with final point $[P_B^* + \delta, P_A^* + \delta]$.*

What is slightly more non-trivial is that given this TDPG one can construct a WCF protocol. This means that by using only "allowed moves" one can be sure that there exists a corresponding sequence of unitaries U_i, the measurements $\Pi_{A/B}$ and the initial state $|\psi_0\rangle$ complemented by the dual variables $Z_{A,i}$ and $Z_{B,i}$ which certify the bias corresponding to the coordinates of the final point in the point game. This establishes the equivalence between TDPGs and WCF protocols. The precise statement is as follows.

Theorem 13 (EBM to protocol)**.** *Given an EBM point game with final point* $[\beta, \alpha]$*, there exists a WCF protocol and two dual feasible points proving that the optimal cheating probabilities are* $P_A^* \leq \alpha$ *and* $P_B^* \leq \beta$*.*

Our first contribution is related to this part. We construct protocols from EBM point games in a slightly different way which results in two important improvements. The first improvement makes the protocol more practical as the message register gets decoupled from Alice and Bob's registers after each round. (In Mochon/Aharonov et al's version the message register is highly entangled and stays that way until the very end.) The second improvement is due to the addition of a cheat detection measurement at every round (similar to Mochon's improved Dip Dip Boom protocol) which allows us to consider certain matrices with infinite eigenvalues in a well defined way. These pave the path for converting the bias 1/10 point game (due to Mochon; will be introduced later) into a protocol.

2.3 (Time Dependent) Point Games with valid transitions

While the problem has been simplified by the removal of the basis information, it is still hard to know which transitions are allowed, i.e. are EBM transitions. This is because finding the matrices certifying that a transition is EBM is not easy. The goal of this section is to find another criterion for establishing that a transition is EBM. This criterion is at the heart of coin flipping. It would turn out that the set of EBM functions (closely related to EBM transitions) form a closed convex cone. The dual of this cone happens to be the set operator monotone functions (as described earlier in Subsection 1.3, these are a generalisation of monotone functions, $x \geq y \implies f(x) \geq f(y)$, to $X \geq Y \implies f(X) \geq f(Y)$ where X and Y are now matrices). These functions have a very nice and simple characterisation. This is what leads to the key simplification for WCF. To be able to harness this, one can use the known fact that for a closed convex cone, the dual of the dual is the original cone itself (also called bi-dual). So this dual of operator monotone functions, i.e. the bi-dual of the cone of EBM functions, equals the cone of EBM functions. The dual of operator monotone functions has an easy description because operator monotone functions have an easy description. Combining these, one obtains an easy characterisation of EBM functions which allows one to construct interesting WCF protocols. The catch is that we establish that the two cones, the cone of EBM functions and the dual of the cone of operator monotone functions, are the same but given a point in the second cone we do not have a recipe for finding the matrices certifying it is an EBM. Without the matrices we can not implement the protocol even though we know the matrices must exist as the cones are the same. These notions are now formalised.

2.3.1 Formalising the equivalence between transitions and functions

Working with functions instead of transitions will be rather useful as will be evident from the next subsection.

Definition 14 (K, EBM functions)**.** A function $a : \mathbb{R}_{\geq 0} \to \mathbb{R}$ with finite support is an *EBM function* if the line transition $a^- \to a^+$ is EBM, where $a^+ : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $a^- : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ denote, respectively, the positive and the negative part of a ($a = a^+ - a^-$).

We denote by K the set of EBM functions.

Definition 15 (K_Λ, EBM functions on $[0, \Lambda]$)**.** For any finite Λ, a function $a : [0, \Lambda) \to \mathbb{R}$ with finite support is an *EBM function with support on* $[0, \Lambda]$ if the line transition $a^- \to a^+$ is EBM with its spectrum in $[0, \Lambda]$, where $a^- : [0, \Lambda) \to \mathbb{R}_{\geq 0}$ and $a^+ : [0, \Lambda) \to \mathbb{R}_{\geq 0}$ denote, respectively, the positive and the negative part of a.

We denote the set of EBM functions with support on $[0, \Lambda]$ by K_Λ.

It is evident that if the functions g, h denoting the transition $g \to h$ have no common support then the function description uniquely captures the said transition. In this section we restrict to such transitions and therefore use them interchangeably. In later sections we revisit this notion.

To be able to talk about different characterisations of EBM functions it is useful to abstract it (the characterisation) into a property $\mathcal{P}$ which the function must satisfy. Using this we can define games which use these $\mathcal{P}$ functions. This is done to be able to handle subtleties which arise in proving that the set of EBM functions is the same as the set of $\mathcal{P}$ functions for specific $\mathcal{P}$s.

Definition 16 (Horizontal and vertical $\mathcal{P}$-functions)**.** A $\mathcal{P}$-function $a : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is a function with finite support that has the property $\mathcal{P}$.

A function $t : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is a

- *horizontal $\mathcal{P}$-function* if for all $y \geq 0$, $t(., y)$ is a $\mathcal{P}$-function;
- *vertical $\mathcal{P}$-function* if for all $x \geq 0$, $t(x, .)$ is a $\mathcal{P}$-function.

Definition 17 (point games with $\mathcal{P}$-functions)**.** A point game with $\mathcal{P}$-functions is a set $\{t_1, \ldots, t_n\}$ of n $\mathcal{P}$-functions alternatively horizontal and vertical such that

- $\frac{1}{2}[0, 1] + \frac{1}{2}[1, 0] + \sum_{i=1}^{n} t_i = [\beta, \alpha]$;
- $\forall j \in \{1, \ldots, n\}$, $\frac{1}{2}[0, 1] + \frac{1}{2}[1, 0] + \sum_{i=1}^{j} t_i \geq 0$.

We call $[\beta, \alpha]$ the final point of the game.

We note the following before looking at $\mathcal{P}$ functions in more detail.

Lemma 18 (point game with EBM functions $\implies$ point game with EBM transitions)**.** *Given a point game with n EBM functions and final point $[\beta, \alpha]$ we can construct a point game with n EBM transitions and final point $[\beta, \alpha]$.*

2.3.2 Operator monotone functions and valid functions

This discussion is essential to understand our second contribution. The set of EBM functions forms a convex cone. To see this we recall the definition of a convex cone.

Definition 19 (convex cone)**.** A set C in a vector space V is a cone if for all $x \in C$ and for all $\lambda > 0$, $\lambda x \in C$. It is convex if for all $x, y \in C$, $x + y \in C$.

Noting that the state $|\psi\rangle$ in the definition of an EBM function (which in turn invokes an EBM transition) is unnormalised the set of EBM functions is easily seen to form a cone. By taking a direct sum one can establish convexity as well. The vector space of interest here is given by the span of the basis $\{[x_g]\}_{x_g \in [0,\infty)}$ where $[x_g](x) = \delta_{x_g, x}$. The notation is similar to the one introduced earlier. We use it shortly.

Lemma 20. *K is a convex cone. Also, for any $\Lambda \in (0, \infty)$, K_Λ is a convex cone.*

To establish an alternative characterisation of the cone of EBM functions we need to define what is called a dual cone.

Definition 21 (dual cone)**.** Let C be a cone in a normed vector space V. We denote by V' the space of continuous linear functionals from V to $\mathbb{R}$. The dual cone of a set $C \subseteq V$ is

$$C^* = \left\{\Phi \in V' | \forall a \in C, \Phi(a) \geq 0\right\}.$$

For our purpose linear functionals can be thought of simply as functions which map objects in the cone to a non-negative real number with the added property of being linear in its argument.

We now formally define operator monotone functions.

Definition 22 (operator monotone functions)**.** A function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is operator monotone if for all $0 \leq X \leq Y$ we have $f(X) \leq f(Y)$.

Definition 23 (operator monotone functions on $[0, \Lambda]$)**.** A function $f : [0, \Lambda] \to \mathbb{R}$ is operator monotone on $[0, \Lambda]$ if for all $0 \leq X \leq Y$ with spectrum in $[0, \Lambda]$ we have $f(X) \leq f(Y)$.

The pivotal result of this (sub)section is the equivalence between the cone of operator monotone functions and the dual cone of EBM functions.

Lemma 24. *$\Phi \in K^*$ if and only if f_Φ is operator monotone in $[0, \infty]$. Also, for any $\Lambda \in (0, \infty)$, $\Phi \in K^*_\Lambda$ if and only if f_Φ is operator monotone on $[0, \Lambda]$.*

(NB: We need to use the bijection between Φ (a linear functional from $V \to \mathbb{R}$) and a function on reals (from $\mathbb{R} \to \mathbb{R}$) given by the identification $f_\Phi(x) = \Phi([x])$ to make such a statement)

The proof of this crucial result is not too hard (almost trivial in one direction) and follows from the respective definitions with some work for unpacking. What makes this connection interesting is the following beautiful characterisation of operator monotone functions introduced by Löwner (in 1934, see [17]).

Lemma 25 (characterisation of operator monotone functions)**.** *Any operator monotone function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ can be written as*

$$f(x) = c_0 + c_1 x + \int_0^\infty \frac{\lambda x}{\lambda + x} d\omega(\lambda)$$

for a measure ω satisfying $\int_0^\infty \frac{\lambda}{1+\lambda} d\omega(\lambda) < \infty$.

Lemma 26 (characterisation of operator monotone functions on $[0, \Lambda]$)**.** *Any operator monotone function $f : [0, \Lambda] \to \mathbb{R}$ can be written as*

$$f(x) = c_0 + c_1 x + \int \frac{\lambda x}{\lambda + x} d\omega(\lambda)$$

with the integral ranging over $\lambda \in (-\infty, -\Lambda) \cup (0, \infty)$ satisfying $\int \frac{\lambda}{1+\lambda} d\omega(\lambda) < \infty$ where ω is a measure.

As will become clear when we discuss the dual of the cone of operator monotones, it suffices to consider operator monotones of the form $\lambda x/(\lambda + x)$ (which basically is because ω is a measure).

So far the statements from Aharonov et al's paper were made in the same order as they had originally appeared. We now re-order these a little with an eye on our end-goal (as opposed to the one of Mochon/Aharonov). It is known that the bi-dual of a cone is the closure of the cone we started with.

Fact 27. *Let $C \subseteq V$ be a convex cone, then $C^{**} = cl(C)$ where C^* is the dual cone of C.*[5]

[5] See [Boyd and Vandenberghe 2004] for proofs of these facts.

The astute reader would have guessed where we are going with this discussion. We define, from hindsight, the bi-dual of EBM functions to be the cone of valid functions. Since the dual of EBM functions has an easy characterisation, the bi-dual also has an easy characterisation which is why we are interested in it.

Definition 28 (Λ valid functions)**.** A function $a : [0, \Lambda] \to \mathbb{R}$ with finite support on $[0, \Lambda]$ is Λ valid if $a \in K_\Lambda^{**}$.

To be able to use the aforementioned fact we note that the cone of interest, the cone of EBM functions, is closed. As one can imagine, proving this is easier if the matrices involved have a bounded spectrum. We consider only these for now. This means that the cone of valid functions is the same as the cone of EBM functions. We state these precisely below.

Lemma 29. *For $\Lambda \in (0, \infty)$, K_Λ is closed (which implies $K_\Lambda^{**} = K_\Lambda$).*

Corollary 30. *For $\Lambda \in (0, \infty)$, $K_\Lambda = \{a \in V | \forall \Phi \in K_\Lambda^*, \Phi(a) \geq 0\}$.*

Corollary 31 (EBM on $[0, \Lambda]$ is equivalent to Λ valid)**.** *A function $a : [0, \Lambda] \to \mathbb{R}$ with finite support on $[0, \Lambda]$ is EBM on $[0, \Lambda]$ if and only if $\sum_x a(x) = 0$, $\sum_x xa(x) \geq 0$ and $\forall \lambda \in (-\infty, -\Lambda] \cup (0, \infty)$, $\sum_x \frac{\lambda x}{\lambda + x} a(x) \geq 0$.*

In the last statement, the characterisation of operator monotone functions was used which we introduced earlier. Note that all the statements made here assume that the matrices used in EBM functions have a finite spectrum. Our EMA algorithm heavily relies on this part of the analysis which is due to Aharonov et al.

It is worth pointing out that Mochon outlines this scheme used by Aharonov et al. but himself proceeds by using matrix perturbation theory for proving a similar result.

2.3.3 Strictly valid functions are EBM functions

To be able to simplify the conditions one needs to check, it is useful to relax the condition on the spectrum of the matrices involved. This is evident from range of λ one needs to use in the characterisation of operator monotone functions (compare Lemma 26 and Lemma 25).

It is easy to describe the interior of the dual of a cone. It is also possible to relate the interior with the closure of the cone, but in finite dimensions. This reasoning fails for infinite dimensions. They still serve as motivation for the definition of valid and strictly valid functions.

Fact 32. *Let C be a convex set, then $int(C) = int(cl(C))$.*

Fact 33. *Let C be a cone in the finite-dimensional vector space V, then $int(C^*) = \{\Phi \in V' | \forall a \in C - \{0\}, \Phi(a) > 0\}$.*

Definition 34 (valid function)**.** A function $a : \mathbb{R}_{\geq 0} \to \mathbb{R}$ with finite support is valid if for every operator monotone function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ we have $\sum_{x \in \text{supp}(h)} f(x)a(x) \geq 0$.

Definition 35 (strictly valid function)**.** A function $a : \mathbb{R}_{\geq 0} \to \mathbb{R}$ with finite support is strictly valid if for every non-constant operator monotone function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ we have $\sum_{x \in \text{supp}(a)} f(x)a(x) > 0$.

One can use the characterisation of operator monotone functions to explicitly characterise the set of valid and strictly valid functions.

Lemma 36. *Let $a : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be a function with finite support such that $\sum_x a(x) = 0$. The function a is a strictly valid function if and only if for all $\lambda > 0$, $\sum_x \frac{-a(x)}{\lambda+x} > 0$ and $\sum_x x.a(x) > 0$.*
(We added the last condition else the merge (discussed later) becomes a strictly valid function but it can be shown that no bounded matrices exist for which it is EBM.)

The function a is valid if and only if for all $\lambda > 0$, $\sum_x \frac{-a(x)}{\lambda+x} \geq 0$ and $\sum_x x.a(x) \geq 0$.

The set of strictly valid functions can be shown to also be Λ valid for some finite Λ. This means that it would also be EBM on $[0, \Lambda]$ which in turn means it would be an EBM function. We hence have the following.

Lemma 37. *Any strictly valid function is an EBM function.*

2.3.4 From valid functions to EBM functions

If we construct a point game with valid functions we can convert it into a game with EBM functions with an arbitrarily small overhead on the bias. The trick is to raise the coordinates of all the final points (ones with positive weight) a little at each step, to convert a valid function into a strictly valid function.

Theorem 38 (valid to EBM)**.** *Given a point game with $2m$ valid functions and final point $[\beta, \alpha]$ and any $\epsilon > 0$, we can construct a point game with $2m$ EBM functions and final point $[\beta + \epsilon, \alpha + \epsilon]$.*

Lemma 39. *Fix $\epsilon > 0$. Given a point game with $2m$ valid functions and final point $[\beta, \alpha]$ we can construct a point game with $2m$ strictly valid functions and final point $[\beta + \epsilon, \alpha + \epsilon]$.*

2.3.5 Examples of valid line transitions

We go back to transitions to discuss some simple valid and strictly valid line transitions which are defined similar to the corresponding functions.

Definition 40 (Valid and strictly valid line transitions)**.** Let $g, h : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be two functions with finite support. The transition $g \to h$ is valid (resp., strictly valid) if the function $h - g$ is valid (resp., strictly valid).

We focus our attention to the simplest cases. The first is to increase the coordinate of a point. The second considers the case of merging two points into one. The third is about splitting a single point into two.

Example 41 (Point raise)**.** $p[x_g] \to p[x_h]$ with $x_h \geq x_g$.

Example 42 (Point merge)**.** $p_{g_1}[x_{g_1}] + p_{g_2}[x_{g_2}] \to (p_{g_1} + p_{g_2})[x_h]$ with $x_h \geq \frac{p_{g_1} x_{g_1} + p_{g_2} x_{g_2}}{p_{g_1} + p_{g_2}}$.

Example 43 (Point split)**.** $p_g[x_g] \to p_{h_1}[x_{h_1}] + p_{h_2}[x_{h_2}]$ with $p_g = p_{h_1} + p_{h_2}$ and $\frac{p_g}{x_g} \geq \frac{p_{h_1}}{x_{h_1}} + \frac{p_{h_2}}{x_{h_2}}$.

The merge and split can be generalised to many points and be shown to have the same form.

2.4 TIPGs.

Mochon's Dip Dip Boom protocol, the one with bias 1/6, can be expressed already as a (time dependent) point game. However, it is possible to simplify the point game formalism even further and it is in this simplified formalism Mochon constructs his family of point games that achieve an arbitrarily small bias. Instead of worrying about the entire sequence of horizontal and vertical transitions, one can focus on just two functions as described below.

Definition 44 (TIPG)**.** A TIPG is a valid horizontal function a and a valid vertical function b such that

$$a + b = 1[\beta, \alpha] - \frac{1}{2}[0, 1] - \frac{1}{2}[1, 0]$$

for some $\alpha, \beta > 1/2$. We call the point $[\beta, \alpha]$ the final point of the game.

The main difference here is that we do not worry about the sequence in which one must apply the transitions to obtain the final configuration. This justifies the name TIPG which stands for a Time Independent Point Game. It is not too hard to see that if we have a valid point game we can combine the horizontal functions and the vertical functions to obtain a and b. It is a little counter-intuitive in fact to learn that one can convert a TIPG into a valid (time dependent) point game with an arbitrarily small cost on the bias. It is counter-intuitive because it is not clear that one can flesh out a time ordered sequence as one can, and in fact does for Mochon's point games, run into causal loops that is you expect a point to be present to create another point which in turn is required to produce the first point. The trick that is used to fix this problem is known as the "catalyst state". One deposits a little bit of weight wherever there is negative weight for a, for instance, and then one can implement a scaled down round of a and b. The scaling is proportional to the weight that is placed to start with. Repeating this procedure multiple times yields the required final state along with the "catalyst state" which stays unchanged. Absorbing the catalyst state leads to a small increase in the bias. The number of rounds increases with how small one wants this increase in bias to be.

Theorem 45 (TIPG to valid point games)**.** *Given a TIPG with a valid horizontal function a and a valid vertical function b such that $a + b = 1[\beta, \alpha] - \frac{1}{2}[0, 1] - \frac{1}{2}[1, 0]$, we can construct, for all $\epsilon > 0$, a valid point game with final point $[\beta + \epsilon, \alpha + \epsilon]$ where the number of transitions depends on ϵ.*

It is important to state that the conversion from a TIPG to a valid (time dependent) point game, TDPG, is easy and explicit.

A word about resource usage. The size (dimension) of the physical system we use depends on the number of points involved in the point games linearly. The number of rounds on the other hand needs to be calculated with more care as it depends on the choice of the catalyst state. These calculations with respect to Mochon's game and in general have been performed by Aharonov et al. in their article and we do not discuss it here.

We have stated enough results to be able to commence the discussion of our work.

Part I
Bias 1/10

3 TDPG → Explicit Protocol, Framework (TEF)

We strongly recommend that the reader looks at the third section titled "The illustrated guide to point games" from Mochon's [1] article, if they have not already, before proceeding.

3.1 Motivation and Conventions

We wish to construct a protocol such that its dual matches a given TDPG. The main difference in our construction, compared to the one used by Aharonov et al. and Mochon, is the introduction of a message register that decouples after each round and of suitably adapted projectors. Consequently, the non-trivial constraint that the dual matrices must satisfy would be similar to, but not quite the same as, the EBM condition.

Keep Definition 8 in mind. Intuitively, the most natural way of constructing Zs and a $|\psi\rangle$ given an arbitrary frame (think of a TDPG as a sequence of frames) is to construct an entangled state that encodes the weight and define Zs to contain the coordinates corresponding to the weight. Let us make this idea more precise.

Definition (Canonical Form)**.** The tuple $(|\psi\rangle, Z^A, Z^B)$ is said to be in the Canonical Form with respect to a set of points in a frame of a TDPG if (see Figure 4a) $|\psi\rangle = \sum_i \sqrt{P_i}\,|ii\rangle_{AB} \otimes |.\rangle_M$, $Z^A = (\sum x_i\,|i\rangle\langle i|_A) \otimes |.\rangle\langle .|_M$ and $Z^B = (\sum y_i\,|i\rangle\langle i|_B) \otimes |.\rangle\langle .|_M$ where $|.\rangle_M$ represent extra uncoupled registers which might be present.

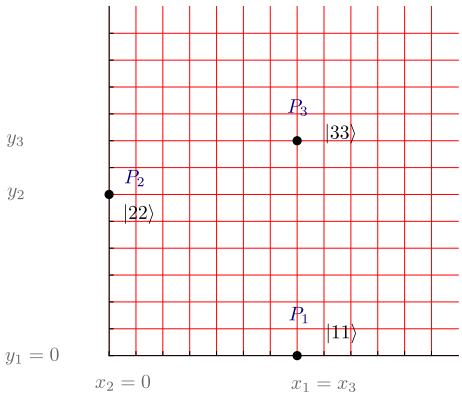
It is easy to see that the 'label' $|ii\rangle$ correspond to a point with coordinates x_i, y_i and weight P_i in the frame (see Definition 8). It is tempting to imagine that we systematically construct, from each frame of a TDPG, a canonical form of $|\psi\rangle$ s and Zs. The unitaries can be deduced from the evolution of $|\psi\rangle$. This approach has two problems, (1) it does not manifestly mean that the unitaries would be decomposable into moves by Alice and Bob who communicate only through the message 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 ensures these issues are dissolved.

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

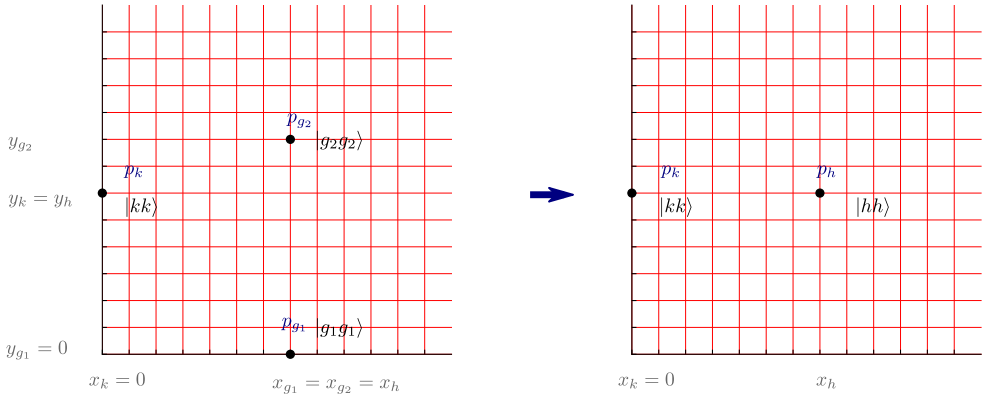
Basic Moves Work Out of the Box

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 $w_2 = 1/p_2$.



(a) Frame of a TDPG



(b) The points which are unchanged from one frame to another are labelled by $\{k_i\}$. Among the points that change, the initial ones are labelled by $\{g_i\}$ and the final ones by $\{h_i\}$.

Figure 4: Illustrations for the Canonical Form

We construct 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.

3.2 The Framework

Intuition

Imagine a 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 performs the non-trivial step). Let the labels of the points that one wishes to leave untouched be given by $\{k_i\}$ (see Figure 4b). We can write the state as

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

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

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

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 assume that h and g index orthonormal vectors. Alice can update the probabilities and labels by locally performing a unitary to obtain

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

It is precisely this step which yields the non-trivial constraint. 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.

Note that we add a few extra frames to the final TDPG to go from a given frame to the next of the initial TDPG. This is irrelevant as the bias stays the same but we mention it to avoid confusion.

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 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) $\left|\psi_{(2)}\right\rangle = U^{(1)}\left|\psi_{(1)}\right\rangle$ 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

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

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,

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

where

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

subject to the condition

$$\sum x_{h_i}\left|h_ih_i\right\rangle\left\langle h_ih_i\right| \geq \sum x_{g_i}E^{(2)}U^{(2)}\left|g_ig_i\right\rangle\left\langle g_ig_i\right|U^{(2)\dagger}E^{(2)} \tag{5}$$

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 $\mathcal{H} = \text{span}\{\left|g_1g_1\right\rangle, \left|g_2g_2\right\rangle\ldots, \left|h_1h_1\right\rangle, \left|h_2,h_2\right\rangle\ldots\}$. We separate all expressions (they are nearly diagonal) into the $\mathcal{H}$ space (which gets non-diagonal) and the rest. We start with the RHS,

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

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,

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

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

$$\sum x_{h_i} |h_i\rangle \langle h_i| \otimes (\mathbb{I} - |h_i\rangle \langle h_i|) \geq 0$$
$$\sum (x_{k_i} - x_{k_i}) |k_i\rangle \langle k_i| \otimes \mathbb{I} \geq 0$$

which essentially only implies

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

Finally the non-trivial part yields

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

which completes the proof. □

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

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

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

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 $U^\dagger E$ on $\left|\psi_{(4)}\right\rangle$ (where E is defined to be $E^{(3)}$ and U to be $U^{(3)}$ for the proof).
(2) Note that

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

Since the other term in $Z_3^B \otimes \mathbb{I}$ is anyway in the non-action space of U it follows that

$$EU(Z_3^B \otimes \mathbb{I})U^\dagger E = y \sum |h_i\rangle \langle h_i| \otimes \mathbb{I}_M^{\{m\}} + \sum y_{k_i} |k_i\rangle \langle k_i| \otimes \mathbb{I}_M.$$

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

We can summarise the condition of interest as follows, the proof of which is a trivial consequence of the aforesaid.

Theorem 46. *For an x-transition (where Alice performs the non-trivial step)*

$$\sum_{i=1}^{n_k} p_{k_i}[x_{k_i}] + \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}] \to \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}] + \sum_{i=1}^{n_k} p_{k_i}[x_{k_i}]$$

to be implementable under the TDPG-to-Explicit-protocol Framework (TEF) one must find a $U^{(2)}$ that satisfies the inequality

$$\sum_{i=1}^{n_h} x_{h_i} |h_i h_i\rangle \langle h_i h_i|_{AM} \geq \sum_{i=1}^{n_g} x_{g_i} E_h^{(2)} U^{(2)} |g_i g_i\rangle \langle g_i g_i|_{AM} U^{(2)\dagger} E_h^{(2)} \tag{6}$$

and the honest action constraint

$$U^{(2)} |v\rangle = |w\rangle$$

where $|h_i\rangle$ and $|g_i\rangle$ are orthonormal basis vectors,

$$|v\rangle = \mathcal{N}\left(\sum \sqrt{p_{g_i}} |g_i g_i\rangle_{AM}\right)$$

and

$$|w\rangle = \mathcal{N}\left(\sum \sqrt{p_{h_i}} |h_i h_i\rangle_{AM}\right)$$

for $\mathcal{N}(|\psi\rangle) = |\psi\rangle / \sqrt{\langle\psi|\psi\rangle}$, $E_h = (\sum_{i=1}^{n_h} |h_i\rangle \langle h_i|_A + \sum |k_i\rangle \langle k_i|_A) \otimes \mathbb{I}_M$ with $U^{(2)}$'s non-trivial action restricted to span$\{\{|g_i g_i\rangle_{AM}\}, \{|h_i h_i\rangle_{AM}\}\}$ (note $|k_i\rangle$ corresponds to the points that are left unchanged in the transition).

3.3 Important Special Case: The Blinkered Unitary

So far we have not specified the non-trivial $U^{(2)}$ (which we 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

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

  We 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 (substituting for 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

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

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

$$\frac{x_{g_1}}{N^2}\left(\frac{p_{h_1}}{x_{h_1}} + \frac{p_{h_2}}{x_{h_2}}\right) \leq 1$$

  where in the second step we used the fact that $F - M \geq 0$ implies $\mathbb{I} - \sqrt{F}^{-1}M\sqrt{F}^{-1} \geq 0$ (if $F > 0$) and the last step is obtained by writing the matrix in the previous step as $|\psi\rangle\langle\psi|$ followed by demanding $1 \geq \langle\psi|\psi\rangle$.
  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.

  It would not be surprising to learn/prove that the class of unitaries with these properties is much more general than the Blinkered Unitaries.

- 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

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

  or more explicitly,

$$\frac{1}{\sum_{i=1}^{m} p_{g_i} x_{g_i}} \geq \sum_{i=1}^{n} p_{h_i} . \frac{1}{x_{h_i}},$$

  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).

4 Games and Protocols

We now describe two games, the bias 1/6 game and the bias 1/10 game, from the family of games constructed by Mochon to show that arbitrarily small bias is achievable. Mochon parametrises his games by k which determines the number of points involved in the non-trivial step. The bias he obtains is $\epsilon = 1/(4k+2)$. We consider games with $k = 1$ and $k = 2$, yielding the aforementioned bias.

4.1 Mochon's Approach

4.1.1 Assignments

Recall that a function

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

is valid if

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

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 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. *We have $\sum_{i=1}^{n} \frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)} = (-1)^{n-1}$ for $n \geq 2$ (proof in section B of the Appendix).*

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} being ± 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.

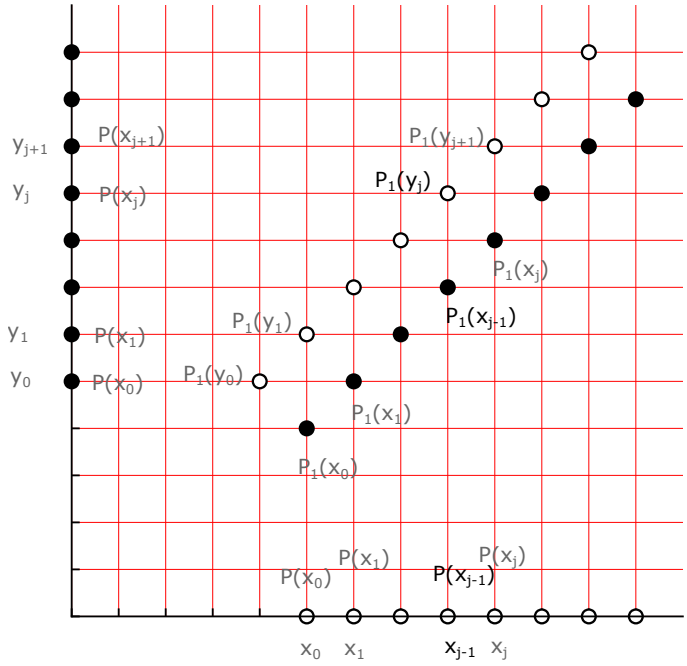


Figure 5: Building a TDPG/TIPG using merge moves

4.1.2 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.

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

Similarly with the point $[0, y_j]$ we associate $P(y_j)$ where $y_j = x_j$ as we 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.

4.2 Bias 1/6

4.2.1 Game

With reference to Figure 5 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

$$0 \leftrightarrow x_{g_1}$$
$$y_j \leftrightarrow x_{g_2}$$
$$y_{j-2} \leftrightarrow x_{h_1}$$

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

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 as it helps approximating $\sum P(x_j)$ by an integral). Plugging this back we have

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

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

$$\sum_{j=0}^{n} P(x_j) = \frac{1}{2} \to c_0 \int_{x_0}^{\Gamma} \frac{dx}{x^3} = \frac{c_0}{(-2)} \left[\frac{1}{\Gamma^2} - \frac{1}{x_0^2} \right] = \frac{1}{2}$$

which entails $c_0 = x_0^2$. The next condition yields 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. These issues have been dealt with in general (see [6]).

4.2.2 Protocol

Although we could only claim that one can construct the protocol once the unitaries are known, the basic idea is that 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 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.

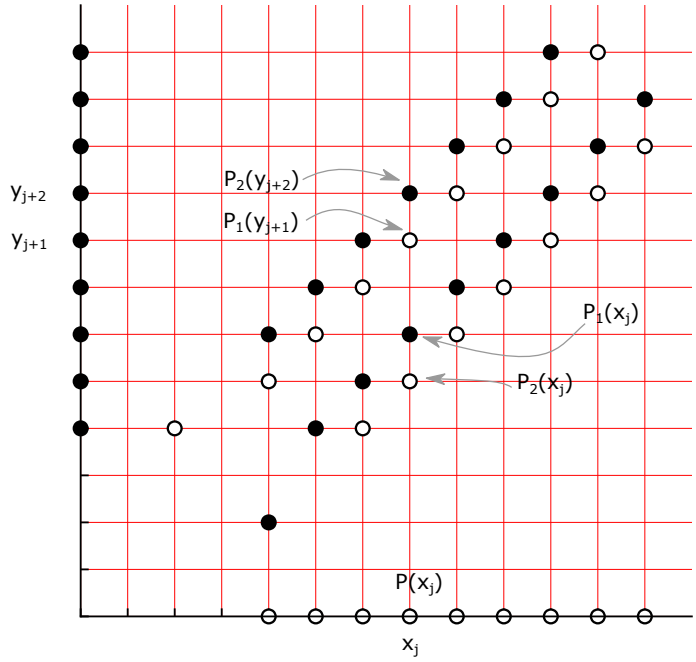


Figure 6: 1/10 game: The $3 \to 2$ move based TIPG for bias 1/10

4.3 Bias 1/10 Game

With respect to Figure 6 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

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

where we added the minus sign to account for the fact that f is negative for coordinates between y_{-2} and Γ_1. Imposing the symmetry constraints $P_1(y_j) = P_1(x_j)$ we 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 $P_2(y_j) = P_2(x_j)$, viz. it satisfies 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

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

4.4 Bias 1/10 Protocol

4.4.1 The $3 \to 2$ Move

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

Recall that

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

and let

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

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

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

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 do not influence the constraint equation. That is what we evaluate now. We need terms of the form $EU\,|g_i\rangle$ with $E = \mathbb{I}^{\{h_i\}}$. This entails that on the $\{|g_i\rangle\}$ space

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

Consequently we have

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

Recall that the constraint equation was

$$\sum x_{h_i}\,|h_i\rangle\langle h_i| - \sum x_{g_i}E_hU\,|g_i\rangle\langle g_i|\,U^\dagger E_h \geq 0$$

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 $|w\rangle, |w_1\rangle$ 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

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

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

which means that the constraint equation becomes

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

We already showed that Mochon's transition is average non-decreasing viz. $\langle x_h \rangle - \langle x_g \rangle \geq 0$. We 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 basis and techniques were chosen based on experience. Now we claim that $\theta \frac{N_g}{N_{v_2}} = \mathcal{O}(\delta y)$ 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 $\theta \frac{N_g}{N_{v_2}}$ limit and to first order in it the constraints become

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

and

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

If our claim is wrong when we evaluate $\theta \frac{N_g}{N_{v_2}}$ we will get zero order terms but as we show in the following section $\theta \frac{N_g}{N_{v_2}} = 0.\delta y + \mathcal{O}(\delta y^2)$ in fact.

4.4.2 Validity of the $3 \to 2$ Move

With respect to Figure 6 we have

$$\begin{aligned}
P_2(y_{j+2}) = p_{h_2} &= \frac{-f(y_{j+2})}{4.3 \delta y^2 y_{j+2}} \\
P_1(y_{j+1}) = p_{g_3} &= \frac{-f(y_{j+1})}{3.2 \delta y^2 y_{j+1}} \\
P_1(x_j) = p_{h_1} &= \frac{-f(y_{j-1})}{3.2 \delta y^2 y_{j-1}} \\
P_2(x_j) = p_{g_2} &= \frac{-f(y_{j-2})}{4.3 \delta y^2 y_{j-2}} \\
P(x_j) = p_{g_1} &= \frac{f(0) \delta y}{y_{j+2} y_{j+1} y_{j-1} y_{j-2}}
\end{aligned}$$

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 now convert all expressions to first order in δy. To this end we note

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

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

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

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

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

and similarly

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

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

$$\begin{aligned}\theta\frac{N_g}{N_{v_2}} &= \frac{4\sqrt{\frac{1}{12.6}}(-3\delta y)\left[f.(\not{1} + \frac{\delta y}{2f}\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right)) - f.(\not{1} - \frac{\delta y}{2f}\left(\frac{\partial f}{\partial y} - \frac{f}{y_j}\right)) + \mathcal{O}(\delta y^2)\right]}{\langle x_g\rangle} \\ &= 0 + \mathcal{O}(\delta y^2).\end{aligned}$$

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

Now we show that the second diagonal element is positive to first order in δy. Using the fact that $\theta\frac{N_g}{N_{v_2}} = \mathcal{O}(\delta y^2)$ 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

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

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.

Note that only the proof of validity was done perturbatively 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 in Figure 6 as they form a $3 \to 1$ merge and a $2 \to 1$ merge (possibly followed by a raise). The only remaining task is implementing the $2 \to 2$ move in the last step because we assumed here that $\sqrt{p_{g_2}} \neq 0$ (else the vectors which we assumed are orthonormal, cease to be so).

4.4.3 The $2 \to 2$ Move and its validity

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

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

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

$$|v\rangle = \frac{1}{N_g}\left(\sqrt{p_{g_1}}|g_1\rangle + \sqrt{p_{g_2}}|g_2\rangle\right),$$

$$|w\rangle = \frac{1}{N_h}\left(\sqrt{p_{h_1}}\,|h_1\rangle + \sqrt{p_{h_2}}\,|h_2\rangle\right),$$
$$|v_1\rangle = \frac{1}{N_g}\left(\sqrt{p_{g_2}}\,|g_1\rangle - \sqrt{p_{g_1}}\,|g_2\rangle\right),$$

and

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

We evaluate the constraint equation using

$$E_h U\,|g_{11}\rangle = \frac{\sqrt{p_{g_1}}\,|w\rangle + \beta e^{-i\phi_g}e^{i\phi_h}\sqrt{p_{g_2}}\,|w_1\rangle}{N_g}$$
$$E_h U\,|g_{22}\rangle = \frac{\sqrt{p_{g_2}}\,|w\rangle - \beta e^{-i\phi_g}e^{i\phi_h}\sqrt{p_{g_1}}\,|w_1\rangle}{N_g}$$

and

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

(similarly for $L\,|g_{22}\rangle\langle g_{22}|\,L^\dagger$) 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 use the same strategy as above and take the first diagonal element to be zero. Our burden would be to first show that

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

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 7 we note that

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

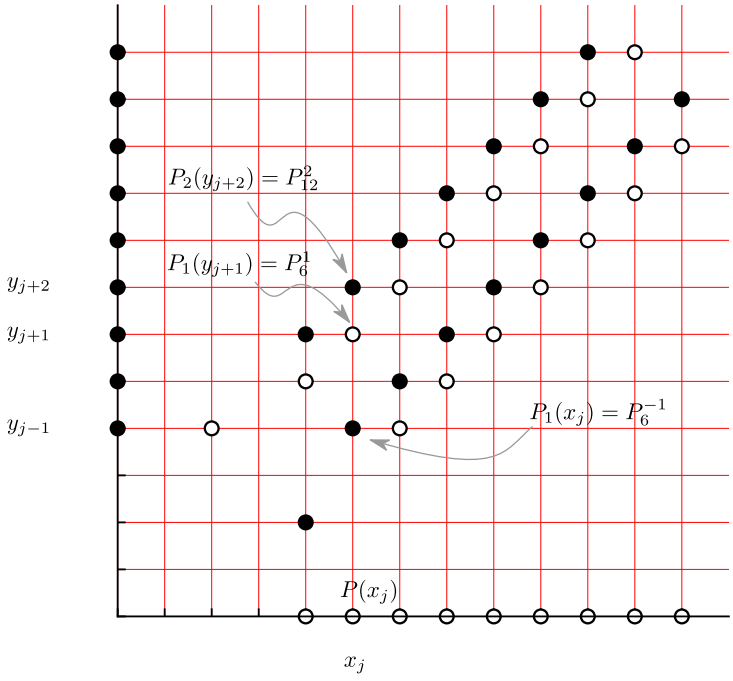


Figure 7: First $2 \to 2$ Transition

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

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

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.

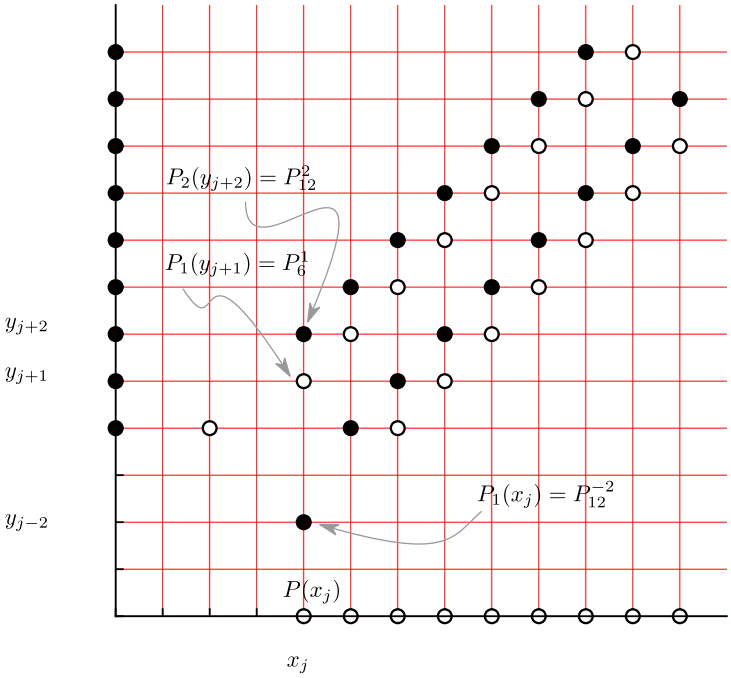


Figure 8: Final $2 \to 2$ Transition.

The $2 \to 2$ move corresponding to the leftmost (see Figure 8) and bottommost set of points can be shown to be implementable similarly.

Part II
Elliptic Monotone Algorithm (EMA)

5 Canonical Forms Revisited

We note that to construct the unitaries involved in the bias 1/10 protocol we did not follow any systematic recipe. We now switch gears and construct an algorithm that can generate the required unitary for any given Λ-valid function (see Definition 28). Note that corresponding to any WCF protocol with valid functions, one can find a WCF protocol with strictly valid functions (see Lemma 39). All strictly valid functions are Λ-valid for some finite Λ (see Lemma 37, Corollary 31). Thus we do not lose generality by restricting to Λ-valid functions.

In this section we formalise the non-trivial constraint Equation (5) into two forms which we call the Canonical Projective Form (CPF) and the Canonical Orthogonal Form (COF). The CPF is always well defined but the corresponding COF may contain diverging eigenvalues. However since we restrict to Λ-valid functions, as we will see, the COF will also always be well defined. We need the COF as it is this that we use in the Elliptic Monotone Algorithm (EMA) algorithm.

We continue to use purple for the intuitive parts and start using blue for the proofs.

5.1 The Canonical Projective Form (CPF) and the Canonical Orthogonal Form (COF)

We always use the convention $p_{g_i}, p_{h_i} > 0$. This is important else in some of the statements one can find trivial counter-examples. Recall Theorem 46 which formally states the main result of Section 3. Note that the number of points initially, n_g, and finally, n_h, may differ. To facilitate further discussion we formalise the aforesaid condition into an object and its property. First, however, we define the following notation.

Definition 47 (Transition)**.** Consider two finitely supported functions $g, h : \mathbb{R}_{\geq} \to \mathbb{R}_{\geq}$. A transition is defined as $g = \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}] \to h = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}]$ where $[y](x) := \delta_{xy}$ and $p_{g_i} > 0$, $p_{h_i} > 0$.

Definition 48 (Canonical Projective Form (CPF) for a give transition)**.** For a given transition the *Canonical Projective Form (CPF)* is given by the set of $m \times m$ matrices X_h, X_g, U, D and m dimensional vectors $|v\rangle$, $|w\rangle$ where

$$X_h := \sum_{i=1}^{n_h} x_{h_i} |h_i\rangle \langle h_i| , \; X_g := \sum_{i=1}^{n_g} x_{g_i} |g_i\rangle \langle g_i| ,$$

$$|w\rangle := \sum_{i=1}^{n_h} \sqrt{p_{h_i}} |h_i\rangle , \; |v\rangle := \sum_{i=1}^{n_g} \sqrt{p_{g_i}} |g_i\rangle ,$$

$$D := X_h - E_h U X_g U^\dagger E_h$$

and U is a unitary which satisfies

$$U |v\rangle = |w\rangle$$

for $E_h = \sum |h_i\rangle \langle h_i|$, orthonormal basis vectors $\{|g_1\rangle , |g_2\rangle \dots |g_{n_g}\rangle , |h_1\rangle , |h_2\rangle \dots |h_{n_h}\rangle\}$, $m = n_g + n_h$.

Definition 49 (legal CPF)**.** A CPF is *legal* if $D \geq 0$.

In this language then our objective is to find a legal CPF for a given transition.

Surprising as it may seem it *suffices to restrict to real unitaries viz. orthogonal matrices*. This will be justified in the next section but we already make this restriction in everything that follows (unless stated otherwise). In this section we try to reach an equivalence between a legal CPF and what we call the legal Canonical Orthogonal Form (COF).

The latter will be, roughly speaking, an inequality of the form $X_h - OX_gO^T \geq 0$ where $X_h = \text{diag}\{x_{h_1}, x_{h_2} \ldots, x_{h_{n_h}}, \xi, \xi \ldots\}$ and $X_g = \text{diag}\{x_{g_1}, x_{g_2} \ldots, x_{g_{n_g}}, 0, 0 \ldots\}$ for a large ξ. It is easy to see that if we can find an O that satisfies the COF for a given transition then the same O would satisfy the TEF inequality. It is almost trivial to note that a Λ valid function admit matrices of the COF form but we will show this later. Proving the other way, i.e. every legal CPF entails the corresponding COF must also be legal, is more non-trivial. Doing this requires handling the infinities and the matrix sizes more carefully. We only sketch an argument for this as we do not use it in the algorithm.

Definition 50 ((n, ξ) Canonical Orthogonal Form (COF) for a transition, ξ COF for a transition)**.** For a given transition and two numbers $n \geq \max(n_h, n_g)$, $\xi \geq \max(x_{h_1}, x_{h_2} \ldots x_{h_{n_h}})$ an (n, ξ) *Canonical Orthogonal Form (COF)* is given by the set of $n \times n$ matrices X_h, X_g, O, D and vectors $|v\rangle$, $|w\rangle$ where

$$X_h := \text{diag}\{x_{h_1}, x_{h_2} \ldots, x_{h_{n_h}}, \xi, \xi \ldots\},$$

$$X_g := \text{diag}\{x_{g_1}, x_{g_2} \ldots, x_{g_{n_g}}, 0, 0 \ldots\},$$

$$|v\rangle := \sum_{i=1}^{n_g} \sqrt{p_{g_i}} \, |i\rangle \,,$$

$$|w\rangle := \sum_{i=1}^{n_h} \sqrt{p_{h_i}} \, |i\rangle \,,$$

$$D := X_h - OX_gO^T$$

and the matrix O is orthogonal which satisfies

$$O \, |v\rangle = |w\rangle \,.$$

A ξ *Canonical Orthogonal Form (COF)* is an (n, ξ) COF with $n = n_h + n_g - 1$.

Definition 51 (n legal COF, legal COF)**.** An (n, ξ) COF is an n *legal COF* if $D \geq 0$ in the limit $\xi \to \infty$. A *legal COF* is a ξ COF such that $D \geq 0$ in the limit $\xi \to \infty$.

Imagine you found a legal COF corresponding to some transition. One can then sandwich D between a positive matrix as EDE to get

$$\left[\begin{array}{c|c} X_h & \\ \hline & \begin{array}{ccc} 1 & & \\ & \ddots & \\ & & 1 \end{array} \end{array}\right] - \underbrace{\left[\begin{array}{ccc|c} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \hline & & & \begin{array}{ccc} \xi^{-1/2} & & \\ & \ddots & \\ & & \xi^{-1/2} \end{array} \end{array}\right]}_{:=E} U X_g U^\dagger \left[\begin{array}{ccc|c} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \hline & & & \begin{array}{ccc} \xi^{-1/2} & & \\ & \ddots & \\ & & \xi^{-1/2} \end{array} \end{array}\right].$$

Note that $D \geq 0 \iff EDE \geq 0$ because E is diagonal (which means one can write $EDE = (E\sqrt{D})(\sqrt{D}E)$ which in turn is of the A^TA form). From the legality of the COF, $D \geq 0$ in the limit $\xi \to \infty$ and in this limit E becomes a projector. After some relabelling (and appropriately expanding the space to $m = n_g + n_h$ dimensions) the inequality reduces to a CPF. This observation readily extends to the n legal case where $n \leq n_g + n_h$. It turns out that one can, and we show this later, always express an n' legal COF as an n legal COF with $n \leq n_g + n_h$ (in fact we can prove that $n \leq n_g + n_h - 1$). We have established the following statement.

Proposition 52. *Consider a transition. If there exists an n legal COF corresponding to it then there exists a legal CPF for the said transition.*

How about the reverse? Given a legal CPF can we find the corresponding n legal COF? We are given

$$D = \left[\begin{array}{c|ccc} X_h & & & \\ \hline & 0 & & \\ & & \ddots & \\ & & & 0 \end{array}\right] - \underbrace{\left[\begin{array}{ccc|ccc} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ \hline & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{array}\right]}_{=E_h} U \left[\begin{array}{c|c} 0 & \\ \hline & X_g \end{array}\right] U^\dagger \left[\begin{array}{ccc|ccc} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ \hline & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{array}\right] \geq 0.$$

Replacing the appended diagonal zeros in the first matrix (one containing X_h) with 1s yields an equivalent inequality. Next note that we can conjugate by a permutation matrix to get

$$\left[\begin{array}{c|c} 0 & \\ \hline & X_g \end{array}\right] = \tilde{U} \left[\begin{array}{c|c} X_g & \\ \hline & 0 \end{array}\right] \tilde{U}.$$

Finally we write the diagonal zeros in E_h as $1/\xi^{1/2}$ and use the reverse of the trick above to recover an m legal COF where recall $m := n_g + n_h$. This sketches the proof of the following statement.

Proposition 53. *Consider a transition. If there exists legal CPF corresponding to it then there exists an m legal COF for the said transition (where recall $m := n_g + n_h$).*

5.2 From EBM to EBRM to COF

We briefly summarise, at the cost being redundant, how Aharonov et al. prove that valid functions are equivalent to the EBM functions (assuming the operator monotones are on/the spectrum of the matrices is in $[0, \Lambda]$). They do this by showing that the set of EBM functions forms a convex cone K. Then they take the dual of this cone to get K^*. *This dual happens to be the set of operator monotone functions.* Then they find the bi-dual K^{**} and define the objects in this to be valid functions. They then show that $K = K^{**}$ which is to say that valid functions are equivalent to EBM functions. Note that all of this is assuming the aforesaid $[0, \Lambda]$ condition.

This is an extremely useful result because checking if a function is EBM is hard. Checking if a function is valid is a piece of cake because mathematical wizards have neatly characterised the set of operator monotone functions.

One can do even better. Instead of EBM functions, consider EBRM functions where the matrices are additionally restricted to be real. Let this set be given by K'. It turns out that its dual K'^* is also the set of operator monotone functions [18] viz. $K'^* = K^*$. Aharonov et al's proof for

$K = K^{**}$ can be applied to the real case as is to get $K' = K^{**}$ (granted we assume the same $[0, \Lambda]$ condition).

Since Mochon's point games (and even the ones built later) use valid functions, the aforesaid simplification justifies why it suffices to restrict to real matrices.

We use the definition of Prob (Definition 8), EBM line transition (Definition 9), EBM function (Definition 14, Definition 15), Operator Monotone functions (Definition 22, Definition 23) and their characterisation (Lemma 25, Lemma 26), Λ valid functions (Definition 28) and finally its equivalence with EBM functions (Corollary 31).

Equivalence of EBM and EBRM

First we define EBRM transitions and EBRM functions similar to their EBM analogues except with the further restriction that the matrices and vectors involved are real.

Definition 54 (EBRM transitions)**.** Let $g, h : [0, \infty) \to [0, \infty)$ be two functions with finite supports. The transition $g \to h$ is EBRM if there exist two real matrices $0 \leq G \leq H$ and a (not necessarily normalised) vector $|\psi\rangle$ such that $g = \text{prob}[G, \psi]$ and $h = \text{prob}[H, \psi]$.

Definition 55 (K', EBRM functions; K'_Λ, EBRM functions on $[0, \Lambda]$)**.** A function $a : [0, \infty) \to \mathbb{R}$ with finite support is an EBRM function if the transition $a^- \to a^+$ is EBRM, where $a^+ : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $a^- : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ denote, respectively, the positive and the negative part of a ($a = a^+ - a^-$). We denote by K' the set of EBRM functions.

For any finite $\Lambda \in (0, \infty)$, a function $a : [0, \Lambda) \to \mathbb{R}$ with finite support is an EBRM function with support on $[0, \Lambda]$ if the transition $a^- \to a^+$ is EBRM with its spectrum in $[0, \Lambda]$, where a^+ and a^- denote, respectively, the positive and the negative part of a (again, $a = a^+ - a^-$). We denote by K'_Λ the set of EBRM functions with support on $[0, \Lambda]$.

Definition 56 (Real operator monotone functions)**.** A function $f : (0, \infty) \to \mathbb{R}$ is a real operator monotone if for all real matrices $0 \leq A \leq B$ we have $f(A) \leq f(B)$.

A function $f : (0, \Lambda) \to \mathbb{R}$ is a real operator monotone on $[0, \Lambda]$ if for all real matrices $0 \leq A \leq B$ with spectrum in $[0, \Lambda]$ we have $f(A) \leq f(B)$.

Lemma. *K'^*_Λ is the set of real operator monotones on $[0, \Lambda]$.*

Proof sketch. Aharonov et al. showed that K^*_Λ (which is, recall, the dual of the set of EBM functions on $[0, \Lambda]$) is the set of operator monotone functions on $[0, \Lambda]$ (see Lemma 3.9 of [6]). Their proof can be adapted here by restricting to real matrices which entails that K'^*_Λ is the set of real operator monotone functions on $[0, \Lambda]$. □

Lemma 57. *$K^*_\Lambda = K'^*_\Lambda$ and $K^* = K'^*$, i.e. the set of operator monotones on $[0, \Lambda]$ equals the set of real operator monotones on $[0, \Lambda]$ and the set of operator monotones equals the set of real operator monotones.*

Corollary 58. *$K'_\Lambda = K'^{**}_\Lambda = K^{**}_\Lambda = K_\Lambda$, i.e. the set of EBRM functions on $[0, \Lambda]$ = the set of Λ valid functions (dual of EBRM functions) = the set of Λ valid functions (dual of EBM functions) = the set of EBM functions on $[0, \Lambda]$.*

Corollary 59. *Any strictly valid function is EBRM.*

We now sketch the proof of Lemma 57. It is clear that the set of real operator monotones must contain the set of operator monotones because operator monotones are by definition required to

work in the restricted real case as well. The set of real operator monotones might be bigger but that does not happen to be the case. This is because we can encode an $n \times n$ hermitian matrix into a $2n \times 2n$ real symmetric matrix. This is achieved by replacing each complex number $\alpha + i\beta$ with the matrix

$$\alpha \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \beta \begin{bmatrix} & -1 \\ 1 & \end{bmatrix}.$$

Note that the matrices have the exact same properties as 1 and i respectively. This corresponds to (after some permutation) writing a complex matrix $W = W_{\Re} + iW_{\Im}$ as a real symmetric

$$W' = \begin{bmatrix} W_{\Re} & -W_{\Im} \\ W_{\Im} & W_{\Re} \end{bmatrix}$$

where $W_{\Re}$ and $W_{\Im}$ are real. For a Hermitian $W^\dagger = W$ we must have $W_{\Re} = W_{\Re}^T$ and $W_{\Im} = -W_{\Im}^T$ which makes $W' = W'^T$ a symmetric matrix. The spectra of W and W' are the same. This is established most easily by looking at the diagonal decomposition. $W = USU^\dagger$ which would get transformed to $W' = U'S'U'^\dagger$. Since S is real S' is also real with doubly degenerate eigenvalues (except for the degeneracy already present in S). Thus if we have $0 \leq A \leq B$ then we would also have $0 \leq A' \leq B'$ as $A - B$ and $A' - B'$ would have the same spectrum where we used A' and B' to represent real symmetric analogues of the hermitian matrices A and B. The other way is trivial. Hence we have an equivalence which means that requiring a function to be operator monotone on complex matrices is the same as requiring it to be operator monotone on real symmetric matrices of twice the size (at most). This means that the set of real operator monotones is the same as the set of operator monotones.

EBRM to COF | Mochon's Variant

We just saw how we can reduce our problem from the set of EBM transitions to the set of EBRM transitions. We now show that each EBRM transition can be written in a standard form, which we call the Canonical Orthogonal Form (COF). The following is actually due to Mochon/Kitaev [1] but there was a minor mistake in one of the arguments which we have corrected. The interesting part is showing that one can always restrict the matrices to a certain size which in turn depends on the number of points involved in the transition.

Lemma 60. *For every EBRM transition $g \to h$ with spectrum in $[a, b]$ there exists an orthogonal matrix O, diagonal matrices X_h, X_g (with no multiplicities except possibly those of a and b) of size at most $n_g + n_h - 1$ such that*

$$O \underbrace{\begin{bmatrix} x_{g_1} & & & & & \\ & \ddots & & & & \\ & & x_{g_{n_g}} & & & \\ & & & a & & \\ & & & & \ddots & \end{bmatrix}}_{:=X_g} O^T \leq \begin{bmatrix} x_{h_1} & & & & & \\ & \ddots & & & & \\ & & x_{h_{n_h}} & & & \\ & & & b & & \\ & & & & \ddots & \end{bmatrix} = X_h,$$

and the vector $|\psi\rangle := \sum_{i=1}^{n_h} \sqrt{p_{h_i}} |i\rangle = \sum_{i=1}^{n_g} \sqrt{p_{g_i}} O |i\rangle$.

Proof. An EBRM entails that we are given $G \leq H$ with their spectrum in $[a, b]$ and a $|\psi\rangle$ such that

$$g = \text{Prob}[G, |\psi\rangle] = \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}]$$

and

$$h = \text{Prob}[H, |\psi\rangle] = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}]$$

with $p_{g_i}, p_{h_i} > 0$ and $x_{g_i} \neq x_{g_j}$, $x_{h_i} \neq x_{h_j}$ for $i \neq j$ but the dimension and multiplicities can be arbitrary. First we show that one can always choose the eigenvectors $|g_i\rangle$ of G with eigenvalue x_{g_i} such that

$$|\psi\rangle = \sum_{i=1}^{n_g} \sqrt{p_{g_i}} |g_i\rangle .$$

Consider P_{g_i} to be the projector on the eigenspace with eigenvalue x_{g_i}. Note that

$$|g_i\rangle := \frac{P_{g_i} |\psi\rangle}{\sqrt{\langle\psi| P_{g_i} |\psi\rangle}}$$

fits the bill. Similarly we choose/define $|h_i\rangle$ so that

$$|\psi\rangle = \sum_{i=1}^{n_h} \sqrt{p_{h_i}} |h_i\rangle .$$

Consider now the projector onto the $\{|g_i\rangle\}$ space

$$\Pi_g = \sum_{i=1}^{n_g} |g_i\rangle \langle g_i| .$$

Note that this will not have all eigenvectors with eigenvalues $\in \{x_{g_i}\}$. Similarly we define

$$\Pi_h = \sum_{i=1}^{n_h} |h_i\rangle \langle h_i| .$$

We further define $G' := \Pi_g G \Pi_g + a(\mathbb{I} - \Pi_g)$ and $H' := \Pi_h H \Pi_h + b(\mathbb{I} - \Pi_h)$. These definitions are useful as we can show

$$G' \leq H'.$$

From $G = \Pi_g G \Pi_g + (\mathbb{I} - \Pi_g) G (\mathbb{I} - \Pi_g)$ we can conclude that $\Pi_g G \Pi_g + a(\mathbb{I} - \Pi_g) \leq G$. This entails $G' \leq G$. Using a similar argument one can also establish that $H \leq H'$. Combining these we get $G' \leq H'$.
Consider the projector

$$\Pi := \text{projector on span}\{\{|g_i\rangle\}_{i=1}^{n_g}, \{|h_i\rangle\}_{i=1}^{n_h}\}$$

and note that this has at most $n_g + n_h - 1$ dimension because $|\psi\rangle$ lives in the span of $\{|g_i\rangle\}$ and in the span of $\{|h_i\rangle\}$ so one of the basis vectors at least is not independent. Now note that

$$G'' := \Pi G' \Pi \leq \Pi H' \Pi =: H''$$

because we can always conjugate an inequality by a positive semi-definite matrix on both sides. Note also that $\Pi |\psi\rangle = |\psi\rangle$ which means the matrices and the vectors have the claimed dimension. We now establish that $\text{Prob}[H'', |\psi\rangle] = h$ and $\text{Prob}[G'', |\psi\rangle] = g$. For this we first write the projector tailored to the g basis as $\Pi = \Pi_g + \Pi_{g_\perp}$ where $\Pi_{g_\perp}$ is meant to enlarge the space to the $\text{span}\{h_i\}_{i=1}^{n_h}$. With this we evaluate

$$\begin{aligned} G'' &= (\Pi_g + \Pi_{g_\perp}) [\Pi_g G \Pi_g + a(\mathbb{I} - \Pi_g)] (\Pi_g + \Pi_{g_\perp}) \\ &= \Pi_g G \Pi_g + a \Pi_{g_\perp} . \end{aligned}$$

Manifestly then $\text{Prob}[G'', |\psi\rangle] = g$. By a similar argument one can establish the h claim. Note that that G'' and H'' have no multiplicities except possibly in a and b respectively. Thus we conclude we can always restrict to the claimed dimension and form. □

Corollary 61. *For every EBRM transition the corresponding COF is legal.*

The COF is of interest because we can use it to interpret our inequalities geometrically and use the tools thereof. We study this connection next.

6 Ellipsoids

6.1 The inequality as containment of ellipsoids

We try to show that the matrix inequality of the form $0 \leq G \leq H$ can be geometrically viewed as the containment of a smaller ellipsoid inside a larger one.

Consider an unnormalised vector $|u\rangle = \sum_j u_j |h_j\rangle$ with $u_j \in \mathbb{R}$. Note that the set

$$\{|u\rangle \,|\, \langle u| X_h |u\rangle = 1\}$$

describes the surface of an ellipsoid where $X_h = \text{diag}(x_{h_1}, x_{h_2} \dots)$. This is easy to see as the constraint corresponds to

$$x_{h_1} u_1^2 + x_{h_2} u_2^2 + \dots = 1$$

which is of the form

$$\frac{u_1^2}{a_1^2} + \frac{u_2^2}{a_2^2} + \dots = 1$$

which, in turn, is the equation of an ellipsoid in the variables $\{u_i\}$ with axes $a_1 = 1/\sqrt{x_{h_1}}, a_2 = 1/\sqrt{x_{h_2}} \dots$. An inequality would correspond to points inside or outside the ellipsoid. It is also useful to note that if we start with some arbitrary (even unnormalised) vector $|u\rangle$ then the point on the ellipse along this direction are given by

$$\mathcal{E}_h(|u\rangle) = \frac{|u\rangle}{\sqrt{\langle u| X_h |u\rangle}}.$$

Finally, note that the set $\left\{|u\rangle \,|\, \langle u| U X_g U^\dagger |u\rangle = 1\right\}$ also corresponds to the equation of an ellipsoid with axes $\left\{1/\sqrt{x_{g_i}}\right\}$ except that it is rotated. This follows from the fact that if we use $|u'\rangle = U |u\rangle$ then the equation reduces to the standard form in the u_i' variables which can then be used to obtain u_is by the aforesaid relations which is a rotation. We can define a similar map from a vector $|u\rangle$ to a point on the rotated ellipse as

$$\mathcal{E}_g(|u\rangle) = \frac{|u\rangle}{\sqrt{\langle u| U X_g U^\dagger |u\rangle}}.$$

With this understanding in place we are ready to get a visual interpretation of our equation. The statement that

$$\begin{aligned}
& X_h - U X_g U^\dagger \geq 0 \\
\iff & \langle u| X_h |u\rangle - \langle u| U X_g U^\dagger |u\rangle \geq 0 && \forall\, |u\rangle \\
\iff & \langle u| U X_g U^\dagger |u\rangle \leq 1 && \forall \{|u\rangle \,|\, \langle u| X_h |u\rangle = 1\}
\end{aligned}$$

which in turn corresponds to the statement that every point denoted by $|u\rangle$ that is on the h ellipse must be on or inside the g ellipse. Note that if $\langle x_h\rangle - \langle x_g\rangle = 0$ then for $|u\rangle = |w\rangle$ the inequality saturates. This in turn means that even for $\mathcal{E}_h(|w\rangle)$ the inequality is saturated as it is the same vector up to a scaling. The difference is that $\mathcal{E}_h(|w\rangle)$ represents a point on the h ellipsoid. Since the inequality is saturated it means that the ellipsoids must touch at this point. Thus $\mathcal{E}_g(|w\rangle) = \mathcal{E}_h(|w\rangle)$ which one can check explicitly as well.

The discussion so far can only give some intuition about the visualisation of our constraint equation. This intuition, as was explained in Subsection 1.3, can be efficiently used but it requires us to precisely specify the notion of curvature.

6.2 Convex Geometry Tools | Weingarten Map and the Support Function

Consider a curve in the plane. One can easily guess that the curvature must be related to the rate of change of tangents. This means we must use the second derivative. This can be generalised to arbitrary dimensions and in this case we obtain a matrix of the form $\partial_i\partial_j f$ for some function f which describes the curve. The eigenvalues of this matrix would tell us the curvature along the principle directions of curvature, given by the corresponding eigenvectors. It is possible to follow this idea through for an ellipsoid but the result becomes rather cumbersome as one must choose a coordinate system with its origin at the point of interest, aligned along the normal and re-express all the quantities of interest.

A concise way of evaluating the same is based on a mathematically sophisticated method applicable to all convex bodies. We state the result for the convex body of interest, an ellipsoid.

For a normalised direction vector $|u\rangle$ the support function corresponding to an ellipsoid X is given by

$$h(u) = \sqrt{\langle u|\, X^{-1}\, |u\rangle} = \sqrt{\sum x_i^{-1} u_i^2}. \tag{7}$$

The derivative of the support function yields the point on the ellipsoid where the tangent plane corresponding to the direction $|u\rangle$ touches the said ellipsoid. It is

$$\partial_i h(u) = \frac{x_i^{-1} u_i}{h(u)}.$$

The most remarkable of all these is the fact is that

$$h\partial_j\partial_i h(u) = \left(-\frac{x_j^{-1} x_i^{-1} u_i u_j}{h^2} + x_i^{-1}\delta_{ij}\right) \tag{8}$$

contains as eigenvalues the radii of curvature at the aforesaid point and as eigenvectors the directions of principle curvature. If instead of the normal you know the point at which you would like to evaluate this object then one can use the gradient to first find this normal and then apply the aforesaid. The normal at a point of contact $|c\rangle = \sum c_i\, |i\rangle$ is $|u(c)\rangle = \mathcal{N}\left(\sum x_i c_i\, |i\rangle\right)$. The results discussed here were deduced as special cases of those discussed in Section 2.5 of the book on convex bodies by R. Schneider [19].

We have stated the basic results needed to proceed with the description of our algorithm.

7 Elliptic Monotone Align (EMA) Algorithm

Solving the weak coin flipping (WCF) problem can be reduced to finding explicit matrices for a given EBM transition $g = \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}] \to h = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}]$ where g and h have disjoint support

or, equivalently, for a given EBM function $a = h - g = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}] - \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}]$. Here we describe our elliptic monotone align (EMA) algorithm, which runs by converting the given problem into the same problem of one less dimension iteratively until it is solved.

7.1 Notation

This subsection might appear to be particularly dry as we almost exclusively only introduce definitions; but it is a necessary evil. We try to motivate the definitions as we move along but things would not make perfect sense until one reaches the description and analysis of the algorithm itself.

At step k of the iteration, the transition $g \to h$ and the associated function $a = h - g$ used below are given by $g^{(k)} \to h^{(k)}$ and $a^{(k)}$ respectively. It remains fixed for the said step. We therefore do not write an explicit dependence on it in the following definitions. This is to facilitate the discussion of the iterative algorithm. We consider the extended real line $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$ with $1/\infty = -1/\infty := 0$. We also need the extended half line $\bar{\mathbb{R}}_{\geq} := \mathbb{R}_{\geq} \cup \{\infty\}$ and $\bar{\mathbb{R}}_{>} := \mathbb{R}_{>} \cup \{\infty\}$. These situations appear unavoidably in the analysis of certain transitions and correspond to one of the directions of the ellipsoids having infinite curvature. We use $\mathcal{N}(|\psi\rangle) := |\psi\rangle = \sqrt{\langle\psi|\psi\rangle}$. We usually denote by $[x_{\min}, x_{\max}]$ the smallest interval that contains $\mathrm{supp}(a)$. We call this interval the *support domain* for a. Similarly, we would refer to the smallest interval containing $\mathrm{supp}(g) \cup \mathrm{supp}(h)$ as the *transition support domain* for (the transition) $g \to h$. We use the variables $\chi, \xi \in \mathbb{R}$ to be such that they denote an interval $[\chi, \xi] \supseteq [x_{\min}, x_{\max}]$. As these χ and ξ would later be associated with an interval containing the spectrum of relevant matrices, we would refer to this interval as the *spectral domain*. The need for distinguishing the three is not hard to justify. A transition $g \to h$ can be such that both g and h have a term $p_k[x_k]$ for some $p_k > 0$ and x_k. This term would be absent from $a = h - g$. Thus the transition support domain and the support domain would be different in general. One might object to this as we started with the assumption that g and h have disjoint support. The issue is that this assumption does not necessarily hold once the problem is reduced to a smaller instance of itself. As for the spectral domain, this is defined from hindsight, as we know we need to use COFs which (see Lemma 60) use matrices with spectra that would usually be larger than the transition support domain.

Recall from the discussion of Subsection 1.3 that we intend to use operator monotone functions to make the ellipsoids touch along a known direction. We already have a characterisation of operator monotone functions (see Lemma 26). The function $\lambda x/(\lambda + x)$ can be turned into $-1/(\lambda + x)$ by adding a constant (we will do this carefully shortly). Further, the characterisation we have expects the input matrices to have their spectrum in the range $[0, \Lambda]$. We must generalise this as this assumption can not be made for smaller instances of the same problem which appear in subsequent iterations. This motivates the following definitions.

Definition 62 (f_λ on (α, β))**.** $f_\lambda : (\alpha, \beta) \to \mathbb{R}$ is defined for $\lambda \in \mathbb{R} \backslash [-\beta, -\alpha]$ as

$$f_\lambda(x) := \frac{-1}{\lambda + x}.$$

Definition 63 (f_λ on $[\alpha, \beta]$)**.** $f_\lambda : [\alpha, \beta] \to \bar{\mathbb{R}}$ is defined for $\lambda \in \mathbb{R} \backslash (-\beta, -\alpha)$. For $\lambda \in \mathbb{R} \backslash [-\beta, -\alpha]$ we define

$$f_\lambda(x) := \frac{-1}{\lambda + x}.$$

For $\lambda = -\beta$ and $-\alpha$ we retain the same definition as above except when $x = \beta$ and α respectively in which case we define

$$f_{-\beta}(\beta) := \infty$$
$$f_{-\alpha}(\alpha) := -\infty.$$

Remark 64. Values for $f_{-\beta}(\beta)$ and $f_{-\alpha}(\alpha)$ are obtained by taking for x, respectively, the left limit (approaching from the left to β) and right limit (approaching from the right to α). Also note that the operator monotone $f(x) = x$ is not included in the aforesaid family of functions.

We had to define f_λ on the two intervals for technical reasons which we can't quite motivate here. We explicitly defined f_λ to be ∞ or $-\infty$ where it would be otherwise undefined (division by zero). These infinities, however, will only appear in the denominator in our algorithm.

Again, from the discussion of Subsection 1.3, we recall that we have to expand the smaller ellipsoid until it touches the larger ellipsoid. From Subsection 6.1 we can see that the ellipsoid corresponding X_h, a positive diagonal matrix, is smaller than the one corresponding to γX_h for $0 < \gamma < 1$. The X_h matrix would correspond to a function h. What would the corresponding function be for γX_h? The following definition of h_γ formalises the answer. We also introduce l_γ which helps us check the validity condition for a transition (similar to Definition 28 and Corollary 31). The basic idea is to take the inner product (sum over the finite support of a) of the function a with a given operator monotone. If this is positive for every operator monotone, then the function a is valid. From hindsight, since we already know the characterisation of these operator monotones, we define $l_\gamma(\lambda)$ to be this inner product which must be positive, labelling the operator monotone by λ and encoding the stretching of the h ellipsoid into γ. This plays a crucial role in our algorithm as we have to make sure we use the right stretching, γ, without actually knowing the ellipsoids completely. We do not expect the details of these statements to be clear just yet but we hope the following definitions appear reasonable.

Definition 65 $(l_\gamma, l^1_\gamma, a_\gamma)$**.** Consider the transition $g \to h$ and let $a = h - g$. For $\gamma \in (0, 1]$ we define the finitely supported functions $h_\gamma : \mathbb{R} \to \mathbb{R}_\geq$ and $a_\gamma(x) : \mathbb{R} \to \mathbb{R}$ as

$$h_\gamma(x) := h(x/\gamma)$$
$$a_\gamma(x) := h_\gamma(x) - g(x).$$

Let $S_\gamma = [x_{\min}(\gamma), x_{\max}(\gamma)]$ be the support domain of a_γ. We define $l_\gamma : \mathbb{R}\backslash[-x_{\max}(\gamma), -x_{\min}(\gamma)] \to \mathbb{R}$

$$l_\gamma(\lambda) := \sum_{x \in \text{supp}(a_\gamma)} a_\gamma(x) f_\lambda(x)$$

where f_λ is defined on S_γ.

We define

$$l^1_\gamma := \sum_{x \in \text{supp}(a_\gamma)} a_\gamma(x) x.$$

Remark 66. h_γ and g might have overlapping support for certain values of γ which justifies the terminology distinguishing support domain and spectral support domain (introduced at the beginning of the section).

We now define a sort of indicator function, m, which tells us, given the transition $g \to h$, if the transition corresponding to the scaled ellipsoid $g \to h_\gamma$ is valid. There are some extra parameters this function needs. Consider the spectrum of the matrices which make this transition EBRM (they must be EBRM if they are valid, similar to Corollary 31). These parameters encode the interval in which this spectrum must be contained.

Definition 67 ($m(\gamma, \chi, \xi)$)**.** We define $m : ((0,1], \mathbb{R}, \mathbb{R}) \to \{0,1\}$ to be

$$m(\gamma, \chi, \xi) := \begin{cases} 0 & \text{if any of the following root conditions hold} \\ 1 & \text{else.} \end{cases}$$

where the first root condition is satisfied if there exists a $\lambda \in \mathbb{R} \backslash (-\xi, -\chi)$ such that $l_\gamma(\lambda) = 0$, and the second root condition is satisfied if $l_\gamma^1 = 0$.

As we are dealing with different representations of the same object, we define a relation between the matrix instance of the problem (which involves matrices) and the function instance thereof (which involves transitions and functions). The matrix instance contains all the information needed and so in the discussion of the algorithm we pack everything into a matrix instance to keep things palpable. The reader can glance through the following and later refer to them when they are used.

Definition 68 (Matrix Instance, $\underline{\mathrm{X}} \to$ Function Instance, $\underline{\mathrm{x}}$)**.** For a *Matrix Instance* defined to be the tuple $\underline{\mathrm{X}} := (X_h, X_g, |w\rangle, |v\rangle)$ where X_h, X_g are diagonal matrices and $|w\rangle, |v\rangle$ are vectors on $\mathbb{R}^n$ for some n with equal norm, i.e. $\langle w|w\rangle = \langle v|v\rangle$, we define the *Function Instance* to be the tuple $\underline{\mathrm{x}} : (g, h, a)$ where $h = \mathrm{Prob}[X_h, |w\rangle]$, $g = \mathrm{Prob}[X_g, |v\rangle]$ and $a = h - g$.

Definition 69 (Attributes of the Function Instance, $\underline{\mathrm{x}}$)**.** For a given tuple $\underline{\mathrm{x}} := (g, h, a)$ as Definition 68 we define the attributes $n_h, n_g, \{p_{g_i}\}, \{p_{h_i}\}, \{x_{g_i}\}, \{x_{h_i}\}$ as they appear by declaring $g \to h$ to be a transition, i.e.,

- n_h as the number of times h is non-zero,
- n_g as the number of times g is non-zero,
- $\{p_{h_i}\}, \{x_{h_i}\}, \{p_{g_i}\}, \{x_{g_i}\}$ implicitly as

$$h = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}],\ g = \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}]$$

  (for $p_{h_i}, p_{g_i} > 0$).

The *support domain* for a is denoted by $[x_{\min}, x_{\max}]$, i.e., the attributes $x_{\min}, x_{\max}$ are defined to be such that $[x_{\min}, x_{\max}]$ is the smallest interval containing $\mathrm{supp}(a)$.

Remark 70. Note that $x_{\min}$ and $x_{\max}$ may not be $x_{\min} = \min\{\{x_{h_i}\}, \{x_{g_i}\}\}$ and $x_{\max} = \max\{\{x_{h_i}\}, \{x_{g_i}\}\}$ respectively because there can be cancellations in the evaluation of $h - g = a$.

Definition 71 (Attributes of the Matrix Instance, $\underline{\mathrm{X}}$)**.** We associate the following with a matrix instance.

- *Spectral domain*: For a tuple $\underline{\mathrm{X}}$ as defined in Definition 68 we denote the *spectral domain* by $[\chi, \xi]$ where the attributes χ, ξ are such that $[\chi, \xi]$ is the smallest interval containing $\mathrm{spec}\{X_g \oplus X_h\}$.

- *Solution*: We say that O is a *solution* to the matrix instance $\underline{\mathrm{X}} = (X_h, X_g, |w\rangle, |v\rangle)$ if $X_h \geq OX_gO^T$ and $O|v\rangle = |w\rangle$.
- *Notation*: With respect to a standard orthonormal basis $\{|i\rangle\}$, we use the *notation* $X_h := \sum_{i=1}^{k} y_{h_i} |i\rangle\langle i|$, $X_g := \sum_{i=1}^{k} y_{g_i} |i\rangle\langle i|$, $|w\rangle := \sum_{i=1}^{k} \sqrt{q_{h_i}} |i\rangle$, $|v\rangle := \sum_{i=1}^{k} \sqrt{q_{g_i}} |i\rangle$.

Remark 72. We index the Matrix Instance and the corresponding function instance as $\underline{\mathrm{X}}^{(k)} = \left(X_h^{(k)}, X_g^{(k)}, \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle\right)$ and $\underline{\mathrm{X}}^{(k)} \to \underline{\mathrm{x}}^{(k)} = \left(h^{(k)}, g^{(k)}, a^{(k)}\right)$ respectively. The associated attributes are implicitly assumed to be correspondingly indexed, e.g., as $\chi^{(k)}, \xi^{(k)}$ and $n_h^{(k)}, n_g^{(k)}, x_{\min}^{(k)}, x_{\max}^{(k)}$.

Remark 73. We introduce two different symbol sets p, x and q, y as it allows us to describe the proof more neatly by allowing two ways of indexing the same object. We use p, x for $\underline{\mathrm{x}}$ and q, y for $\underline{\mathrm{X}}$ which are essentially the same.

7.2 Lemmas for EMA

With the notation in place, we can now state and prove some results which we would need in our algorithm. We do this in three steps. First, we generalise the results obtained by Aharonov et al. about operator monotones and their relation with EBM functions. This is the workhorse of our algorithm. Second, we prove some results which formalise our intuitive notion of tightening—stretching the smaller h ellipsoid until it touches the larger g ellipsoid. Finally, we prove a generalisation thereof in the case where the curvature of the smaller h ellipsoid becomes infinite.

For a first reading, it might be better to focus on the statements, and come back to the proofs after reading the algorithm.

7.2.1 Generalisations

Keep the bigger picture, Figure 9, in mind to retain a sense of direction. Our main objective here would be twofold. First, we wish to generalise Corollary 58 from being restricted to matrices with their spectrum in $[0, \Lambda]$ to being applicable for matrices with their spectrum in $[\chi, \xi]$. Second, we wish to extend the result from valid functions to valid transitions, including the case of overlapping support.

To establish the first, our strategy would be to find a relation between $[0, \Lambda]$ valid functions and $[\chi, \xi]$ valid functions (which we will define carefully soon) and then a relation between $[0, \Lambda]$ EBRM functions and $[\chi, \xi]$ EBRM functions. Then we use the link between $[0, \Lambda]$ valid and $[0, \Lambda]$ EBRM functions to establish the equivalence of $[\chi, \xi]$ valid functions and $[\chi, \xi]$ EBRM functions. Along the way we sharpen our understanding of operator monotone functions which should make the definitions of f_λ, l and m (see Definition 65 and Definition 67) obvious. The second objective can be met with by a single, albeit, slightly long argument.

Let us start with extending our definition of the Canonical Orthogonal Form to accommodate matrices with their spectrum in $[\chi, \xi]$.

Definition 74 (Canonical Orthogonal Form (COF) with spectrum in $[\chi, \xi]$)**.** For a given transition $g \to h$ let $[\chi, \xi]$ be such that it contains $\mathrm{supp}(g)$ and $\mathrm{supp}(h)$. We define the Canonical Orthogonal Form (COF) with its spectrum in $[\chi, \xi]$ by the set of $n \times n$ matrices X_h, X_g, O, D and vectors $|v\rangle, |w\rangle$ where

$$X_h := \mathrm{diag}\{x_{h_1}, x_{h_2} \ldots, x_{h_{n_h}}, \xi, \xi \ldots\},$$

$$X_g := \mathrm{diag}\{x_{g_1}, x_{g_2} \ldots, x_{g_{n_g}}, \chi, \chi \ldots\},$$

$$|v\rangle := \sum_{i=1}^{n_g} \sqrt{p_{g_i}} \, |i\rangle \,,$$

$$|w\rangle := \sum_{i=1}^{n_h} \sqrt{p_{h_i}} \, |i\rangle \,,$$

$$D := X_h - O X_g O^\dagger,$$

the matrix O is orthogonal which satisfies

$$|v\rangle = O \, |w\rangle$$

and $n = n_g + n_h - 1$.

Definition 75 (Legal COF with spectrum in $[\chi, \xi]$)**.** A COF with spectrum in $[\chi, \xi]$ is legal if $D \geq 0$.

We obviously need to generalise the notion of operator monotone functions to the range $[\chi, \xi]$ as well to achieve our first objective.

Definition 76 (Operator monotone functions on $[\chi, \xi]$)**.** A function $f : [\chi, \xi] \to \mathbb{R}$ is operator monotone on $[\chi, \xi]$ if for all real symmetric matrices H, G with $\text{spec}(H \oplus G) \in [\chi, \xi]$ and $H \geq G$ we have $f(H) \geq f(G)$.

What happens if we try to shift/translate the interval on which an operator monotone is defined? This is a natural question to ask, an answer to which would also directly relate our new definition to the previous one.

Claim 77. $f(x)$ is an operator monotone function on $[\chi, \xi]$ if and only if $f'(x') = f(x' - x_0)$ is an operator monotone function on $[\chi + x_0, \xi + x_0]$.

Proof. Consider real symmetric matrices $H \geq G$ with $\text{spec}(H \oplus G) \in [\chi, \xi]$ and let $f(x)$ be operator monotone on $[\chi, \xi]$. We must consider $f'(x') = f(x = x' - x_0)$ which is the same as $f'(x + x_0) = f(x)$. We show that f' is an operator monotone on $[\chi + x_0, \xi + x_0]$. Note that $H' := H + x_0 \mathbb{I}$ and $G' := G + x_0 \mathbb{I}$ are such that $H' \geq G'$ and $\text{spec}(H' \oplus G') \in [\chi + x_0, \xi + x_0]$. Note that $f'(H') = f(H)$ and $f'(G') = f(G)$ because

$$\begin{aligned} f'(H') &= f'(H + x_0 \mathbb{I}) \\ &= O_h f'(H_d + x_0 \mathbb{I}) O_h^T \\ &= O_h f(H_d) O_h^T \\ &= f(H) \end{aligned}$$

and similarly for G where $H = O_h H_d O_h^T$ for O_h orthogonal and H_d diagonal. Since f is operator monotone on $[\chi, \xi]$ we have $f(H) \geq f(G)$ which entails $f'(H') \geq f'(G')$. Since this holds for all H', G' with their $\text{spec}(H' \oplus G') \in [\chi + x_0, \xi + x_0]$ we can conclude that f' is an operator monotone on $[\chi + x_0, \xi + x_0]$. The other way follows by setting $\chi + x_0$ to χ, $\xi + x_0$ to ξ, x_0 to $-x_0$ but since all these were arbitrary to start with, the reasoning goes through unchanged. □

We now note that from the characterisation of operator monotone functions we initially had (see Lemma 26), we can construct one which is easier to shift/translate (in the aforesaid sense).

Corollary 78 (Characterisation of operator monotone functions on $[0,\Lambda]$)**.** *Any operator monotone function $f:[0,\Lambda]\to\mathbb{R}$ can be written as*

$$f(x)=c_0+c_1x-\int\frac{1}{\lambda+x}d\tilde{\omega}(\lambda)$$

with the integral ranging over $\lambda\in(-\infty,-\Lambda)\cup(0,\infty)$ satisfying $\int\frac{1}{\lambda(1+\lambda)}d\tilde{\omega}(\lambda)<\infty$.

Proof. Consider the characterisation given in Lemma 26 according to which we had $f(x)=c_0'+c_1x+\int\frac{\lambda x}{\lambda+x}d\omega(\lambda)$ with $\int\frac{\lambda}{1+\lambda}d\omega(\lambda)<\infty$. We can write

$$\begin{aligned}f(x)&=c_0'+c_1x+\int\left(\lambda-\frac{\lambda^2}{\lambda+x}\right)d\omega(\lambda)\\&=c_0+c_1x-\int\frac{\lambda^2d\omega(\lambda)}{\lambda+x}\end{aligned}$$

where with $d\tilde{\omega}=\lambda^2d\omega(\lambda)$ we obtain the claimed form. Note that the finiteness of $\int\frac{\lambda}{1+\lambda}d\omega$ is necessary to conclude that $c_0=c_0'+\int\frac{\lambda}{1+\lambda}d\omega$ is also finite. □

Note that this form also makes it easier for us to handle any divergences as there is only the denominator one has to deal with.

This can now be shifted/translated to allow for a characterisation of our shifted/translated operator monotones.

Corollary 79 (Characterisation of operator monotone functions on $[\chi,\xi]$)**.** *Any operator monotone function $f':[\chi,\xi]\to\mathbb{R}$ can be written as*

$$f'(x')=c_0'+c_1'x'-\int\frac{1}{\lambda'+x'}d\tilde{\omega}'(\lambda')$$

with the integral ranging over $\lambda'\in(-\infty,-\xi)\cup(-\chi,\infty)$ satisfying $\int\frac{1}{(\lambda'+\chi)(1+\lambda'+\chi)}d\tilde{\omega}'(\lambda')<\infty$.

Proof. We follow the convention that $x'\in[\chi,\xi]$ while the unprimed $x\in[0,\xi-\chi]$. From Claim 77 we know that $f(x)$ is operator monotone on $[0,\xi-\chi]$ if and only if $f'(x')=f(x'-\chi)$ is operator monotone on $[\chi,\xi]$ where $x'=x+\chi$. Since we already have a characterisation for $f(x)$ we can characterise $f'(x')$ as $f(x'-\chi)$. From Corollary 78 we have

$$\begin{aligned}f'(x')&=c_0+c_1(x'-\chi)-\int\frac{d\tilde{\omega}(\lambda)}{\lambda+x'-\chi}\\&=c_0'+c_1x'-\int\frac{d\tilde{\omega}'(\lambda')}{\lambda'+x'}\end{aligned}$$

where $\lambda'=\lambda-\chi$. Since we had $\lambda\in(-\infty,-(\xi-\chi))\cup(0,\infty)$ it entails $\lambda'\in(-\infty,-\xi)\cup(-\chi,\infty)$. The condition on the integral $\int\frac{d\tilde{\omega}(\lambda)}{\lambda(\lambda+x)}<\infty$ can be expressed in terms of λ' as $\int\frac{d\tilde{\omega}'(\lambda')}{(\lambda'+\chi)(1+\lambda'+\chi)}<\infty$ with $d\tilde{\omega}'(\lambda')=d\tilde{\omega}(\lambda'+\chi)$. With $c_1=c_1'$ and $c_0'=c_0-c_1\chi$ we obtain the claimed form. □

We now generalise Definition 28 to (χ,ξ) valid functions (we intended (χ,ξ) to indicate a tuple and not an open set so do not read too much into this notation).

Definition 80 ((χ,ξ) valid function)**.** A finitely supported function $a:\mathbb{R}\to\mathbb{R}$ with $\mathrm{supp}(a)\in[\chi,\xi]$ is (χ,ξ) valid if for every operator monotone function f on $[\chi,\xi]$ we have $\sum_{x\in\mathrm{supp}(a)}a(x)f(x)\geq 0$.

Remark 81. Since in Corollary 79 $d\tilde{\omega}'$ is a measure, to establish (χ,ξ) validity of functions, it would suffice to restrict our attention to operator monotones $f'(x') = x'$, $f'(x') = -\frac{1}{\lambda'+x'}$ with $x' \in [\chi,\xi]$, $\lambda' \in (-\infty,-\xi)\cup(-\chi,\infty)$.

By shifting/translating the characterisation of operator monotone functions we can shift/translate valid functions as well.

Corollary 82 ($a(x)$ is (χ,ξ) valid $\iff$ $a(x'-x_0)$ is $(\chi+x_0,\xi+x_0)$ valid)**.** *A finitely supported function $a:\mathbb{R}\to\mathbb{R}$ with $supp(a)\in[\chi,\xi]$ is (χ,ξ) valid if and only if the function $a'(x') := a(x'-x_0):\mathbb{R}_{\geq}\to\mathbb{R}$ is $(\chi-x_0,\xi-x_0)$ valid.*

Proof. a is (χ,ξ) valid entails $\sum_{x\in supp(a)} a(x)f(x)\geq 0$ for all f operator monotone on $[\chi,\xi]$. We can write the sum as $\sum a(x'-x_0)f(x'-x_0)\geq 0$. Using Claim 77 we note that $f'(x') = f(x'-x_0)$ is operator monotone on $[\chi+x_0,\xi+x_0]$. For $a'(x') = a(x'-x_0)$ we thus have $\sum a'(x')f'(x')\geq 0$ which means $a'(x')$ is a $(\chi+x_0,\xi+x_0)$ valid function. The other way follows similarly. □

In accordance with our strategy, we have established a relation between $(0,\Lambda)$ valid functions and (χ,ξ) valid functions (in fact we have a more general result). We now proceed with establishing its analogue for EBRM functions.

Definition 83 (EBRM on $[\chi,\xi]$)**.** A finitely supported function $a:\mathbb{R}\to\mathbb{R}$ is EBRM on $[\chi,\xi]$ if there exist real symmetric matrices $H\geq G$ with their spectrum in $[\chi,\xi]$ and a vector $|w\rangle$ such that $a = \mathrm{Prob}[H,|w\rangle] - \mathrm{Prob}[G,|w\rangle]$.

Corollary 84 ($a(x)$ is EBRM on $[\chi,\xi]$ $\iff$ $a(x+\chi)$ is EBRM on $[0,\xi-\chi]$)**.** *A finitely supported function $a:\mathbb{R}\to\mathbb{R}$ with $supp(a)\in[\chi,\xi]$ is EBRM on $[\chi,\xi]$ if and only if the function $a'(x) := (x+\chi):\mathbb{R}_{\geq}\to\mathbb{R}$ is EBRM on $[0,\xi-\chi]$.*

Proof. If a is EBRM on $[\chi,\xi]$ it follows that there exist real symmetric matrices with $H\geq G$ and a vector $|w\rangle$ such that $\mathrm{spec}[H\oplus G]\in[\chi,\xi]$ and $a = \mathrm{Prob}[H,|w\rangle] - \mathrm{Prob}[G,|w\rangle]$. Clearly, $H' := H-\chi\mathbb{I}\geq G-\chi\mathbb{I} =: G'$ and $a'(x) = \mathrm{Prob}[H',|w\rangle] - \mathrm{Prob}[G',|w\rangle] = a(x+\chi)$ with $\mathrm{spec}[H'\oplus G']\in[0,\xi-\chi]$. This means a' is EBRM on $[0,\xi-\chi]$. The other way follows similarly. □

We have done all the hard work for meeting the first objective. We now simply combine our results so far to prove the desired equivalence.

Lemma 85 ($a(x)$ is (χ,ξ) valid function $\iff$ $a(x)$ is EBRM on $[\chi,\xi]$)**.** *A finitely supported function $a:\mathbb{R}\to\mathbb{R}$ with $supp(a)\in[\chi,\xi]$ being (χ,ξ) valid is equivalent to it being EBRM on $[\chi,\xi]$.*

Proof. From Corollary 82 we know that $a(x)$ being (χ,ξ) valid is equivalent to $a(x+\chi)$ being $\Lambda = \xi-\chi$ valid. From Corollary 58 we know that $a(x+\chi)$ is equivalently EBRM on $[0,\xi-\chi]$. Finally using Corollary 84 we know that $a(x+\chi)$ being EBRM on $[0,\xi-\chi]$ is equivalent to $a(x)$ being EBRM on $[\chi,\xi]$. □

The second objective will be achieved in a single shot.

Lemma 86 (EBRM function $\iff$ EBRM transition even with common support)**.** *If we write an EBRM function a with spectrum in $[\chi',\xi']$ as $a = h-g$ with $h,g:\mathbb{R}_{\geq}\to\mathbb{R}_{\geq}$ which may have common support then $g\to h$ is an EBRM transition with spectrum in $[\chi,\xi]$ and with (the smallest) matrix size (at most) n_g+n_h-1 where $[\chi,\xi]$ is the smallest interval containing $[\chi',\xi']$ and $supp(h)\cup supp(g)$.*

Conversely, if $g\to h$ is an EBRM transition with spectrum in $[\chi,\xi]$ with $h,g:\mathbb{R}_{\geq}\to\mathbb{R}_{\geq}$ which may have common support then $a = h-g$ is an EBRM function with its spectrum in $[\chi,\xi]$ (the smallest) matrix size at most n_g+n_h-1.

Proof. To prove the first statement we write $a = a^+ - a^-$ where $a^+ = \sum_{i=1}^{n'_h} p'_{h_i}[x_{h_i}], a^- = \sum_{i=1}^{n'_g} p'_{g_i}[x_{g_i}]$, for $a^+, a^- : \mathbb{R}_{\geq} \to \mathbb{R}_{\geq}$, represent the positive and the negative parts of a. Note that a^+ and a^- by virtue of this definition can't have any common support. Consider $\Delta = \sum_{i=1}^{n_\Delta} c_i[x_i] : \mathbb{R}_{\geq} \to \mathbb{R}_{\geq}$ to be such that $h = a^+ + \Delta$ and $g = a^- + \Delta$. This is always the case because $h - g = a$. Consider the case where $\text{supp}(\Delta) \cap \text{supp}(a) = \emptyset$. In this case $n_g = n'_g + n_\Delta$ and $n_h = n'_h + n_\Delta$. Since a is an EBRM function we have a legal COF, viz $O' X'_g O'^T \leq X'_h$ and $|w'\rangle = O' |v'\rangle$, of dimension $(n' = n'_g + n'_h - 1)$ from Lemma 60. To obtain the matrices corresponding to $g \to h$ we expand the space to $n = n_g + n_h - 1$ dimensions and define $X_g = X'_g \oplus X$, $X_h = X'_h \oplus X$, $O = O' \oplus \mathbb{I}$, $|v\rangle = |v'\rangle + \sum_{i=n'}^{n} \sqrt{c_{i+1-n'}} |i\rangle$ where $X = \text{diag}\{x_1, x_2 \dots x_{n_\Delta}\}$. This is just an elaborate way of adding the points in Δ to the matrices and the vectors in such a way that the part corresponding to Δ remains unchanged. The other cases can be similarly demonstrated with the only difference being in the relation between n_g, n'_g and n_h, n'_h. Suppose Δ is non-zero only at one point. If Δ adds a point where a^- had a point then it does not contribute to increasing the number of points in g that is $n_g = n'_g$ but it does increase the number in h that is $n_h = n'_h + 1$. This means that we have one extra dimension to find the matrices certifying $g \to h$ is EBRM which is precisely what is needed to append that extra idle point as described above. Similarly one can reason for adding a point where a^+ had a point and finally extend it to the most general case of $\text{supp}(\Delta) \cap \text{supp}(a) \neq \emptyset$ which may involve multiple points.

We now prove the converse. Since $g \to h$ is an EBRM transition from Lemma 60 we know that it admits a legal COF, that is $O X_g O^T \leq X_h$ and $O |v\rangle = |w\rangle$ with dimension $n_g + n_h - 1$. To be able to show that $a = h - g = a^+ - a^-$ (where a^+ and a^- are again the positive and negative part of a) is an EBRM function it suffices to show that a is a valid function. This follows directly from the COF and operator monotones as $O f(X_g) O^T \leq f(X_h)$ implies $\langle v| f(X_g) |v\rangle \leq \langle w| f(X_h) |w\rangle$ which in turn is $\sum h(x) f(x) - \sum g(x) f(x) \geq 0$ and that is the same as $\sum a(x) f(x) \geq 0$ for all f operator monotone on the spectrum of $X_h \oplus X_g$, viz. a is valid. From Lemma 85 we conclude that a is also EBRM with size at most $n_g + n_h - 1$ (actually we can make a stronger statement by saying the size should be at most $n'_g + n'_h - 1$ where $|\text{supp}(a^+)| = n'_h$ and $|\text{supp}(a^-)| = n'_g$). □

This completes the first, and longest, step of our groundwork for discussing the algorithm. Our achievement so far has been schematized in Figure 9.

7.2.2 For the finite part

For the second step, we state the following fact (see [17]). The proof of this statement is interesting in its own right but here we only state it and use it for terminating our recursive algorithm.

Fact 87 (Weyl's Monotonicity Theorem)**.** *If H is positive semi-definite and A is Hermitian then $\lambda^{\downarrow}_j(A + H) \geq \lambda^{\downarrow}_j(A)$ for all j where $\lambda^{\downarrow}_j(M)$ represents the j^{th} largest eigenvalue of the Hermitian matrix M.*

Corollary 88. *If $H \geq G$ then $\lambda^{\downarrow}_j(H) \geq \lambda^{\downarrow}_j(G)$ for all j.*

At some point, if the algorithm reaches a point where there is no vector constraint (that is the vector $|v\rangle = |w\rangle = 0$) then one can use the aforesaid result to conclude that the solution, orthogonal matrix O, must be a permutation matrix (this will become clear later, we mention it to motivate the relevance of the result).

We now state a continuity condition which we subsequently use to establish that when we stretch the h ellipsoid, there would always exist the perfect amount of stretching that makes the h ellipsoid just touch the g ellipsoid. The non-triviality here is that we have to conclude this without fully knowing the ellipsoids.

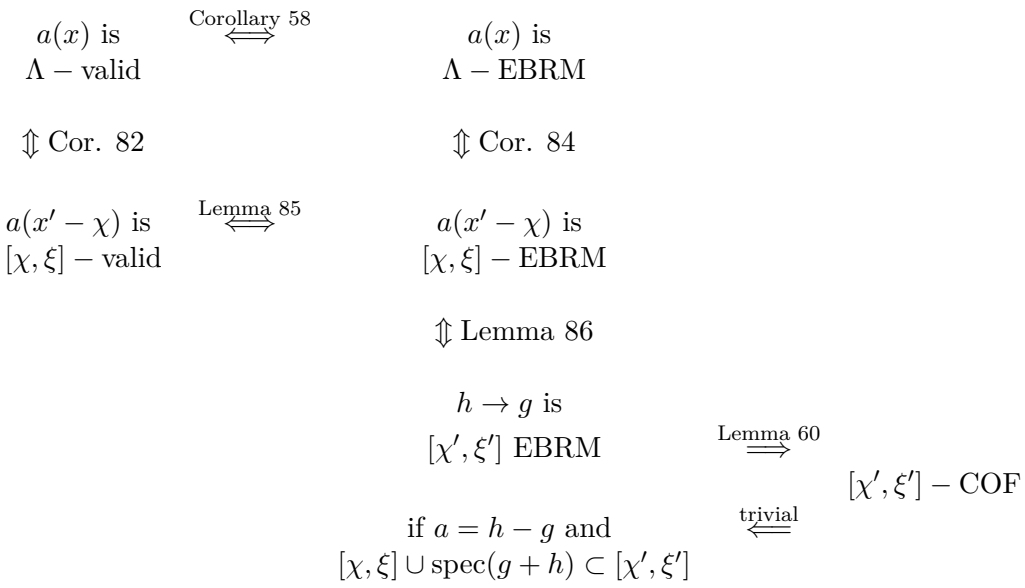


Figure 9: Generalisation schematized.

Claim 89 (Continuity of l). Let $[x_{\min}, x_{\max}]$ be the smallest interval containing $\mathrm{supp}(a)$. $l(\lambda)$ is continuous in the intervals $\lambda \in (-x_{\min}, \infty]$ and $\lambda \in [-\infty, -x_{\max})$ (see Definition 65).

Proof. Since $l(\lambda)$ is just a rational function of λ it suffices to show that the denominator doesn't become zero in the said range. The roots of the denominator are of the form $\lambda + x = 0$ for $x \in \{\{x_{g_i}\}, \{x_{h_i}\}\}$. Hence the largest root will be $\lambda = -x_{\min}$ and the smallest $\lambda = -x_{\max}$. Neither of the intervals defined in the statement contain any roots and therefore we can conclude that $l(\lambda)$ will be continuous therein. Note that the function f_λ on $[x_{\min}, x_{\max}]$ is not even defined for λ in $(-x_{\max}, -x_{\min})$. □

Lemma 90 (Tightening with the matrix spectrum unknown)**.** *Consider a finitely supported valid function a. Let $[x_{\min}(\gamma), x_{\max}(\gamma)]$ be the smallest interval containing $supp(a_\gamma)$. Consider $m(\gamma, x_{\min}(\gamma), x_{\max}(\gamma))$ as a function of γ (see Definition 67). m has at least one root in the interval* $(0, 1]$.

Proof. To prove the claim it suffices to show that $l_\gamma(\lambda)$ has a root in the range $(0, \infty)$ for some $\gamma \in (0, 1]$. Note that we are given a valid function a which means $\mathrm{supp}(a) \in \mathbb{R}_{\geq}$.

We assume that $l_{\gamma=1}(\lambda) > 0$ for all $\lambda \in (0, \infty)$ because if this was not the case then we trivially have $\gamma = 1$ as a root, i.e. $m(1, x_{\min}(1), x_{\max}(1)) = 0$.

Notice that since $\sum h(x) = \sum g(x)$ we have

$$\begin{aligned}\lambda l(\lambda) &= \sum h(x)(\lambda f_\lambda(x) + 1) - \sum g(x)(\lambda f_\lambda(x) + 1) \\ &= \sum h(x)\frac{x}{\lambda + x} - \sum g(x)\frac{x}{\lambda + x}.\end{aligned}$$

Therefore for the remainder of this proof we redefine $f_\lambda = \frac{1}{\lambda}\frac{x}{\lambda+x}$ without changing the value of l or by extension l_γ (the $1/\lambda$ factor is partly why we restricted λ to $(0, \infty)$ instead of the more general $(-x_{\min}, \infty)$). Note that $\lim_{\gamma\to 0^+} l_\gamma(\lambda) < 0$ for all $\lambda \in (0, \infty)$ because $h_\gamma(x) = h(x/\gamma)$ which means $\lim_{\gamma\to 0} \sum h_\gamma(x) f_\lambda(x) = \lim_{\gamma\to 0} \sum h(x) f_\lambda(\gamma x) = 0$ since $\lim_{x\to 0} f_\lambda(x) = 0$. This in turn means $\lim_{\gamma\to 0^+} l_\gamma(\lambda) = -\sum g(x) f_\lambda(x) < 0$.

Further, each term constituting $l_\gamma(\lambda)$ is finite for $\lambda \in (0, \infty)$ since for $\lambda > 0$ the denominators are of the form $\lambda + x$ which are always positive. Hence $l_\gamma(\lambda)$ as a function of $\lambda \in [0, \infty)$ and $\gamma \in (0, 1]$ is continuous. By continuity then between $\gamma = 0^+$ and $\gamma = 1$ there should be a root.

It remains to justify why we extended the range of λ from $(0, \infty)$ to $(-\infty, -x_{\max}) \cup (-x_{\min}, \infty)$ in the definition of m (see Definition 67) as it appears in the statement of the lemma. This is due to the fact that $l_\gamma(\lambda)$ is continuous for λ in the stated range, see Lemma 89, and so there might be a root which appears in the extended range. If this is the case we would like to use this possibly higher value of γ (because for a small enough value a non-negative root must appear due to the aforesaid reasoning). This can help us avoid infinities (we will explain this later). □

Once we are guaranteed that there is at least one perfect stretching amount, we want to know the spectrum of the matrices. We state a slightly more general result which is a direct consequence of the results from the previous subsubsection. The difference is that it is stated in a form that would be useful for the algorithm.

Lemma 91 (Matrix spectrum from a valid function)**.** *Consider a valid function a, i.e. an a such that $l(\lambda) \geq 0$ and $l^1 \geq 0$ for $\lambda \in [0, \infty)$ (see Definition 65) and let $[\chi, \xi]$ be such that for $\lambda \in [-\infty, -\xi) \cup (-\chi, \infty]$ we have $l(\lambda) \geq 0$.*

There exists a legal COF, corresponding to the function a, with its spectrum contained in $[\chi, \xi]$.

Proof. Since $l(\lambda) \geq 0$ for $\lambda \in (-\infty, -\xi) \cup (-\chi, \infty)$ and $l^1 \geq 0$ we know from Corollary 79 that a is (χ, ξ) valid. From Lemma 85 we know that a is EBRM on $[\chi, \xi]$. Finally from Lemma 60 we know that there exists a legal COF with spectrum in $[\chi, \xi]$. □

Recall that in Subsection 1.3 we said that we focus on operator monotone functions f with the special property that f^{-1} is also an operator monotone. We now establish that f_λs (see Definition 63) have this property.

Lemma 92 ($H \geq G \iff f_\lambda(H) \geq f_\lambda(G)$)**.** *Let H, G be real symmetric matrices and $[\chi, \xi]$ be the smallest interval containing spec$[H \oplus G]$ and f_λ be on (χ, ξ). $H \geq G$ if and only if $f_\lambda(H) \geq f_\lambda(G)$.*

Proof. $H \geq G \implies f_\lambda(H) \geq f_\lambda(G)$ because f_λ is an operator monotone function for matrices with spectrum in $[\chi, \xi]$. We prove the converse. We find the inverse function of f_λ and show that it is also an operator monotone. Start with recalling that for $x \in [\chi, \xi]$ we have

$$y = f_\lambda(x) = \frac{-1}{\lambda + x} \implies x = -\frac{1}{y} - \lambda$$

where $\lambda \in \mathbb{R} \backslash [\chi, \xi]$. Thus $f_\lambda^{-1}(y) = -\frac{1}{y} - \lambda$. For a given λ either $f_\lambda(\chi)$ and $f_\lambda(\xi)$ are both greater than zero or both less than zero. Hence the operator monotones $f'_{\lambda'}(y)$ on $[f_\lambda(\chi), f_\lambda(\xi)]$ permit $\lambda' = 0$. Consequently $f'_{\lambda'=0}(y) = \frac{-1}{y}$ is an operator monotone on $[f_\lambda(\chi), f_\lambda(\xi)]$. A constant is also an operator monotone. Thus we conclude $f_\lambda^{-1}(y)$ is an operator monotone on the required interval establishing the converse. □

This completes the second step of the groundwork.

7.2.3 For Wiggle-v; the infinite part

For the final step of the groundwork, we need to establish a result which lets us tackle the divergences head-on. What is this divergence issue? Recall form our discussion in Subsection 1.3, our strategy was to tighten (we saw hints of how that can be done in the previous sections) and then find the operator monotone f_λ for which the ellipsoids touch along the $|w\rangle$ direction. What happens if one of the ellipsoids under the action of this operator monotone has infinite curvature along some directions? One can in fact show that there are cases where this would necessarily happen. Having infinite curvature means that the corresponding matrix has a divergence. It is like an ellipse gets mapped to a line. Our algorithm will fail in this situation because the normal at the tip of a line is ill defined.

Our strategy is to show that tightness is preserved under the action of f_λ. This means that if for some λ' we consider the ellipsoids obtained by applying $f_{\lambda'}$ and we find that they touch along $|w\rangle$ then for some other λ'' ($\neq \lambda'$) they would continue to touch but along some other direction. This allows us to consider the sequence leading to the divergence. We use this sequence in the analysis of the algorithm.

Here we start by showing this result in the case where everything is well defined and then extend it to the divergent case.

Lemma 93 (Strict inequality under f_λ)**.** *$H > G$ if and only if $f_\lambda(H) > f_\lambda(G)$ where f_λ is on $(\chi, \xi) \supset spec[H \oplus G]$.*

Proof. Note that $H > G \iff H' := H + \lambda\mathbb{I} > G + \lambda\mathbb{I} =: G'$ where $\lambda \in \bar{\mathbb{R}}\backslash[-\xi, -\chi]$ (by definition of f_λ on (χ, ξ)). There can be two cases, either both the matrices are strictly positive or both are strictly negative. Let us assume the former (the other follows similarly). We have

$$\begin{aligned} & H' > G' > 0 \\ \iff & \mathbb{I} > H'^{-1/2} G' H'^{-1/2} \\ \iff & \mathbb{I} < H'^{1/2} G'^{-1} H'^{1/2} \\ \iff & H'^{-1} < G'^{-1} \end{aligned}$$

where the first inequality follows from the fact that multiplication by a positive matrix doesn't affect the inequality (hint: it only changes the vectors $|w\rangle$ we use to show $\langle w| (H' - G') |w\rangle > 0$ but maps the set of rays to themselves as the norm of the vector might change), the second follows from the fact that one can diagonalise the matrices (identity stays the same) and then it is just a set of inequalities involving real numbers, and the third follows from again multiplication by a positive matrix. The last one is the same as $f_\lambda(H) > f_\lambda(G)$. □

Corollary 94 (Tightness preservation under f_λ)**.** *Let $H \geq G$ and f_λ be on $(\chi, \xi) \supset spec[H \oplus G]$. There exists a $|w\rangle$ such that $\langle w| (H - G) |w\rangle = 0$ if and only if there exists a $|w_\lambda\rangle$ such that $\langle w_\lambda| (f_\lambda(H) - f_\lambda(G)) |w_\lambda\rangle = 0$.*

Proof. The contrapositive of the aforesaid condition is that $f_\lambda(H) > f_\lambda(G)$ if and only if $H > G$ which holds due to Lemma 93. □

Lemma 95 (Extending tightness preservation under f_λ to apparently divergent situations)**.** *Let X_h, X_g be diagonal matrices with $spec[X_h] \in (\chi, \xi]$, $spec[X_g] \in [\chi, \xi)$ and let f_λ be on $[\chi, \xi]$. Let, further, O be an orthogonal matrix such that $X_h \geq O X_g O^T$.*

There exists a vector $|w\rangle$ such that $\langle w| \left(f_{-\xi}(X_h) - O f_{-\xi}(X_g) O^T\right) |w\rangle = 0$ if and only if there exists a $|w_\lambda\rangle$ such that $\langle w_\lambda| \left(f_\lambda(X_h) - O f_\lambda(X_g) O^T\right) |w_\lambda\rangle = 0$ for $\lambda \in \mathbb{R}\backslash[\chi, \xi]$.

Similarly, there exists a vector $|w\rangle$ *such that* $\langle w|\left(f_{-\chi}(X_h) - Of_{-\chi}(X_g)O^T\right)|w\rangle = 0$ *if and only if there exists a* $|w_\lambda\rangle$ *such that* $\langle w_\lambda|\left(f_\lambda(X_h) - Of_\lambda(X_g)O^T\right)|w_\lambda\rangle = 0$ *for* $\lambda \in \mathbb{R}\backslash[\chi,\xi]$.

Proof. The trouble with this version of the tightness statement is that X_h has an eigenvalue ξ (if it doesn't then it reduces to the previous statement) which means that $f_{-\xi}(X_h)$ is not well defined. We assume that X_h can be expressed as

$$X_h = \begin{bmatrix} X_h' & \\ & \xi\mathbb{I}'' \end{bmatrix}$$

where X_h' has no eigenvalue equal to ξ and $\mathbb{I}''$ is the identity matrix in the subspace. We can write

$$\begin{aligned} & X_h > OX_gO^T \\ \iff & \begin{bmatrix} f_\lambda(X_h') & \\ & f_\lambda(\xi\mathbb{I}'') \end{bmatrix} > Of_\lambda(X_g)O^T \text{ for } \lambda \in \mathbb{R}\backslash[-\xi,-\chi] \\ \iff & \begin{bmatrix} f_\lambda(X_h') & \\ & \mathbb{I}'' \end{bmatrix} > \begin{bmatrix} \mathbb{I}' & \\ & f_\lambda(\xi\mathbb{I}'')^{-1/2} \end{bmatrix} Of_\lambda(X_g)O^T \begin{bmatrix} \mathbb{I}' & \\ & f_\lambda(\xi\mathbb{I}'')^{-1/2} \end{bmatrix} \text{ for } \lambda \in \mathbb{R}\backslash[-\xi,-\chi] \end{aligned}$$

where in the last line the expression has a well defined limit for $\lambda = -\xi$. This establishes the contrapositive variant of the statement we wanted to prove (similar to the strategy used for proving Corollary 94) once we note the following. If $\langle w|\left(f_{-\xi}(X_h) - Of_{-\xi}(X_g)O^T\right)|w\rangle = 0$ it is easy to see that $\begin{bmatrix} 0 & \\ & \mathbb{I}'' \end{bmatrix}|w\rangle = 0$ otherwise due to the constraint on the spectrum of X_g the aforesaid expression would be ∞. This entails that

$$\langle w|\left(\begin{bmatrix} f_{-\xi}(X_h') & \\ & \mathbb{I}'' \end{bmatrix} - \begin{bmatrix} \mathbb{I}' & \\ & f_{-\xi}(\xi\mathbb{I}'')^{-1/2} \end{bmatrix} Of_{-\xi}(X_g)O^T \begin{bmatrix} \mathbb{I}' & \\ & f_{-\xi}(\xi\mathbb{I}'')^{-1/2} \end{bmatrix}\right)|w\rangle = 0.$$

One can similarly prove the case for $f_{-\chi}(X_g)$. □

7.3 The Algorithm

We now state our algorithm and formally state its correctness. Thereafter, we motivate each step of the algorithm and prove its correctness.

Definition 96 (EMA Algorithm)**.** Given a finitely supported function a (we assume it is Λ-valid) proceed in the following three phases.

PHASE 1: INITIALISATION

- **Tightening procedure**: Let $[x_{\min}(\gamma'), x_{\max}(\gamma')]$ be the support domain for $a_{\gamma'}$. Let $\gamma \in (0,1]$ be the largest root of $m(\gamma', x_{\min}(\gamma'), x_{\max}(\gamma'))$. Let $x_{\max} := x_{\max}(\gamma)$ and $x_{\min} := x_{\min}(\gamma)$.
- **Spectral domain for the representation**: Find the smallest interval $[\chi,\xi]$ such that $l_\gamma(\lambda) \geq 0$ for $\lambda \in \bar{\mathbb{R}}\backslash[\chi,\xi]$. If $\text{supp}(g), \text{supp}(h)$ is not contained in $[\chi,\xi]$ then from all expansions of $[\chi,\xi]$ that contain the aforesaid sets, pick the smallest. Relabel this interval to $[\chi,\xi]$.

- **Shift**: Transform
$$a(x) \to a'(x') := a(x' + \chi - 1)$$
where instead of 1 any positive constant would do (justified by Corollary 84). Similarly transform
$$g(x) \to g'(x') := g(x' + \chi - 1)$$
$$h(x) \to h'(x') := h(x' + \chi - 1).$$
Relabel a' to be a, g' to be g and h' to be h. (Remark: We do not deduce h and g from a as its positive and negative part because they might now have common support due to the tightening procedure.)

- **The matrices**: For $n := n_g + n_h - 1$ we define $n \times n$ matrices with spectrum in $[\chi, \xi]$ and n dimensional vectors as
$$\begin{aligned} X_g^{(n)} &= \text{diag}[\chi, \chi, \dots x_{g_1}, x_{g_2} \dots, x_{g_{n_g}}], \\ X_{h_\gamma}^{(n)} &= \text{diag}[\gamma x_{h_1}, \gamma x_{h_2}, \dots, \gamma x_{h_{n_h}}, \xi, \xi, \dots], \\ \left|v^{(n)}\right\rangle &\doteq \left[0, 0 \dots, \sqrt{p_{g_1}}, \sqrt{p_{g_2}}, \dots, \sqrt{p_{g_{n_g}}}\right], \\ \left|w^{(n)}\right\rangle &\doteq \left[\sqrt{p_{h_1}}, \sqrt{p_{h_2}}, \dots, \sqrt{p_{h_{n_h}}}, 0, 0 \dots\right] \end{aligned}$$
where $g = \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}]$ and $h = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}]$. Note that n_g and n_h may be different.

- **Bootstrapping the iteration**:
  - Basis: $\left\{\left|t_{h_i}^{(n+1)}\right\rangle\right\}$ where $\left|t_{h_i}^{(n+1)}\right\rangle := |i\rangle$ for $i = 1, 2 \dots n$ where $|i\rangle$ refers to the standard basis in which the matrices and the vectors were originally written.
  - Matrix Instance: $\underline{\mathrm{X}}^{(n)} = \{X_h^{(n)}, X_g^{(n)}, \left|w^{(n)}\right\rangle, \left|v^{(n)}\right\rangle\}$.

PHASE 2: ITERATION

- Objective: Find the objects $\left|u_h^{(k)}\right\rangle, \bar{O}_g^{(k)}, \bar{O}_h^{(k)}$ and $s^{(k)}$ (which together relate $O^{(k)}$ to $O^{(k-1)}$ where $O^{(k)}$ solves $\underline{\mathrm{X}}^{(k)}$ and $O^{(k-1)}$ solves $\underline{\mathrm{X}}^{(k-1)}$ that is yet to be defined)

- Input: We will assume we are given
  - Basis: $\left\{\left|t_{h_i}^{(k+1)}\right\rangle\right\}$
  - Matrix Instance: $\underline{\mathrm{X}}^{(k)} = \left(X_h^{(k)}, X_g^{(k)}, \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle\right)$ with attribute $\chi^{(k)} > 0$
  - Function Instance: $\underline{\mathrm{X}}^{(k)} \to \underline{\mathrm{x}}^{(k)} = \left(h^{(k)}, g^{(k)}, a^{(k)}\right)$

- Output:
  - Basis: $\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_i}^{(k)}\right\rangle\right\}$
  - Matrix Instance: $\underline{\mathrm{X}}^{(k-1)} = \left(X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle\right)$ with attribute $\chi^{(k-1)} > 0$
  - Function Instance: $\underline{\mathrm{X}}^{(k-1)} \to \underline{\mathrm{x}}^{(k-1)} = \left(h^{(k-1)}, g^{(k-1)}, a^{(k-1)}\right)$

– Unitary Constructors: Either $\bar{O}_g^{(k)}$ and $\bar{O}_h^{(k)}$ are returned or $\bar{O}^{(k)}$ is returned. If $\bar{O}^{(k)}$ is returned, set $\bar{O}_g^{(k)} := \bar{O}^{(k)}$ and $\bar{O}_h^{(k)} = \mathbb{I}$.

– Relation: If $s^{(k)}$ is not specified, define $s^{(k)} := 1$.
If $s^{(k)} = 1$ then use

$$O^{(k)} := \bar{O}_h^{(k)} \left(\left| u_h^{(k)} \right\rangle \left\langle u_h^{(k)} \right| + O^{(k-1)} \right) \bar{O}_g^{(k)}$$

else use

$$O^{(k)} := \left[\bar{O}_h^{(k)} \left(\left| u_h^{(k)} \right\rangle \left\langle u_h^{(k)} \right| + O^{(k-1)} \right) \bar{O}_g^{(k)} \right]^T.$$

- Algorithm:

– **Boundary condition: If** $n_g = 0$ and $n_h = 0$ **then** set $k_0 = k$ and **jump to** phase 3.

– **Tighten**: Define $X_{h_{\gamma'}}^{(k)} := \gamma' X^{(k)}$. Let γ be the largest root of $m(\gamma', \chi_{\gamma'}^{(k)}, \xi_{\gamma'}^{(k)})$ for $a^{(k)}$ where $\chi_{\gamma'}^{(k)}, \xi_{\gamma'}^{(k)}$ are such that $[\chi_{\gamma'}^{(k)}, \xi_{\gamma'}^{(k)}]$ is the smallest interval containing $\text{spec}[X_{h_{\gamma'}}^{(k)} \oplus X_g^{(k)}]$. Relabel $X_{h_\gamma}^{(k)}$ to $X_h^{(k)}$, $\chi_\gamma^{(k)}$ to $\chi^{(k)}$ and $\xi_\gamma^{(k)}$ to $\xi^{(k)}$ for notational ease. Similarly relabel $a_\gamma^{(k)}$ to $a^{(k)}$, $h_\gamma^{(k)}$ to $h^{(k)}$, $l_\gamma^{(k)}$ to $l^{(k)}$. Update $x_{\min}$ and $x_{\max}$ to be such that $\text{supp}(a^{(k)}) \in [x_{\min}^{(k)}, x_{\max}^{(k)}]$ is the smallest such interval. Define $s^{(k)} := 1$.

– **Honest align**: **If** $l^{1(k)} = 0$ **then** define $\eta = -\chi^{(k)} + 1$

$$X_h'^{(k)} := X_h^{(k)} + \eta, \quad X_g'^{(k)} := X_g + \eta.$$

Else: Pick a root λ of the function $l^{(k)}(\lambda')$ in the domain $\mathbb{R} \backslash (-\xi^{(k)}, -\chi^{(k)})$. In the following two cases we consider the function f_λ on $[\chi^{(k)}, \xi^{(k)}]$.

∗ If $\lambda \neq -\chi^{(k)}$ then: Let $\eta = -f_\lambda(\chi^{(k)}) + 1$ where any positive constant could be chosen instead of 1. Define

$$X_h'^{(k)} := f_\lambda(X_h^{(k)}) + \eta, \quad X_g'^{(k)} := f_\lambda(X_g^{(k)}) + \eta.$$

∗ If $\lambda = -\chi^{(k)}$ then: Update $s^{(k)} = -1$. Let $\eta = -f_\lambda(\xi^{(k)}) - 1$ where any positive constant could be chosen instead of 1. Define

$$X_h'^{(k)} := X_g''^{(k)}, \quad X_g'^{(k)} := X_h''^{(k)},$$

where

$$X_h''^{(k)} := -f_\lambda(X_h^{(k)}) - \eta, \quad X_g''^{(k)} := -f_\lambda(X_g^{(k)}) - \eta$$

and make the replacement

$$\left| v^{(k)} \right\rangle \to \left| w^{(k)} \right\rangle$$
$$\left| w^{(k)} \right\rangle \to \left| v^{(k)} \right\rangle.$$

– **Remove spectral collision: If** $\lambda = -\chi^{(k)}$ or $\lambda = -\xi^{(k)}$ **then**

1. **Idle point: If** for some j', j, we have $q_{g_{j'}}^{(k)} = q_{h_j}^{(k)}$ and $y_{g_{j'}}^{(k)} = y_{h_j}^{(k)}$ **then** the solution is given by Definition 98
Jump to **End**.

2. **Final Extra**: **If** for some j, j' we have $q^{(k)}_{g_{j'}} > q^{(k)}_{h_j}$ and $y^{(k)}_{g_{j'}} = y^{(k)}_{h_j}$ **then** the solution is given by Definition 99
   Jump to **End**.
3. **Initial Extra**: **If** for some j, j' we have $q^{(k)}_{g_{j'}} < q^{(k)}_{h_j}$ and $y^{(k)}_{g_{j'}} = y^{(k)}_{h_j}$ **then** the solution is given by Definition 100
   Jump to **End**.

– **Evaluate the Reverse Weingarten Map**:

1. Consider the point $\left|w^{(k)}\right\rangle / \sqrt{\langle w^{(k)}| X'^{(k)}_h |w^{(k)}\rangle}$ on the ellipsoid $X'^{(k)}_h$. Evaluate the normal at this point as $\left|u^{(k)}_h\right\rangle = \mathcal{N}\left(\sum_{i=1}^{n^{(k)}_h} \sqrt{p^{(k)}_{h_i}} x'^{(k)}_{h_i} \left|t^{(k+1)}_{h_i}\right\rangle\right)$. Similarly evaluate $\left|u^{(k)}_g\right\rangle$, the normal at the point $\left|v^{(k)}\right\rangle / \sqrt{\langle w^{(k)}| X'^{(k)}_g |w^{(k)}\rangle}$ on the ellipsoid $X'^{(k)}_g$.
2. Recall that for a given diagonal matrix $X = \sum_i y_i |i\rangle \langle i| > 0$ and normal vector $|u\rangle = \sum_i u_i |i\rangle$ the Reverse Weingarten map is given by $W_{ij} = \left(-\frac{y_j^{-1} y_i^{-1} u_i u_j}{r^2} + y_i^{-1}\delta_{ij}\right)$ where $r = \sqrt{\sum y_i^{-1} u_i^2}$. Evaluate the Reverse Weingarten maps $W'^{(k)}_h$ and $W'^{(k)}_g$ along $\left|u^{(k)}_h\right\rangle$ and $\left|u^{(k)}_g\right\rangle$ respectively.
3. Find the eigenvectors and eigenvalues of the Reverse Weingarten maps. The eigenvectors of W'_h form the h tangent (and normal) vectors $\left\{\left\{\left|t^{(k)}_{h_i}\right\rangle\right\}, \left|u^{(k)}_h\right\rangle\right\}$. The corresponding radii of curvature are obtained from the eigenvalues $\left\{\{r^{(k)}_{h_i}\}, 0\right\} = \left\{\{c^{(k)-1}_{h_i}\}, 0\right\}$ which are inverses of the curvature values. The tangents are labelled in the decreasing order of radii of curvature (increasing order of curvature). Similarly for the g tangent (and normal) vectors. Fix the sign freedom in the eigenvectors by requiring $\left\langle t^{(k)}_{h_i} | w^{(k)}\right\rangle \geq 0$ and $\left\langle t^{(k)}_{g_i} | v^{(k)}\right\rangle \geq 0$.

– **Finite Method: If** $\lambda \neq -\xi^{(k)}$ and $\lambda \neq -\chi^{(k)}$, i.e. if it is the finite case **then**

1. $\bar{O}^{(k)} := \left|u^{(k)}_h\right\rangle \left\langle u^{(k)}_g\right| + \sum_{i=1}^{k-1} \left|t^{(k)}_{h_i}\right\rangle \left\langle t^{(k)}_{g_i}\right|$
2. $\left|v^{(k-1)}\right\rangle := \bar{O}^{(k)} \left|v^{(k)}\right\rangle - \left\langle u^{(k)}_h\right| \bar{O}^{(k)} \left|v^{(k)}\right\rangle \left|u^{(k)}_h\right\rangle$ and $\left|w^{(k-1)}\right\rangle := \left|w^{(k)}\right\rangle - \left\langle u^{(k)}_h | w^{(k)}\right\rangle \left|u^{(k)}_h\right\rangle$.
3. Define $X^{(k-1)}_h := \mathrm{diag}\{c^{(k)}_{h_1}, c^{(k)}_{h_2} \ldots, c^{(k)}_{h_{k-1}}\}$, $X^{(k-1)}_g := \mathrm{diag}\{c^{(k)}_{g_1}, c^{(k)}_{g_2} \ldots c^{(k)}_{g_{k-1}}\}$.
4. **Jump** to **End**.

– **Wiggle-v Method: If** $\lambda = -\xi^{(k)}$ or $\lambda = -\chi^{(k)}$ **then**

1. $\left|u^{(k)}_h\right\rangle$ is renamed to $\left|\bar{u}^{(k)}_h\right\rangle$, $\left|u^{(k)}_g\right\rangle$ remains the same.
2. Let $\tau = \cos\theta := \left\langle u^{(k)}_g | v^{(k)}\right\rangle / \left\langle \bar{u}^{(k)}_h | w^{(k)}\right\rangle$. Let $\left|\bar{t}^{(k)}_h\right\rangle$ be an eigenvector of $X'^{(k)-1}_h$ with zero eigenvalue (comment: this is also perpendicular to $\left|w^{(k)}\right\rangle$). Redefine

$$\left|u^{(k)}_h\right\rangle := \cos\theta \left|\bar{u}^{(k)}_h\right\rangle + \sin\theta \left|\bar{t}^{(k)}_h\right\rangle,$$

$$\left|t^{(k)}_{h_k}\right\rangle = s\left(-\sin\theta \left|\bar{u}^{(k)}_h\right\rangle + \cos\theta \left|\bar{t}^{(k)}_h\right\rangle\right)$$

   where the sign $s \in \{1, -1\}$ is fixed by demanding $\left\langle t^{(k)}_{h_k} | w^{(k)}\right\rangle \geq 0$.
3. $\bar{O}^{(k)}$ and $\left|v^{(k-1)}\right\rangle, \left|w^{(k-1)}\right\rangle$ are evaluated as step i and ii of the finite case (a).

4. Define

$$X_h'^{(k-1)} := \mathrm{diag}\{c_{h_1}^{(k)}, c_{h_2}^{(k)}, \ldots, c_{h_{k-1}}^{(k)}\}, \qquad X_g'^{(k-1)} := \mathrm{diag}\{c_{g_1}^{(k)}, c_{g_2}^{(k)}, \ldots, c_{g_{k-1}}^{(k)}\}.$$

Let $[\chi'^{(k-1)}, \xi'^{(k-1)}]$ denote the smallest interval containing $\mathrm{spec}[X_h'^{(k-1)} \oplus X_g'^{(k-1)}]$. Let $\lambda' = -\chi'^{(k-1)} + 1$ where instead of 1 any positive number would also work. Consider $f_{\lambda''}$ on $[\chi'^{(k-1)}, \xi'^{(k-1)}]$. Let $\eta = -f_{\lambda'}(\chi'^{(k-1)}) + 1$. Define

$$X_h^{(k-1)} := f_{\lambda'}(X_h'^{(k-1)}) + \eta, \qquad X_g^{(k-1)} := f_{\lambda'}(X_g'^{(k-1)}) + \eta.$$

5. **Jump** to **End**.

– **End**: Restart the current phase (phase 2) with the newly obtained $(k-1)$ sized objects.

PHASE 3: RECONSTRUCTION

Let k_0 be the iteration at which the algorithm stops. Using the relation

$$O^{(k)} = \bar{O}_g^{(k)} \left(\left| u_h^{(k)} \right\rangle \left\langle u_h^{(k)} \right| + O^{(k-1)} \right) \bar{O}_h^{(k)}$$

(or its transpose if $s^{(k)} = -1$), evaluate $O^{(k_1)}$ from $O^{(k_0)} := \mathbb{I}_{k_0}$, then $O^{(k_2)}$ from $O^{(k_1)}$, then $O^{(k_3)}$ from $O^{(k_2)}$ and so on until $O^{(n)}$ is obtained which solves the matrix instance $\underline{\mathrm{X}}^{(n)}$ we started with. In terms of EBRM matrices, the solution is given by $H = X_h^{(n)}$, $G = O^{(n)} X_g O^{(n)T}$, and $|w\rangle = \left| w^{(n)} \right\rangle$.

Theorem 97 (Correctness of the EMA Algorithm)**.** *Given a Λ-valid function, the EMA algorithm (see Definition 96) always finds an orthogonal matrix O of size at most $n \times n$ where $n = n_g + n_h$, such that the constraints on O stated in Theorem 1 corresponding to the function a, are satisfied.*

Definition 98 (Spectral Collision: Case Idle Point)**.**

$$\left\{ \left| u_h^{(k)} \right\rangle, \left| t_{h_1}^{(k)} \right\rangle, \left| t_{h_2}^{(k)} \right\rangle, \ldots \left| t_{h_{k-1}}^{(k)} \right\rangle \right\} \overset{\text{componentwise}}{:=}$$
$$\left\{ \left| t_{h_j}^{(k+1)} \right\rangle, \left| t_{h_1}^{(k+1)} \right\rangle, \left| t_{h_2}^{(k+1)} \right\rangle, \ldots \left| t_{h_{j-1}}^{(k+1)} \right\rangle, \left| t_{h_{j+1}}^{(k+1)} \right\rangle, \ldots \left| t_{h_k}^{(k+1)} \right\rangle \right\},$$

$$\bar{O}^{(k)} := \sum_{i=1}^{k} |a_i\rangle \left\langle t_{h_i}^{(k+1)} \right|,$$

where

$$\{|a_1\rangle, |a_2\rangle \ldots |a_k\rangle\} \overset{\text{componentwise}}{:=}$$

$$\begin{cases} \left\{ \left| t_{h_1}^{(k+1)} \right\rangle, \left| t_{h_2}^{(k+1)} \right\rangle, \ldots \left| t_{h_{j-1}}^{(k+1)} \right\rangle, \left| t_{h_{j'}}^{(k+1)} \right\rangle, \left| t_{h_j}^{(k+1)} \right\rangle, \left| t_{h_{j+1}}^{(k+1)} \right\rangle, \right. \\ \qquad \left. \ldots, \left| t_{h_{j'-1}}^{(k+1)} \right\rangle, \left| t_{h_{j'+1}}^{(k+1)} \right\rangle \ldots \left| t_{h_k}^{(k+1)} \right\rangle \right\} & j < j' \\ \\ \left\{ \left| t_{h_1}^{(k+1)} \right\rangle, \left| t_{h_2}^{(k+1)} \right\rangle, \ldots \left| t_{h_{j'-1}}^{(k+1)} \right\rangle, \left| t_{h_{j'+1}}^{(k+1)} \right\rangle \ldots \right. \\ \left. \left| t_{h_{j-1}}^{(k+1)} \right\rangle, \left| t_{h_{j'}}^{(k+1)} \right\rangle, \left| t_{h_j}^{(k+1)} \right\rangle, \left| t_{h_{j+1}}^{(k+1)} \right\rangle \ldots \left| t_{h_k}^{(k+1)} \right\rangle \right\} & j > j' \\ \\ \left\{ \left| t_{h_1}^{(k+1)} \right\rangle, \left| t_{h_2}^{(k+1)} \right\rangle, \ldots \left| t_{h_k}^{(k+1)} \right\rangle \right\} & j = j', \end{cases}$$

and

$$X_h^{(k-1)} := \sum_{i \neq j} y_{h_i}^{(k)} \left| t_{h_i}^{(k+1)} \right\rangle \left\langle t_{h_i}^{(k+1)} \right|,$$

$$X_g^{(k-1)} := \bar{O}^{(k)} X_g^{(k)} \bar{O}^{(k)T} - y_{h_j} \left| t_{h_j}^{(k+1)} \right\rangle \left\langle t_{h_j}^{(k+1)} \right|,$$

$$\left| w^{(k-1)} \right\rangle = \mathcal{N} \left[\left| w^{(k)} \right\rangle - \sqrt{p_{h_j}} \left| t_{h_j}^{(k+1)} \right\rangle \right], \left| v^{(k-1)} \right\rangle = \mathcal{N} \left[\bar{O}^{(k)} \left| v^{(k)} \right\rangle - \sqrt{p_{h_j}} \left| t_{h_j}^{(k+1)} \right\rangle \right].$$

(This specifies $\underline{\mathrm{X}}^{(k-1)} := \{X_h^{(k-1)}, X_g^{(k-1)}, \left| w^{(k-1)} \right\rangle, \left| v^{(k-1)} \right\rangle\}$.)

Definition 99 (Spectral Collision: Case Final Extra)**.**

$\underline{\mathrm{X}}^{(k-1)} := (X_h^{(k-1)}, X_g^{(k-1)}, \left| w^{(k-1)} \right\rangle, \left| v^{(k-1)} \right\rangle)$ where $X_h^{(k-1)} = \sum_{i=1}^{k-1} y_{h_i}^{(k-1)} \left| t_{h_i}^{(k)} \right\rangle \left\langle t_{h_i}^{(k)} \right|$, $X_g^{(k-1)} = \sum_{i=1}^{k-1} y_{g_i}^{(k-1)} \left| t_{h_i}^{(k)} \right\rangle \left\langle t_{h_i}^{(k)} \right|$, $\left| v^{(k-1)} \right\rangle = \mathcal{N} \left[\sum_{i=1}^{k-1} \sqrt{q_{g_i}^{(k-1)}} \left| t_{h_i}^{(k)} \right\rangle \right]$, $\left| w^{(k-1)} \right\rangle = \mathcal{N} \left[\sum_{i=1}^{k-1} \sqrt{q_{h_i}^{(k-1)}} \left| t_{h_i}^{(k)} \right\rangle \right]$ where the coordinates and weights are given by

$$\begin{aligned}
\left\{ q_{h_1}^{(k-1)}, \ldots q_{h_{k-1}}^{(k-1)} \right\} &\overset{\text{componentwise}}{=} \left\{ q_{h_1}^{(k)}, q_{h_2}^{(k)} \ldots, q_{h_{j-1}}^{(k)}, q_{h_{j+1}}^{(k)}, \ldots q_{h_k}^{(k)} \right\} \\
\left\{ q_{g_1}^{(k-1)}, \ldots q_{g_{k-1}}^{(k-1)} \right\} &\overset{\text{componentwise}}{=} \left\{ q_{g_2}^{(k)} \ldots, q_{g_{j'-1}}^{(k)}, q_{g_{j'}}^{(k)} - q_{h_j}^{(k)}, q_{g_{j'+1}}^{(k)}, q_{g_{j'+2}}^{(k)} \ldots q_{g_k}^{(k)} \right\} \\
\left\{ y_{g_1}^{(k-1)}, \ldots y_{g_{k-1}}^{(k-1)} \right\} &\overset{\text{componentwise}}{=} \left\{ y_{g_2}^{(k)}, \ldots y_{g_k}^{(k)} \right\} \\
\left\{ y_{h_1}^{(k-1)}, \ldots y_{h_{k-1}}^{(k-1)} \right\} &\overset{\text{componentwise}}{=} \left\{ y_{h_1}^{(k)}, \ldots y_{h_{j-1}}^{(k)}, y_{h_{j+1}}^{(k)} \ldots, y_{h_k}^{(k)} \right\},
\end{aligned}$$

the basis is given by

$$\begin{aligned}
\left\{ \left| u_h^{(k)} \right\rangle, \left| t_{h_1}^{(k)} \right\rangle \ldots \left| t_{h_{k-1}}^{(k)} \right\rangle \right\} &\overset{\text{componentwise}}{=} \\
\left\{ \left| t_{h_j}^{(k+1)} \right\rangle, \left| t_{h_1}^{(k+1)} \right\rangle, \left| t_{h_2}^{(k+1)} \right\rangle, \ldots \left| t_{h_{j-1}}^{(k+1)} \right\rangle, \left| t_{h_{j+1}}^{(k+1)} \right\rangle, \left| t_{h_{j+2}}^{(k+1)} \right\rangle \ldots \left| t_{h_k}^{(k+1)} \right\rangle \right\}.
\end{aligned}$$

The orthogonal matrices are given by $\bar{O}_h^{(k)} := \sum \left| t_{h_i}^{(k+1)} \right\rangle \langle a_i|$ where

$$\{|a_1\rangle, \ldots |a_k\rangle\} \to \left\{ \left| u_h^{(k)} \right\rangle, \left| t_{h_1}^{(k)} \right\rangle \ldots \left| t_{h_{k-1}}^{(k)} \right\rangle \right\},$$

$\bar{O}_g^{(k)} := \tilde{O}^{(k)} \bar{O}_h^{(k)}$ where

$$\begin{aligned}
\tilde{O}^{(k)} :=& \mathcal{N} \left[\sqrt{q_{h_j}^{(k)}} \left| u_h^{(k)} \right\rangle + \sqrt{q_{g_{j'}}^{(k)} - q_{h_j}^{(k)}} \left| t_{h_{j'}}^{(k)} \right\rangle \right] \mathcal{N} \left[\sqrt{q_{g_1}^{(k)}} \left\langle u_h^{(k)} \right| + \sqrt{q_{g_{j'}}^{(k)}} \left\langle t_{h_{j'}}^{(k)} \right| \right] \\
&+ \mathcal{N} \left[\sqrt{q_{g_{j'}}^{(k)} - q_{h_j}^{(k)}} \left| u_h^{(k)} \right\rangle - \sqrt{q_{h_j}^{(k)}} \left| t_{h_{j'}}^{(k)} \right\rangle \right] \mathcal{N} \left[\sqrt{q_{g_{j'}}^{(k)}} \left\langle u_h^{(k)} \right| - \sqrt{q_{g_1}^{(k)}} \left\langle t_{h_{j'}}^{(k)} \right| \right] \\
&+ \sum_{i \in \{1, \ldots k\} \setminus j'} \left| t_{h_i}^{(k)} \right\rangle \left\langle t_{h_i}^{(k)} \right|.
\end{aligned}$$

Definition 100 (Spectral Collision: Case Initial Extra)**.**

$\underline{X}^{(k-1)} := (X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle)$ where $X_h^{(k-1)} = \sum_{i=1}^{k-1} y_{h_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|$, $X_g^{(k-1)} = \sum_{i=1}^{k-1} y_{g_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|$, $\left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q_{g_i}^{(k-1)}} \left|t_{h_i}^{(k)}\right\rangle\right]$, $\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q_{h_i}^{(k-1)}} \left|t_{h_i}^{(k)}\right\rangle\right]$ where the coordinates and weights are given by

$$\begin{aligned}
\left\{q_{h_1}^{(k-1)}, \dots q_{h_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{q_{h_1}^{(k)} \dots, q_{h_{j-1}}^{(k)}, q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}, q_{h_{j+1}}^{(k)}, q_{h_{j+2}}^{(k)} \dots q_{h_{k-1}}^{(k)}\right\} \\
\left\{q_{g_1}^{(k-1)}, \dots q_{g_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{q_{g_1}^{(k)}, q_{g_2}^{(k)} \dots, q_{g_{j'-1}}^{(k)}, q_{g_{j'+1}}^{(k)}, \dots q_{g_k}^{(k)}\right\} \\
\left\{y_{g_1}^{(k-1)}, \dots y_{g_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{y_{g_1}^{(k)}, \dots y_{g_{j'-1}}^{(k)}, y_{g_{j'+1}}^{(k)} \dots, y_{g_k}^{(k)}\right\} \\
\left\{y_{h_1}^{(k-1)}, \dots y_{h_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{y_{h_1}^{(k)}, \dots y_{h_{k-1}}^{(k)}\right\},
\end{aligned}$$

the basis is given by

$$\begin{aligned}
&\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{componentwise}}{=} \\
&\left\{\left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \left|t_{h_{j+2}}^{(k+1)}\right\rangle \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\}.
\end{aligned}$$

The orthogonal matrices are given by $\bar{O}_h^{(k)} := \tilde{O}^{(k)} \sum |a_i\rangle \left\langle t_{h_i}^{(k+1)}\right|$ where

$$\{|a_1\rangle, \dots |a_k\rangle\} \overset{\text{componentwise}}{=} \left\{\left|t_{h_1}^{(k)}\right\rangle, \left|t_{h_2}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle, \left|u_h^{(k)}\right\rangle\right\}.$$

$$\begin{aligned}
\tilde{O}^{(k)} :=& \mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}} \left|u_h^{(k)}\right\rangle + \sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}} \left|t_{h_j}^{(k)}\right\rangle\right] \mathcal{N}\left[\sqrt{q_{h_k}^{(k)}} \left\langle u_h^{(k)}\right| + \sqrt{q_{g_j}^{(k)}} \left\langle t_{h_j}^{(k)}\right|\right] \\
&+ \mathcal{N}\left[\sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}} \left|u_h^{(k)}\right\rangle - \sqrt{q_{g_{j'}}^{(k)}} \left|t_{h_j}^{(k)}\right\rangle\right] \mathcal{N}\left[\sqrt{q_{g_j}^{(k)}} \left\langle u_h^{(k)}\right| - \sqrt{q_{h_k}^{(k)}} \left\langle t_{h_j}^{(k)}\right|\right] \\
&+ \sum_{i \in \{1,\dots k\} \setminus j} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|
\end{aligned}$$

and $\bar{O}_h^{(k)}$ is given by the basis change $\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\} \to \left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle \dots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\}$.

We start with motivating the exact step of the algorithm and then provide a proof or justification for the claims made in that step.

7.3.1 Phase 1: Initialisation

We are given a Λ-valid transition $g \to h$ and the EBRM function $a = h - g$. (Remark: We use below the notation used in the definition of a transition.)

Since the function is EBRM we know there are matrices $H \geq G$ and a vector $|\psi\rangle$ such that $a = \text{Prob}[H, |\psi\rangle] - \text{Prob}[G, |\psi\rangle]$. We also know that the maximum matrix size we need to consider is $n_g + n_h - 1$. We want to know the spectrum of the matrices involved to proceed.

The picture we have in mind is the following. We know that $H \geq G$ in terms of ellipsoids means that the H ellipsoid is inside the G ellipsoid (the order gets reversed). We try to expand the H ellipsoid (which means scaling down the matrix H) until it touches the G ellipsoid. When they touch we know that the corresponding spectrum of the matrices is optimal in some sense. This would be trivial if we already knew H and G but it serves as a good picture nonetheless.

What we do know is the function $a = h - g$. We use the equivalence between EBRM and valid functions to perform the aforesaid tightening procedure even without knowing the matrices. We use $a_\gamma = h_\gamma - g$ where $h_\gamma(x) = h(x/\gamma)$ and check if a_γ stays valid as we shrink γ from one to zero. We stop the moment we see any signature of tightness. Using this a_γ we determine the spectrum of the matrices certifying the EBRM claim.

We start with tightening till we find some operator monotone labelled by λ for which $l_{\gamma'}(\lambda)$ disappears. This captures the notion of the ellipsoids touching as after applying this operator monotone, along the $|w\rangle$ direction, the ellipsoids must touch.

Tightening procedure: Let $[x_{\min}(\gamma'), x_{\max}(\gamma')]$ be the support domain for $a_{\gamma'}$. Let $\gamma \in (0, 1]$ be the largest root of $m(\gamma', x_{\min}(\gamma'), x_{\max}(\gamma'))$. Let $x_{\max} := x_{\max}(\gamma)$ and $x_{\min} := x_{\min}(\gamma)$.

First we must show that there would indeed be a root of m as a function of γ' in the range $(0, 1]$. This is a direct consequence of Lemma 90. Second we must show that if we can find the matrices certifying a_γ is EBRM we can find the matrices certifying a is EBRM. This follows from the observation that $\gamma X_h \geq O X_g O^T$ implies that $X_h \geq \gamma X_h \geq O X_g O^T$.

We found a signature of tightness. Now we find the spectrum of the matrices involved.

Spectral domain for the representation: Find the smallest interval $[\chi, \xi]$ such that $l_\gamma(\lambda) \geq 0$ for $\lambda \in \bar{\mathbb{R}} \backslash [\chi, \xi]$. If $\text{supp}(g), \text{supp}(h)$ is not contained in $[\chi, \xi]$ then from all expansions of $[\chi, \xi]$ that contain the aforesaid sets, pick the smallest. Relabel this interval to $[\chi, \xi]$.

The interval so obtained will contain the spectrum of the matrices that certify a_γ is EBRM. This is justified by Lemma 91 using the fact that $l^1_\gamma \geq 0$ due to the previous step.

We need our matrices to be positive to be able to use the elliptic picture. We therefore shift the spectrum of the matrices so that the smallest eigenvalue required is one (where we could have used any positive number).

Shift: Transform

$$a(x) \to a'(x') := a(x' + \chi - 1)$$

where instead of 1 any positive constant would do (justified by Corollary 84). Similarly transform

$$g(x) \to g'(x') := g(x' + \chi - 1)$$
$$h(x) \to h'(x') := h(x' + \chi - 1).$$

Relabel a' to be a, g' to be g and h' to be h. (Remark: We do not deduce h and g from a as its positive and negative part because they might now have common support due to the tightening procedure.)

We use Corollary 84 to deduce that if $a(x)$ is EBRM with spectrum in $[\chi, \xi]$ then $a'(x') = a(x' + \chi - 1)$ is EBRM with spectrum in $[1, \xi - \chi + 1]$. We must also show that if we can find the matrices certifying a' is EBRM then we can find the matrices certifying a is EBRM. This is a direct consequence of the fact that $X'_h \geq O X'_g O^T \iff X_h - (\chi - 1)\mathbb{I} \geq O(X_g - (\chi - 1)\mathbb{I})O^T$. The orthogonal matrix, O, which is of primary interest remains unchanged.

With the spectrum determined and adjusted to the elliptic picture, which we put to use soon, we fix everything except the orthogonal matrix by using the Canonical Orthogonal Form (up to a permutation).

The matrices: For $n := n_g + n_h - 1$ we define $n \times n$ matrices with spectrum in $[\chi, \xi]$ and n dimensional vectors as

$$\begin{aligned}
X_g^{(n)} &= \text{diag}[\chi, \chi, \dots x_{g_1}, x_{g_2} \dots, x_{g_{n_g}}], \\
X_{h_\gamma}^{(n)} &= \text{diag}[\gamma x_{h_1}, \gamma x_{h_2}, \dots, \gamma x_{h_{n_h}}, \xi, \xi, \dots], \\
\left|v^{(n)}\right\rangle &\doteq \left[0, 0 \dots, \sqrt{p_{g_1}}, \sqrt{p_{g_2}}, \dots, \sqrt{p_{g_{n_g}}}\right], \\
\left|w^{(n)}\right\rangle &\doteq \left[\sqrt{p_{h_1}}, \sqrt{p_{h_2}}, \dots, \sqrt{p_{h_{n_h}}}, 0, 0 \dots\right]
\end{aligned}$$

where $g = \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}]$ and $h = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}]$. Note that n_g and n_h may be different.

We use Lemma 86 to deduce that $g \to h$ is a valid transition from the validity of a. Then we use Lemma 60 to write the diagonal matrices as described above given the valid transition $g \to h$, upto a permutation. Our objective is to find a matrix $O^{(n)}$ such that $O^{(n)} \left|v^{(n)}\right\rangle = \left|w^{(n)}\right\rangle$ while satisfying the inequality $X_h^{(n)} \geq O^{(n)} X_g^{(n)} O^{(n)T}$.

We now remove all the redundant information and pack it into a form which we can iteratively reduce to a simpler form.

Bootstrapping the iteration:

- Basis: $\left\{\left|t_{h_i}^{(n+1)}\right\rangle\right\}$ where $\left|t_{h_i}^{(n+1)}\right\rangle := |i\rangle$ for $i = 1, 2 \dots n$ where $|i\rangle$ refers to the standard basis in which the matrices and the vectors were originally written.
- Matrix Instance: $\underline{\mathrm{X}}^{(n)} = \{X_h^{(n)}, X_g^{(n)}, \left|w^{(n)}\right\rangle, \left|v^{(n)}\right\rangle\}$.

7.3.2 Phase 2: Iteration

We take as input the matrices X_g, X_h together with the vectors $|w\rangle, |v\rangle$ and churn out the same objects with one less dimension. We also output objects that, once we have iteratively reduced the problem to triviality, can be put together to find the orthogonal matrix O. See Figure 10 for a schematic reference.

- Objective: Find the objects $\left|u_h^{(k)}\right\rangle, \bar{O}_g^{(k)}, \bar{O}_h^{(k)}$ and $s^{(k)}$ (which together relate $O^{(k)}$ to $O^{(k-1)}$ where $O^{(k)}$ solves $\underline{\mathrm{X}}^{(k)}$ and $O^{(k-1)}$ solves $\underline{\mathrm{X}}^{(k-1)}$ that is yet to be defined)
- Input: We assume we are given
  - Basis: $\left\{\left|t_{h_i}^{(k+1)}\right\rangle\right\}$
  - Matrix Instance: $\underline{\mathrm{X}}^{(k)} = \left(X_h^{(k)}, X_g^{(k)}, \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle\right)$ with attribute $\chi^{(k)} > 0$
  - Function Instance: $\underline{\mathrm{X}}^{(k)} \to \underline{\mathrm{x}}^{(k)} = \left(h^{(k)}, g^{(k)}, a^{(k)}\right)$
- Output:
  - Basis: $\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_i}^{(k)}\right\rangle\right\}$
  - Matrix Instance: $\underline{\mathrm{X}}^{(k-1)} = \left(X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle\right)$ with attribute $\chi^{(k-1)} > 0$

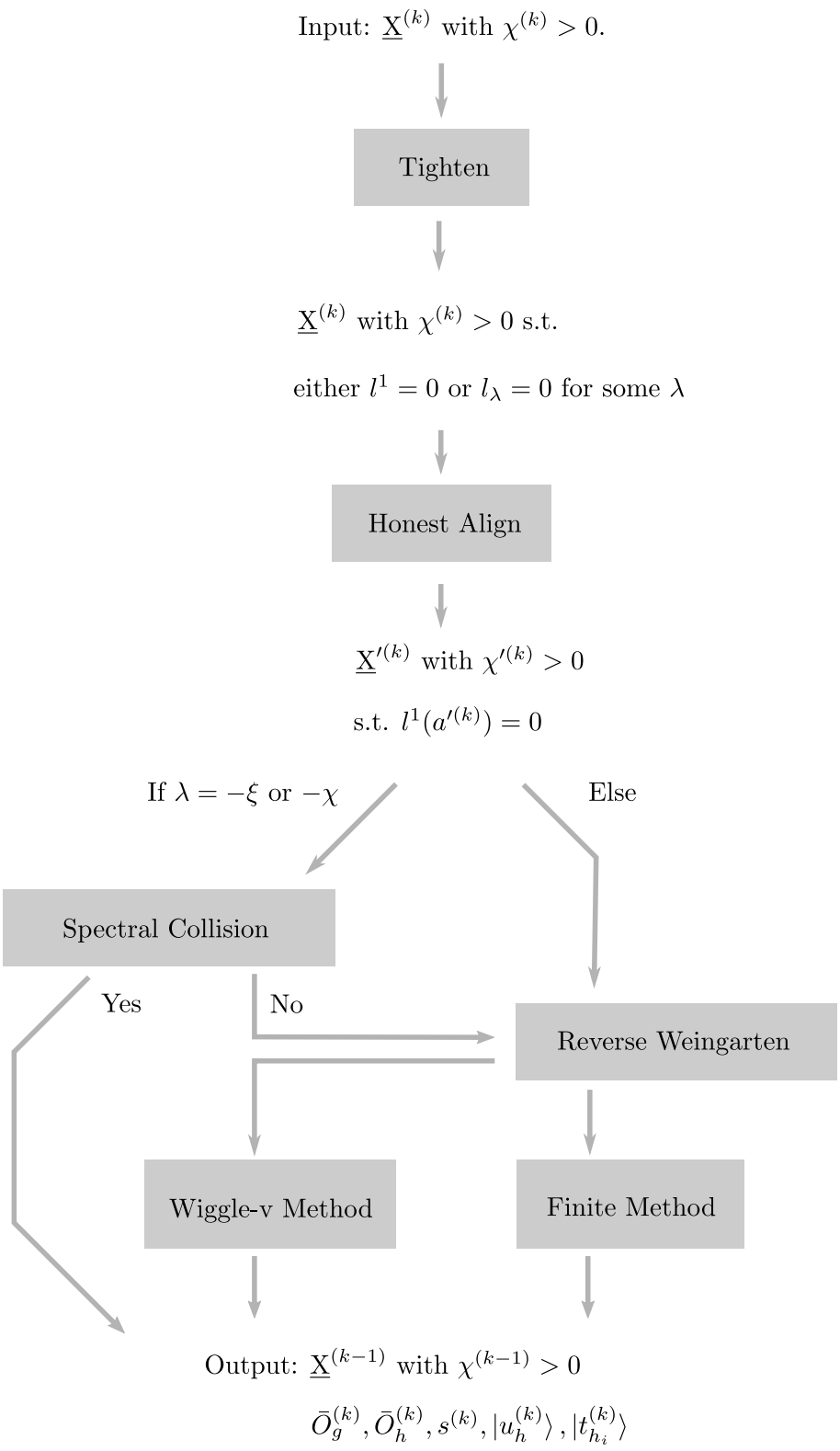


Figure 10: Overview of the main step, the iteration, of the algorithm (excluding the boundary condition).

– Function Instance: $\underline{\mathrm{X}}^{(k-1)} \to \underline{\mathrm{x}}^{(k-1)} = \left(h^{(k-1)}, g^{(k-1)}, a^{(k-1)}\right)$

– Unitary Constructors: Either $\bar{O}_g^{(k)}$ and $\bar{O}_h^{(k)}$ are returned or $\bar{O}^{(k)}$ is returned. If $\bar{O}^{(k)}$ is returned, set $\bar{O}_g^{(k)} := \bar{O}^{(k)}$ and $\bar{O}_h^{(k)} = \mathbb{I}$.

– Relation: If $s^{(k)}$ is not specified, define $s^{(k)} := 1$.
If $s^{(k)} = 1$ then use

$$O^{(k)} := \bar{O}_h^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}_g^{(k)}$$

else use

$$O^{(k)} := \left[\bar{O}_h^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}_g^{(k)}\right]^T.$$

Our task is to solve the matrix instance $\underline{\mathrm{X}}^{(k)}$, i.e. find a real unitary $O^{(k)}$ such that $X_h^{(k)} \geq O^{(k)} X_g^{(k)} O^{(k)T}$ and $O^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$. We assume that the solution exists and show that the solution to the smaller instance, denoted by $\underline{\mathrm{X}}^{(k-1)}$ must also exist. More precisely, we show that $O^{(k)}$ must have the form $O^{(k)} = \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}^{(k)}$ (for a solution to exist) which satisfies the aforesaid constraints granted we can find $O^{(k-1)}$ which acts on a $k-1$ dimensional Hilbert space orthogonal to $\left|u_h^{(k)}\right\rangle$ and satisfies constraints of the same form in the smaller dimension, viz. $X_h^{(k-1)} \geq O^{(k-1)} X_g^{(k-1)} O^{(k-1)T}$ and $O^{(k-1)} \left|v^{(k-1)}\right\rangle = \left|w^{(k-1)}\right\rangle$. Hence the assumption that $O^{(k)}$ has a solution allows us to deduce that $O^{(k-1)}$ must also have a solution. This allow us to iteratively solve the problem.
In certain trivial cases, where a point appears both before and after a transition viz. $X_g^{(k)}$ and $X_h^{(k)}$ have a common eigenvalue, the solution has the form $O^{(k)} = \bar{O}_h^{(k)} \left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}_g^{(k)}$. Finally, in one of the "infinite" cases denoted by the "Wiggle-v method" the solution will have the form $O^{(k)} = \left[\left(\left|u_h^{(k)}\right\rangle \left\langle u_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}^{(k)}\right]^T$.

- Algorithm:

If we reach a stage where the vector constraints have disappeared then we can simply stop.

– **Boundary condition: If** $n_g = 0$ and $n_h = 0$ **then** set $k_0 = k$ and **jump to** phase 3.

We assumed that an $O^{(k)}$ satisfying the constraints (listed right after the input/output section) exists. In this case it means that there exists an $O^{(k)}$ such that $X_h^{(k)} \geq O^{(k)} X_g^{(k)} O^{(k)T}$ as there is no vector $\left|v^{(k)}\right\rangle, \left|w^{(k)}\right\rangle$ to impose further constraints. Using Corollary 88 with $H = X_h^{(k)}$ and $G = O^{(k)} X_g^{(k)} O^{(k)T}$ we conclude that $O^{(k)}$ need only be a permutation matrix. Note that this step can never be entered right after the $\underline{\mathrm{X}}^{(n)}$ instance as we start with assuming $n_g, n_h > 0$. Further, since the protocol by construction always returns X_h and X_g in the ascending order the permutation matrix will be $\mathbb{I}$.

Finally, we deal with the interesting case. We again use the picture where the H ellipsoid is contained inside the G ellipsoid. We expand the H ellipsoid (which corresponds to shrinking the H matrix) until it touches the G ellipsoid as before by working with the function a.

– **Tighten**: Define $X^{(k)}_{h_{\gamma'}} := \gamma' X^{(k)}$. Let γ be the largest root of $m(\gamma', \chi^{(k)}_{\gamma'}, \xi^{(k)}_{\gamma'})$ for $a^{(k)}$ where $\chi^{(k)}_{\gamma'}, \xi^{(k)}_{\gamma'}$ are such that $[\chi^{(k)}_{\gamma'}, \xi^{(k)}_{\gamma'}]$ is the smallest interval containing $\mathrm{spec}[X^{(k)}_{h_{\gamma'}} \oplus X^{(k)}_g]$. Relabel $X^{(k)}_{h_\gamma}$ to $X^{(k)}_h$, $\chi^{(k)}_\gamma$ to $\chi^{(k)}$ and $\xi^{(k)}_\gamma$ to $\xi^{(k)}$ for notational ease. Similarly relabel $a^{(k)}_\gamma$ to $a^{(k)}$, $h^{(k)}_\gamma$ to $h^{(k)}$, $l^{(k)}_\gamma$ to $l^{(k)}$. Update $x_{\min}$ and $x_{\max}$ to be such that $\mathrm{supp}(a^{(k)}) \in [x^{(k)}_{\min}, x^{(k)}_{\max}]$ is the smallest such interval. Define $s^{(k)} := 1$.

Our burden again is to show that m as a function of γ' has a root. Unlike the first tightening procedure this time we know the spectrum of the matrices involved. Since we are given (by assumption) that the matrix instance has a solution we know that $l_{\gamma'=1}(\lambda) \geq 0$ and $l^1_{\gamma'=1} \geq 0$ for $\lambda \in \bar{\mathbb{R}} \backslash [\chi^{(k)}_{\gamma'=1}, \xi^{(k)}_{\gamma'=1}]$ using Lemma 85. We also know that $\chi^{(k)}_{\gamma'} > 0$ which means that $a^{(k)}$ (as deduced from the function instance of $\underline{\mathrm{X}}^{(k)}$) is a valid function. This observation lets us conclude that m as a function of γ' has a root in the required range because the reasoning behind a similar claim proved in Lemma 90 goes through unchanged.

The tightening procedure guarantees we will be able to find a λ which corresponds to an operator monotone such that after applying this function the ellipsoids, which we do not even know completely yet, must touch along the $|w\rangle$ direction. This piece of information is key to reducing the problem to a smaller instance of itself. Recall the picture with the H ellipsoid contained inside the G ellipsoid. If we know that they, in addition, touch at some known point then it must be so that the inner ellipsoid is more curved than the outer ellipsoid. When expressed algebraically, this condition essentially becomes that requirement that an ellipsoid $H^{(k-1)}$ that encodes the curvature of the ellipsoid $H^{(k)}$ at the point of contact must be contained inside the corresponding $G^{(k-1)}$ ellipsoid which encodes the curvature of the $G^{(k)}$ ellipsoid. The vector condition also reduces similarly. Subtleties arise when λ happens to have boundary values in its allowed range as this yields infinities and this has an interesting consequence.

– **Honest align**: **If** $l^{1(k)} = 0$ **then** define $\eta = -\chi^{(k)} + 1$

$$X'^{(k)}_h := X^{(k)}_h + \eta, \quad X'^{(k)}_g := X_g + \eta.$$

Else: Pick a root λ of the function $l^{(k)}(\lambda')$ in the domain $\mathbb{R} \backslash (-\xi^{(k)}, -\chi^{(k)})$. In the following two cases we consider the function f_λ on $[\chi^{(k)}, \xi^{(k)}]$.

* If $\lambda \neq -\chi^{(k)}$ then: Let $\eta = -f_\lambda(\chi^{(k)}) + 1$ where any positive constant could be chosen instead of 1. Define

$$X'^{(k)}_h := f_\lambda(X^{(k)}_h) + \eta, \quad X'^{(k)}_g := f_\lambda(X^{(k)}_g) + \eta.$$

* If $\lambda = -\chi^{(k)}$ then: Update $s^{(k)} = -1$. Let $\eta = -f_\lambda(\xi^{(k)}) - 1$ where any positive constant could be chosen instead of 1. Define

$$X'^{(k)}_h := X''^{(k)}_g, \quad X'^{(k)}_g := X''^{(k)}_h,$$

where

$$X''^{(k)}_h := -f_\lambda(X^{(k)}_h) - \eta, \quad X''^{(k)}_g := -f_\lambda(X^{(k)}_g) - \eta$$

and make the replacement

$$\left|v^{(k)}\right\rangle \to \left|w^{(k)}\right\rangle$$
$$\left|w^{(k)}\right\rangle \to \left|v^{(k)}\right\rangle.$$

If we have $\lambda = -\chi^{(k)}$ or $-\xi^{(k)}$ it means that at least one of the matrices (among $X_g^{(k)}$ and $X_h^{(k)}$ under f_λ) would diverge. We must remove eigenvalues common to both matrices as isolating the divergence makes it easier to handle.

– **Remove spectral collision: If** $\lambda = -\chi^{(k)}$ or $\lambda = -\xi^{(k)}$ **then**

If it so happens that the coordinate and the probability associated is the same we must leave the associated vector unchanged (up to a relabelling). The following simply formalises this procedure and encodes the remaining non-trivial part into a problem of one less dimension.

1. **Idle point: If** for some j', j, we have $q_{g_{j'}}^{(k)} = q_{h_j}^{(k)}$ and $y_{g_{j'}}^{(k)} = y_{h_j}^{(k)}$ **then** the solution is given by

$$\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle, \left|t_{h_2}^{(k)}\right\rangle, \dots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{componentwise}}{:=}$$
$$\left\{\left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\},$$

$$\bar{O}^{(k)} := \sum_{i=1}^{k} |a_i\rangle \left\langle t_{h_i}^{(k+1)}\right|,$$

where

$$\{|a_1\rangle, |a_2\rangle \dots |a_k\rangle\} \overset{\text{componentwise}}{:=}$$
$$\begin{cases} \left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'}}^{(k+1)}\right\rangle, \left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \right. & \\ \left. \dots, \left|t_{h_{j'-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'+1}}^{(k+1)}\right\rangle \dots \left|t_{h_k}^{(k+1)}\right\rangle \right\} & j < j' \\ \left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_{j'-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'+1}}^{(k+1)}\right\rangle \dots \right. & \\ \left. \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'}}^{(k+1)}\right\rangle, \left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle \dots \left|t_{h_k}^{(k+1)}\right\rangle \right\} & j > j' \\ \left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \dots \left|t_{h_k}^{(k+1)}\right\rangle\right\} & j = j', \end{cases}$$

and

$$X_h^{(k-1)} := \sum_{i \neq j} y_{h_i}^{(k)} \left|t_{h_i}^{(k+1)}\right\rangle \left\langle t_{h_i}^{(k+1)}\right|,$$
$$X_g^{(k-1)} := \bar{O}^{(k)} X_g^{(k)} \bar{O}^{(k)T} - y_{h_j} \left|t_{h_j}^{(k+1)}\right\rangle \left\langle t_{h_j}^{(k+1)}\right|,$$

$$\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\left|w^{(k)}\right\rangle - \sqrt{p_{h_j}} \left|t_{h_j}^{(k+1)}\right\rangle\right], \left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\bar{O}^{(k)} \left|v^{(k)}\right\rangle - \sqrt{p_{h_j}} \left|t_{h_j}^{(k+1)}\right\rangle\right].$$

(This specifies $\underline{X}^{(k-1)} := \{X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle\}$.)

Jump to **End**.

In this proof by x_{h_i} we mean y_{h_i}, similarly by x_{g_i} we mean y_{h_i}; we apologise for the inconvenience. We want to find an $O^{(k)}$ such that $X_h^{(k)} \geq O^{(k)} X_g^{(k)} O^{(k)T}$ and $O^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$. We do this in two stages. First, we re-arrange the entries of $X_g^{(k)}$ as $X_g^{\prime(k)} := O_p^{(k)} X_g^{(k)} O_p^{(k)T}$ and define $\left|v_p^{(k)}\right\rangle := O_p^{(k)} \left|v\right\rangle$ for an $O_p^{(k)}$ to be specified later. The re-arrangement will be such that $x_{g_{j'}}$ sits at the j, j location while the rest of the elements of $X_g^{\prime(k)}$ are arranged in the increasing order. Second, we solve our initial problem under the assumption that $j = j'$. The non-trivial part here would be showing that we can take $O^{(k)}$ to have the form $\left(\left|j\right\rangle \left\langle j\right| + O^{(k-1)}\right) \bar{O}^{(k)}$ without loss of generality.

Let us start with the first step. We denote the orthogonal matrix $O = \sum_i \left|b_i\right\rangle \left\langle a_i\right|$ by $\{\left|a_1\right\rangle, \left|a_2\right\rangle, \ldots \left|a_k\right\rangle\} \rightarrow \{\left|b_1\right\rangle, \left|b_2\right\rangle, \ldots \left|b_k\right\rangle\}$ where $\{\left|b_i\right\rangle\}$ and $\{\left|a_i\right\rangle\}$ each constitute an orthonormal basis. Using this notation then for the case $j < j'$, we define $O_p^{(k)}$ by

$$\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \ldots \left|t_{h_k}^{(k+1)}\right\rangle\right\} \rightarrow$$
$$\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \ldots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'}}^{(k+1)}\right\rangle, \left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \ldots, \left|t_{h_{j'-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'+1}}^{(k+1)}\right\rangle \ldots \left|t_{h_k}^{(k+1)}\right\rangle\right\},$$

for $j' < j$ we define it by

$$\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \ldots \left|t_{h_k}^{(k+1)}\right\rangle\right\} \rightarrow$$
$$\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \ldots \left|t_{h_{j'-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'+1}}^{(k+1)}\right\rangle \ldots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j'}}^{(k+1)}\right\rangle, \left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle \ldots \left|t_{h_k}^{(k+1)}\right\rangle\right\}$$

and if $j' = j$ we set $O_p^{(k)} = \mathbb{I}^{(k)}$.

For the second step, we solve the main problem under the assumption that $j' = j$. We are given $X_g^{'(k)} = \text{diag}\{x'_{g_1}, x'_{g_2} \ldots x'_{g_k}\}$ and $X_h^{(k)} = \text{diag}\{x_{h_1}, x_{h_2} \ldots x_{h_k}\}$ which are such that $x_{h_j} = x'_{g_j}$; $\left|v'^{(k)}\right\rangle \doteq (\sqrt{q'_{g_1}}, \sqrt{q'_{g_2}}, \ldots \sqrt{q'_{g_k}})^T$ and $\left|w^{(k)}\right\rangle \doteq (\sqrt{q_{h_1}}, \sqrt{q_{h_2}}, \ldots \sqrt{q_{h_k}})^T$ are such that $q_{h_j} = q'_{g_j}$. Let us define the matrix instance to be $\underline{X}^{\prime(k)} = \{X_h^{(k)}, X_g^{\prime(k)}, \left|v'^{(k)}\right\rangle, \left|w^{(k)}\right\rangle\}$. We have to find an $O'^{(k)}$ such that $X_h^{(k)} \geq O'^{(k)} X_g^{\prime(k)} O'^{(k)T}$ and $O'^{(k)} \left|v'^{(k)}\right\rangle = \left|w\right\rangle$. Let $\underline{X}^{\prime(k-1)} = \left\{X_h^{(k-1)}, X_g^{\prime(k-1)}, \left|v'^{(k-1)}\right\rangle, \left|w^{(k-1)}\right\rangle\right\}$ be the matrix instance obtained after removing the j^{th} entry from the vectors, viz. $\left|v'^{(k-1)}\right\rangle := \sum_{i \neq j} \sqrt{q'_{g_i}} \left|t_{h_i}^{(k+1)}\right\rangle$, $\left|w^{(k-1)}\right\rangle := \sum_{i \neq j} \sqrt{q_{h_i}} \left|t_{h_i}^{(k+1)}\right\rangle$ and similarly defining $X_g^{\prime(k-1)} = \text{diag}\{x'_{g_1}, x'_{g_2} \ldots x'_{g_{j-1}}, x'_{g_{j+1}}, \ldots x_{g_k}\}$, $X_h^{(k-1)} = \text{diag}\left\{x_{h_1}, x_{h_2} \ldots x_{h_{j-1}}, x_{h_{j+1}}, \ldots x_{h_k}\right\}$. Note that $a^{(k)} = a^{(k-1)}$ as the j^{th} point gets cancelled. This means that if there is an $O'^{(k)}$ satisfying the aforementioned constraints $a^{(k)}$ is EBRM on the spectral domain of $\underline{X}^{(k)}$. Since $a^{(k)} = a^{(k-1)}$ we know that $a^{(k-1)}$ is also EBRM on the same domain. From Lemma 86 (we will justify that k is large enough separately) we conclude that there must also exist an $O'^{(k-1)}$ which satisfies $X_h^{(k-1)} \geq O'^{(k-1)} X_g^{\prime(k-1)} O'^{(k-1)T}$ and $O'^{(k-1)} \left|v'^{(k-1)}\right\rangle = \left|w^{(k)}\right\rangle$.

With all this in place we can claim that without loss of generality we can write $O'^{(k)} = \left|t_{h_j}\right\rangle \left\langle t_{h_j}\right| + O'^{(k-1)}$ because if we can find some other $\tilde{O}'^{(k)}$ which satisfies

the required constraints then there exists an $O'^{(k-1)}$ which satisfies the corresponding constraints in the smaller dimension and that means we can show $O'^{(k)}$ also satisfies the required constraints,

$$X_h^{(k)} = x_{h_j} \left|t_{h_j}^{(k+1)}\right\rangle \left\langle t_{h_j}^{(k+1)}\right| + X_h^{(k-1)} \geq$$
$$x_{g_j} \left|t_{h_j}^{(k+1)}\right\rangle \left\langle t_{h_j}^{(k+1)}\right| + O'^{(k-1)} X_g'^{(k-1)} O'^{(k-1)} =$$
$$\left(\left|t_{h_j}^{(k+1)}\right\rangle \left\langle t_{h_j}^{(k+1)}\right| + O'^{(k-1)}\right) X_g'^{(k)} \left(\left|t_{h_j}^{(k+1)}\right\rangle \left\langle t_{h_j}^{(k+1)}\right| + O'^{(k-1)}\right)^T =$$
$$O'^{(k)} X_g'^{(k)} O'^{(k)T},$$

along with

$$O'^{(k)} \left|v'^{(k)}\right\rangle = \sqrt{q'_{g_j}} \left|t_{h_j}^{(k+1)}\right\rangle + O'^{(k-1)} \left|v'^{(k-1)}\right\rangle = \sqrt{q'_{g_j}} \left|t_{h_j}^{(k+1)}\right\rangle + \left|w^{(k-1)}\right\rangle = \left|w^{(k-1)}\right\rangle.$$

It remains to combine the two steps to produce the matrix $\bar{O}^{(k)}$, the vectors $\left\{\left|n_h^{(k)}\right\rangle, \left\{\left|t_{h_i}^{(k)}\right\rangle\right\}\right\}$, along with $\underline{\mathrm{X}}^{(k-1)}$. We use $X_g'^{(k)} = O_p^{(k)} X_g^{(k)} O_p^{(k)T}$ from the first step and substitute it in the inequality which we showed would hold, i.e.

$$X_h^{(k)} \geq O'^{(k)} X_g'^{(k)} O'^{(k)T} = O'^{(k)} O_p^{(k)} X_g O_p^{(k)T} O'^{(k)T}$$

and using $O_p^{(k)} \left|v^{(k)}\right\rangle = \left|v'^{(k)}\right\rangle$ we have

$$O'^{(k)} \left|v'^{(k)}\right\rangle = O'^{(k)} O_p^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle.$$

Comparing the inequality to the form $X_h^{(k)} \geq O^{(k)} X_g^{(k)} O^{(k)T}$, $O^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$ for

$$O^{(k)} = \left(\left|n_h^{(k)}\right\rangle \left\langle n_h^{(k)}\right| + O^{(k-1)}\right) \bar{O}^{(k)}$$

we get $\bar{O}^{(k)} = O_p^{(k)}$, $\left|n_h^{(k)}\right\rangle = \left|t_{h_j}^{(k+1)}\right\rangle$ and $O^{(k-1)} = O'^{(k-1)}$. Note that this $O^{(k)}$ is consistent with comparing the equality with $O^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$. The basis for the sub-problem, i.e. the $(k-1)$ dimensional problem, was the same as before except for the fact that we removed $\left|t_{h_j}^{(k+1)}\right\rangle$. Thus we define $\left\{\left|t_{h_1}^{(k)}\right\rangle, \left|t_{h_2}^{(k)}\right\rangle \ldots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\} = \left\{t_{h_1}^{(k+1)}, t_{h_2}^{(k+1)} \ldots t_{h_{j-1}}^{(k+1)}, t_{h_{j+1}}^{(k+1)}, \ldots t_{h_k}^{(k+1)}\right\}$. Identifying

$$\underline{\mathrm{X}}^{(k-1)} = \left\{X_h^{(k-1)}, X_g^{(k-1)}, \left|v^{(k-1)}\right\rangle, \left|w^{(k-1)}\right\rangle\right\}$$

with

$$\underline{\mathrm{X}}'^{(k-1)} = \left\{X_h^{(k-1)}, X_g'^{(k-1)}, \left|v'^{(k-1)}\right\rangle, \left|w^{(k-1)}\right\rangle\right\}$$

completes the argument since $O^{(k-1)}$ was already identified with $O'^{(k-1)}$ so we are just labelling here.

2. **Final Extra**: **If** for some j, j' we have $q_{g_{j'}}^{(k)} > q_{h_j}^{(k)}$ and $y_{g_{j'}}^{(k)} = y_{h_j}^{(k)}$ **then** the solution is given by $\underline{\mathrm{X}}^{(k-1)} := (X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle)$ where $X_h^{(k-1)} = \sum_{i=1}^{k-1} y_{h_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|$, $X_g^{(k-1)} = \sum_{i=1}^{k-1} y_{g_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|$, $\left|v^{(k-1)}\right\rangle =$

$\mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q_{g_i}^{(k-1)}}\left|t_{h_i}^{(k)}\right\rangle\right],\ \left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q_{h_i}^{(k-1)}}\left|t_{h_i}^{(k)}\right\rangle\right]$ where the coordinates and weights are given by

$$\begin{aligned}
\left\{q_{h_1}^{(k-1)},\dots q_{h_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{q_{h_1}^{(k)},q_{h_2}^{(k)}\dots,q_{h_{j-1}}^{(k)},q_{h_{j+1}}^{(k)},\dots q_{h_k}^{(k)}\right\}\\
\left\{q_{g_1}^{(k-1)},\dots q_{g_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{q_{g_2}^{(k)}\dots,q_{g_{j'-1}}^{(k)},q_{g_{j'}}^{(k)}-q_{h_j}^{(k)},q_{g_{j'+1}}^{(k)},q_{g_{j'+2}}^{(k)}\dots q_{g_k}^{(k)}\right\}\\
\left\{y_{g_1}^{(k-1)},\dots y_{g_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{y_{g_2}^{(k)},\dots y_{g_k}^{(k)}\right\}\\
\left\{y_{h_1}^{(k-1)},\dots y_{h_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{y_{h_1}^{(k)},\dots y_{h_{j-1}}^{(k)},y_{h_{j+1}}^{(k)}\dots,y_{h_k}^{(k)}\right\},
\end{aligned}$$

the basis is given by

$$\begin{aligned}
&\left\{\left|u_h^{(k)}\right\rangle,\left|t_{h_1}^{(k)}\right\rangle\dots\left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{componentwise}}{=}\\
&\left\{\left|t_{h_j}^{(k+1)}\right\rangle,\left|t_{h_1}^{(k+1)}\right\rangle,\left|t_{h_2}^{(k+1)}\right\rangle,\dots\left|t_{h_{j-1}}^{(k+1)}\right\rangle,\left|t_{h_{j+1}}^{(k+1)}\right\rangle,\left|t_{h_{j+2}}^{(k+1)}\right\rangle\dots\left|t_{h_k}^{(k+1)}\right\rangle\right\}.
\end{aligned}$$

The orthogonal matrices are given by $\bar{O}_h^{(k)} := \sum\left|t_{h_i}^{(k+1)}\right\rangle\langle a_i|$ where

$$\{|a_1\rangle,\dots|a_k\rangle\}\to\left\{\left|u_h^{(k)}\right\rangle,\left|t_{h_1}^{(k)}\right\rangle\dots\left|t_{h_{k-1}}^{(k)}\right\rangle\right\},$$

$\bar{O}_g^{(k)} := \tilde{O}^{(k)}\bar{O}_h^{(k)}$ where

$$\begin{aligned}
\tilde{O}^{(k)} :=& \mathcal{N}\left[\sqrt{q_{h_j}^{(k)}}\left|u_h^{(k)}\right\rangle+\sqrt{q_{g_{j'}}^{(k)}-q_{h_j}^{(k)}}\left|t_{h_{j'}}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{g_1}^{(k)}}\left\langle u_h^{(k)}\right|+\sqrt{q_{g_{j'}}^{(k)}}\left\langle t_{h_{j'}}^{(k)}\right|\right]\\
&+\mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}-q_{h_j}^{(k)}}\left|u_h^{(k)}\right\rangle-\sqrt{q_{h_j}^{(k)}}\left|t_{h_{j'}}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left\langle u_h^{(k)}\right|-\sqrt{q_{g_1}^{(k)}}\left\langle t_{h_{j'}}^{(k)}\right|\right]\\
&+\sum_{i\in\{1,\dots k\}\setminus j'}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|.
\end{aligned}$$

Jump to **End**.

We are given $\underline{\mathrm{X}}^{(k)} = (X_h^{(k)},X_g^{(k)},\left|w^{(k)}\right\rangle,\left|v^{(k)}\right\rangle)$ where $X_h^{(k)} = \sum_{i=1}^k y_{h_i}^{(k)}\left|t_{h_i}^{(k+1)}\right\rangle\left\langle t_{h_i}^{(k+1)}\right|$, $X_g^{(k)} = \sum_{i=1}^k y_{g_i}^{(k)}\left|t_{h_i}^{(k+1)}\right\rangle\left\langle t_{h_i}^{(k+1)}\right|$, $\left|v^{(k)}\right\rangle = \sum_{i=1}^k q_{g_i}^{(k)}\left|t_{h_i}^{(k+1)}\right\rangle$, $\left|w^{(k)}\right\rangle = \sum_{i=1}^k q_{h_i}^{(k)}\left|t_{h_i}^{(k+1)}\right\rangle$ which means the corresponding function instance $\underline{\mathrm{x}}^{(k)} = (h^{(k)},g^{(k)},a^{(k)})$ where, in particular we have, $a^{(k)} = \sum_{i\in\{1,\dots k\}\setminus j}q_{h_i}^{(k)}[y_{h_i}] - \sum_{i\in\{1,\dots k\}\setminus j'}q_{g_i}^{(k)}[y_{g_i}] - (q_{g_{j'}}^{(k)}-q_{h_j}^{(k)})[y_{h_j}]$. Since we assume $\underline{\mathrm{X}}^{(k)}$ has a solution it follows that $a^{(k)}$ is $[\chi,\xi]$ valid. Thus the transition $g^{(k-1)} := a_-^{(k)} \to a_+^{(k)} =: h^{(k-1)}$ is also $[\chi,\xi]$ valid where $g^{(k-1)}$ comprises $n_g^{(k-1)} = n_g^{(k)}$ points and $h^{(k-1)}$ comprises $n_h^{(k-1)} = n_h^{(k)} - 1$ points (using the attributes corresponding to the function instance $(h^{(k-1)},g^{(k-1)},h^{(k-1)}-g^{(k-1)})$; The notation would be of the form $g = \sum_{i=1}^{n_g}p_{g_i}[x_{g_i}]$ and $h = \sum_{i=1}^{n_h}p_{h_i}[x_{h_i}]$). Since $k = n_g^{(k)} + n_h^{(k)} - 1$ the aforesaid relation yields $k - 1 = n_g^{(k-1)} + n_h^{(k-1)} - 1$. We conclude that $\underline{\mathrm{X}}^{(k-1)} := (X_h^{(k-1)},X_g^{(k-1)},\left|w^{(k-1)}\right\rangle,\left|v^{(k-1)}\right\rangle)$ where $X_h^{(k-1)} = \sum_{i=1}^{k-1}y_{h_i}^{(k-1)}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|$, $X_g^{(k-1)} = \sum_{i=1}^{k-1}y_{g_i}^{(k-1)}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|$,

$\left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q_{g_i}^{(k-1)}}\left|t_{h_i}^{(k)}\right\rangle\right]$, $\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q_{h_i}^{(k-1)}}\left|t_{h_i}^{(k)}\right\rangle\right]$ has a solution for

$$\left\{q_{h_1}^{(k-1)},\ldots q_{h_{k-1}}^{(k-1)}\right\} \overset{\text{componentwise}}{=} \left\{q_{h_1}^{(k)},q_{h_2}^{(k)}\ldots,q_{h_{j-1}}^{(k)},q_{h_{j+1}}^{(k)},\ldots q_{h_k}^{(k)}\right\}$$
$$\left\{q_{g_1}^{(k-1)},\ldots q_{g_{k-1}}^{(k-1)}\right\} \overset{\text{componentwise}}{=} \left\{q_{g_2}^{(k)}\ldots,q_{g_{j'-1}}^{(k)},q_{g_{j'}}^{(k)}-q_{h_j}^{(k)},q_{g_{j'+1}}^{(k)},q_{g_{j'+2}}^{(k)}\ldots q_{g_k}^{(k)}\right\}$$
$$\left\{y_{g_1}^{(k-1)},\ldots y_{g_{k-1}}^{(k-1)}\right\} \overset{\text{componentwise}}{=} \left\{y_{g_2}^{(k)},\ldots y_{g_k}^{(k)}\right\}$$
$$\left\{y_{h_1}^{(k-1)},\ldots y_{h_{k-1}}^{(k-1)}\right\} \overset{\text{componentwise}}{=} \left\{y_{h_1}^{(k)},\ldots y_{h_{j-1}}^{(k)},y_{h_{j+1}}\ldots,y_{h_k}^{(k)}\right\}$$

as the corresponding function instance $\underline{\mathrm{x}}^{(k-1)}$ is indeed given by $(h^{(k-1)},g^{(k-1)},a^{(k-1)}=a^{(k)})$. Here $\{\left|t_{h_i}^{(k)}\right\rangle\}$ constitute an orthonormal basis which we relate to $\left|t_{h_i}^{(k+1)}\right\rangle$ shortly. We used the fact that $q_{g_1}^{(k)}=0$ as $y_{g_1}^{(k)}=\chi$ (To see this note that $k-1>n_g^{(k-1)}$ which means that many q_{g_i} are zero; by convention we write the smallest eigenvalue, χ first to increase the matrix size so the first $i=1,2\ldots\left(k-1-n_g^{(k-1)}\right)$ q_is are zero.). This means that there must exist an $O^{(k-1)}$ which solves $\underline{\mathrm{X}}^{(k-1)}$.

Let us take a moment to note the following basis change manoeuvre. Note that $X_h'\geq O'X_g'O'^T$ with $O'\left|v'\right\rangle=\left|w'\right\rangle$ is equivalent to $X_h\geq OX_gO^T$ with $O\left|v\right\rangle=\left|w\right\rangle$ where $O=\bar{O}_h^TO'\bar{O}_g$, $\bar{O}_g\left|v\right\rangle=\left|v'\right\rangle$, $\bar{O}_h\left|w\right\rangle=\left|w'\right\rangle$, $\bar{O}_hX_h\bar{O}_h^T=X_h'$, $\bar{O}X_g\bar{O}_g^T=X_g'$ which is easy to see by a simple substitution.

We first expand the matrix $\underline{\mathrm{X}}^{(k-1)}$ to k dimensions as follows. We already had $X_h^{(k-1)}\geq O^{(k-1)}X_g^{(k-1)}O^{(k-1)T}$ with $O^{(k-1)}\left|v^{(k-1)}\right\rangle=\left|w^{(k-1)}\right\rangle$ which we expand as

$$\underbrace{y_{h_j}^{(k)}\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right|+X_h^{(k-1)}}_{:=X_h'^{(k)}}\geq$$
$$\underbrace{\left(\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right|+O^{(k-1)}\right)}_{:=O'^{(k)}}\underbrace{\left(y_{h_j}^{(k)}\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right|+X_g^{(k-1)}\right)}_{:=X_g'^{(k)}}\left(\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right|+O^{(k-1)}\right)^T$$

with $\left|v'^{(k)}\right\rangle = \mathcal{N}\left[\sqrt{q_{h_j}^{(k)}}\left|u_h^{(k)}\right\rangle+\left|v^{(k-1)}\right\rangle\right]$ and $\left|w'^{(k)}\right\rangle = \mathcal{N}\left[\sqrt{q_{h_j}^{(k)}}\left|u_h^{(k)}\right\rangle+\left|w^{(k-1)}\right\rangle\right]$. Note that the matrix instance $\underline{\mathrm{X}}'^{(k)} := (X_h'^{(k)},X_g'^{(k)},\left|v'^{(k)}\right\rangle,\left|w'^{(k)}\right\rangle)$ yields $\underline{\mathrm{x}}'^{(k)}=\underline{\mathrm{x}}^{(k)}$. We can now use the equivalence we pointed out above to establish a relation between $X_h^{(k)}\geq O^{(k)}X_g^{(k)}O^{(k)T}$ and $X_h'^{(k)}\geq O'^{(k)}X_g'^{(k)}O'^{(k)T}$ by finding $\bar{O}_g$ and $\bar{O}_h$. We define, somewhat arbitrarily,

$$\left\{\left|u_h^{(k)}\right\rangle,\left|t_{h_1}^{(k)}\right\rangle\ldots\left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{componentwise}}{=}$$
$$\left\{\left|t_{h_j}^{(k+1)}\right\rangle,\left|t_{h_1}^{(k+1)}\right\rangle,\left|t_{h_2}^{(k+1)}\right\rangle,\ldots\left|t_{h_{j-1}}^{(k+1)}\right\rangle,\left|t_{h_{j+1}}^{(k+1)}\right\rangle,\left|t_{h_{j+2}}^{(k+1)}\right\rangle\ldots\left|t_{h_k}^{(k+1)}\right\rangle\right\}.$$

We require $\bar{O}_h^{(k)}\left|w^{(k)}\right\rangle$ to be $\left|w'^{(k)}\right\rangle$. This is simply a permutation matrix given by $\left\{\left|t_{h_1}^{(k+1)}\right\rangle,\ldots\left|t_{h_k}^{(k+1)}\right\rangle\right\}\rightarrow\left\{\left|u_h^{(k)}\right\rangle,\left|t_{h_1}^{(k)}\right\rangle\ldots\left|t_{h_{k-1}}^{(k)}\right\rangle\right\}$. Note that this

yields $\bar{O}_h^{(k)T} X_h'^{(k)} \bar{O}_h^{(k)} = X_h^{(k)}$. It remains to find $\bar{O}_g^{(k)}$ which we demand must satisfy $\bar{O}_g^{(k)} \left|v^{(k)}\right\rangle = \left|v'^{(k)}\right\rangle$. Observe first that $\bar{O}_h^{(k)} \left|v^{(k)}\right\rangle = \sqrt{q_{g_1}^{(k)}} \left|u_h^{(k)}\right\rangle + \sum_{i=2}^k \sqrt{q_{g_i}^{(k)}} \left|t_{h_{i-1}}^{(k)}\right\rangle$. We must now apply

$$\begin{aligned}\tilde{O}^{(k)} :=& \mathcal{N}\left[\sqrt{q_{h_j}^{(k)}}\left|u_h^{(k)}\right\rangle + \sqrt{q_{g_{j'}}^{(k)} - q_{h_j}^{(k)}}\left|t_{h_{j'}}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{g_1}^{(k)}}\left\langle u_h^{(k)}\right| + \sqrt{q_{g_{j'}}^{(k)}}\left\langle t_{h_{j'}}^{(k)}\right|\right] \\ &+ \mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)} - q_{h_j}^{(k)}}\left|u_h^{(k)}\right\rangle - \sqrt{q_{h_j}^{(k)}}\left|t_{h_{j'}}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left\langle u_h^{(k)}\right| - \sqrt{q_{g_1}^{(k)}}\left\langle t_{h_{j'}}^{(k)}\right|\right] \\ &+ \sum_{i\in\{1,\ldots k\}\setminus j'} \left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|\end{aligned}$$

to get $\bar{O}_g^{(k)} \left|v^{(k)}\right\rangle = \left|v'^{(k)}\right\rangle$ where we defined $\bar{O}_g^{(k)} := \tilde{O}^{(k)} \bar{O}_h^{(k)}$. (Note the expression could be simplified by using $q_{g_1} = 0$ which in fact is necessary for probability conservation.) Using $y_{h_j}^{(k)} = y_{g_{j'}}^{(k)}$ we can also see that $\bar{O}_g^{(k)T} X_g'^{(k)} \bar{O}_g^{(k)}$ is essentially $X_g^{(k)}$ with $\chi^{(k)}$ at $\left|t_{h_1}^{(k+1)}\right\rangle$ replaced by $y_{g_{j'}} (= y_{h_j})$. One can conclude therefore that $X_g'^{(k)} \geq \bar{O}_g^{(k)} X_g^{(k)} \bar{O}_g^{(k)T}$. Following the substitution manoeuvre we have

$$\begin{aligned} & X_h'^{(k)} \geq O'^{(k)} X_g'^{(k)} O'^{(k)T} \geq O'^{(k)} \bar{O}_g^{(k)} X_g^{(k)} \bar{O}_g^{(k)T} O'^{(k)T} \\ \iff & \bar{O}_h^{(k)T} X_h'^{(k)} \bar{O}_h^{(k)} \geq \underbrace{\bar{O}_h^{(k)T} O'^{(k)} \bar{O}_g^{(k)}}_{:=O^{(k)}} X_g^{(k)} \bar{O}_g^{(k)T} O'^{(k)T} \bar{O}_h^{(k)} \\ \iff & X_h^{(k)} \geq O^{(k)} X_g^{(k)} O^{(k)T}\end{aligned}$$

and similarly

$$\begin{aligned} & O'^{(k)} \left|v'^{(k)}\right\rangle = \left|w'^{(k)}\right\rangle \\ \iff & O'^{(k)} \bar{O}_g^{(k)} \left|v^{(k)}\right\rangle = \bar{O}_h^{(k)} \left|w^{(k)}\right\rangle \\ \iff & O^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle.\end{aligned}$$

This completes the proof.

3. **Initial Extra**: **If** for some j, j' we have $q_{g_{j'}}^{(k)} < q_{h_j}^{(k)}$ and $y_{g_{j'}}^{(k)} = y_{h_j}^{(k)}$ **then** the solution is given by $\underline{X}^{(k-1)} := (X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle)$ where $X_h^{(k-1)} = \sum_{i=1}^{k-1} y_{h_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|$, $X_g^{(k-1)} = \sum_{i=1}^{k-1} y_{g_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|$, $\left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q_{g_i}^{(k-1)}} \left|t_{h_i}^{(k)}\right\rangle\right]$, $\left|w^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q_{h_i}^{(k-1)}} \left|t_{h_i}^{(k)}\right\rangle\right]$ where the coordinates and weights are given by

$$\begin{aligned} \left\{q_{h_1}^{(k-1)}, \ldots q_{h_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{q_{h_1}^{(k)} \ldots, q_{h_{j-1}}^{(k)}, q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}, q_{h_{j+1}}^{(k)}, q_{h_{j+2}}^{(k)} \ldots q_{h_{k-1}}^{(k)}\right\} \\ \left\{q_{g_1}^{(k-1)}, \ldots q_{g_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{q_{g_1}^{(k)}, q_{g_2}^{(k)} \ldots, q_{g_{j'-1}}^{(k)}, q_{g_{j'+1}}^{(k)}, \ldots q_{g_k}^{(k)}\right\} \\ \left\{y_{g_1}^{(k-1)}, \ldots y_{g_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{y_{g_1}^{(k)}, \ldots y_{g_{j'-1}}^{(k)}, y_{g_{j'+1}}^{(k)} \ldots, y_{g_k}^{(k)}\right\} \\ \left\{y_{h_1}^{(k-1)}, \ldots y_{h_{k-1}}^{(k-1)}\right\} &\overset{\text{componentwise}}{=} \left\{y_{h_1}^{(k)}, \ldots y_{h_{k-1}}^{(k)}\right\},\end{aligned}$$

the basis is given by

$$\left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle \ldots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\} \overset{\text{componentwise}}{=}$$
$$\left\{\left|t_{h_j}^{(k+1)}\right\rangle, \left|t_{h_1}^{(k+1)}\right\rangle, \left|t_{h_2}^{(k+1)}\right\rangle, \ldots \left|t_{h_{j-1}}^{(k+1)}\right\rangle, \left|t_{h_{j+1}}^{(k+1)}\right\rangle, \left|t_{h_{j+2}}^{(k+1)}\right\rangle \ldots \left|t_{h_k}^{(k+1)}\right\rangle\right\}.$$

The orthogonal matrices are given by $\bar{O}_h^{(k)} := \tilde{O}^{(k)} \sum |a_i\rangle \left\langle t_{h_i}^{(k+1)}\right|$ where

$$\{|a_1\rangle, \ldots |a_k\rangle\} \overset{\text{componentwise}}{=} \left\{\left|t_{h_1}^{(k)}\right\rangle, \left|t_{h_2}^{(k)}\right\rangle \ldots \left|t_{h_{k-1}}^{(k)}\right\rangle, \left|u_h^{(k)}\right\rangle\right\}.$$

$$\begin{aligned}\tilde{O}^{(k)} :=& \mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle + \sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}}\left|t_{h_j}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{h_k}^{(k)}}\left\langle u_h^{(k)}\right| + \sqrt{q_{g_j}^{(k)}}\left\langle t_{h_j}^{(k)}\right|\right] \\ &+ \mathcal{N}\left[\sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle - \sqrt{q_{g_{j'}}^{(k)}}\left|t_{h_j}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{g_j}^{(k)}}\left\langle u_h^{(k)}\right| - \sqrt{q_{h_k}^{(k)}}\left\langle t_{h_j}^{(k)}\right|\right] \\ &+ \sum_{i\in\{1,\ldots k\}\setminus j}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|\end{aligned}$$

and $\bar{O}_h^{(k)}$ is given by the basis change $\left\{\left|t_{h_1}^{(k+1)}\right\rangle, \ldots \left|t_{h_k}^{(k+1)}\right\rangle\right\} \to \left\{\left|u_h^{(k)}\right\rangle, \left|t_{h_1}^{(k)}\right\rangle \ldots \left|t_{h_{k-1}}^{(k)}\right\rangle\right\}$.
Jump to **End**.

This proof will be very similar to the previous one. We are given $\underline{\mathrm{X}}^{(k)} = (X_h^{(k)}, X_g^{(k)}, \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle)$ where $X_h^{(k)} = \sum_{i=1}^k y_{h_i}^{(k)} \left|t_{h_i}^{(k+1)}\right\rangle \left\langle t_{h_i}^{(k+1)}\right|$, $X_g^{(k)} = \sum_{i=1}^k y_{g_i}^{(k)} \left|t_{h_i}^{(k+1)}\right\rangle \left\langle t_{h_i}^{(k+1)}\right|$, $\left|v^{(k)}\right\rangle = \sum_{i=1}^k q_{g_i}^{(k)} \left|t_{h_i}^{(k+1)}\right\rangle$, $\left|w^{(k)}\right\rangle = \sum_{i=1}^k q_{h_i}^{(k)} \left|t_{h_i}^{(k+1)}\right\rangle$ which means the corresponding function instance $\underline{\mathrm{x}}^{(k)} = (h^{(k)}, g^{(k)}, a^{(k)})$ where, in particular we have,

$$a^{(k)} = \sum_{i\in\{1,\ldots k\}\setminus j} q_{h_i}^{(k)}[y_{h_i}] + (q_{h_j}^{(k)} - q_{g_{j'}}^{(k)})[y_{h_j}] - \sum_{i\in\{1,\ldots k\}\setminus j'} q_{g_i}^{(k)}[y_{g_i}].$$

Since we assume $\underline{\mathrm{X}}^{(k)}$ has a solution it follows that $a^{(k)}$ is $[\chi, \xi]$ valid. Thus the transition $g^{(k-1)} := a_-^{(k)} \to a_+^{(k)} =: h^{(k-1)}$ is also $[\chi, \xi]$ valid where $g^{(k-1)}$ comprises $n_g^{(k-1)} = n_g^{(k)} - 1$ points and $h^{(k-1)}$ comprises $n_h^{(k-1)} = n_h^{(k)}$ points (using the attributes corresponding to the function instance $(h^{(k-1)}, g^{(k-1)}, h^{(k-1)} - g^{(k-1)})$; The notation would be of the form $g = \sum_{i=1}^{n_g} p_{g_i}[x_{g_i}]$ and $h = \sum_{i=1}^{n_h} p_{h_i}[x_{h_i}]$). Since $k = n_g^{(k)} + n_h^{(k)} - 1$ the aforesaid relation yields $n_g^{(k-1)} + n_h^{(k-1)} - 1 = k - 1$. We conclude that $\underline{\mathrm{X}}^{(k-1)} := (X_h^{(k-1)}, X_g^{(k-1)}, \left|w^{(k-1)}\right\rangle, \left|v^{(k-1)}\right\rangle)$ where $X_h^{(k-1)} = \sum_{i=1}^{k-1} y_{h_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|$, $X_g^{(k-1)} = \sum_{i=1}^{k-1} y_{g_i}^{(k-1)} \left|t_{h_i}^{(k)}\right\rangle \left\langle t_{h_i}^{(k)}\right|$, $\left|v^{(k-1)}\right\rangle = \mathcal{N}\left[\sum_{i=1}^{k-1} \sqrt{q_{g_i}^{(k-1)}} \left|t_{h_i}^{(k)}\right\rangle\right]$, $\left|w^{(k-1)}\right\rangle =$

$\mathcal{N}\left[\sum_{i=1}^{k-1}\sqrt{q_{h_i}^{(k-1)}}\left|t_{h_i}^{(k)}\right\rangle\right]$ have a solution for

$$\left\{q_{h_1}^{(k-1)},\ldots q_{h_{k-1}}^{(k-1)}\right\} \overset{\text{componentwise}}{=} \left\{q_{h_1}^{(k)}\ldots,q_{h_{j-1}}^{(k)},q_{h_j}^{(k)}-q_{g_{j'}}^{(k)},q_{h_{j+1}}^{(k)},q_{h_{j+2}}^{(k)}\ldots q_{h_{k-1}}^{(k)}\right\}$$
$$\left\{q_{g_1}^{(k-1)},\ldots q_{g_{k-1}}^{(k-1)}\right\} \overset{\text{componentwise}}{=} \left\{q_{g_1}^{(k)},q_{g_2}^{(k)}\ldots,q_{g_{j'-1}}^{(k)},q_{g_{j'+1}}^{(k)},\ldots q_{g_k}^{(k)}\right\}$$
$$\left\{y_{g_1}^{(k-1)},\ldots y_{g_{k-1}}^{(k-1)}\right\} \overset{\text{componentwise}}{=} \left\{y_{g_1}^{(k)},\ldots y_{g_{j'-1}}^{(k)},y_{g_{j'+1}}^{(k)}\ldots,y_{g_k}^{(k)}\right\}$$
$$\left\{y_{h_1}^{(k-1)},\ldots y_{h_{k-1}}^{(k-1)}\right\} \overset{\text{componentwise}}{=} \left\{y_{h_1}^{(k)},\ldots y_{h_{k-1}}^{(k)}\right\},$$

as the corresponding function instance $\underline{\mathrm{x}}^{(k-1)}$ is indeed given by $(h^{(k-1)},g^{(k-1)},a^{(k-1)}=a^{(k)})$. Here $\{\left|t_{h_i}^{(k)}\right\rangle\}$ constitute an orthonormal basis which we relate to $\left|t_{h_i}^{(k+1)}\right\rangle$ shortly. We used the fact that $q_{h_k}^{(k)}=0$ as $y_{h_k}^{(k)}=\xi$. (To see this note that $k-1>n_h^{(k-1)}$ which means that many q_{h_i} are zero; by convention we write the smallest eigenvalue, x_{h_1} first all the way till $x_{h_{n_h}}$ and then to increase the matrix size we append zeros so the $i=n_h,n_h+1\ldots k$ yield $q_{h_i}=0$.) This means that there must exist an $O^{(k-1)}$ which solves $\underline{\mathrm{X}}^{(k-1)}$.

Let us take a moment to note the following basis change manoeuvre. $X_h'\geq O'X_g'O'^T$ with $O'\left|v'\right\rangle=\left|w'\right\rangle$ is equivalent to $X_h\geq OX_gO^T$ with $O\left|v\right\rangle=\left|w\right\rangle$ where $O=\bar{O}_h^TO'\bar{O}_g$, $\bar{O}_g\left|v\right\rangle=\left|v'\right\rangle$, $\bar{O}_h\left|w\right\rangle=\left|w'\right\rangle$, $\bar{O}_hX_h\bar{O}_h^T=X_h'$, $\bar{O}X_g\bar{O}_g^T=X_g'$ which is easy to see by a simple substitution.

We first expand the matrix $\underline{\mathrm{X}}^{(k-1)}$ to k dimensions as follows. We already had $X_h^{(k-1)}\geq O^{(k-1)}X_g^{(k-1)}O^{(k-1)T}$ with $O^{(k-1)}\left|v^{(k-1)}\right\rangle=\left|w^{(k-1)}\right\rangle$ which we expand as

$$\underbrace{y_{h_j}^{(k)}\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right|+X_h^{(k-1)}}_{:=X_h'^{(k)}}\geq$$
$$\underbrace{\left(\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right|+O^{(k-1)}\right)}_{:=O'^{(k)}}\underbrace{\left(y_{h_j}^{(k)}\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right|+X_g^{(k-1)}\right)}_{:=X_g'^{(k)}}\left(\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right|+O^{(k-1)}\right)^T$$

with $\left|v'^{(k)}\right\rangle=\mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle+\left|v^{(k-1)}\right\rangle\right]$ and $\left|w'^{(k)}\right\rangle=\mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle+\left|w^{(k-1)}\right\rangle\right]$. Note that the matrix instance $\underline{\mathrm{X}}'^{(k)}:=(X_h'^{(k)},X_g'^{(k)},\left|v'^{(k)}\right\rangle,\left|w'^{(k)}\right\rangle)$ yields $\underline{\mathrm{x}}'^{(k)}=\underline{\mathrm{x}}^{(k)}$. We can now use the equivalence we pointed out above to establish a relation between $X_h^{(k)}\geq O^{(k)}X_g^{(k)}O^{(k)T}$ and $X_h'^{(k)}\geq O'^{(k)}X_g'^{(k)}O'^{(k)T}$ by finding $\bar{O}_g$ and $\bar{O}_h$. We define, somewhat arbitrarily,

$$\left\{\left|u_h^{(k)}\right\rangle,\left|t_{h_1}^{(k)}\right\rangle\ldots\left|t_{h_{k-1}}^{(k)}\right\rangle\right\}\overset{\text{componentwise}}{=}$$
$$\left\{\left|t_{h_j}^{(k+1)}\right\rangle,\left|t_{h_1}^{(k+1)}\right\rangle,\left|t_{h_2}^{(k+1)}\right\rangle,\ldots\left|t_{h_{j-1}}^{(k+1)}\right\rangle,\left|t_{h_{j+1}}^{(k+1)}\right\rangle,\left|t_{h_{j+2}}^{(k+1)}\right\rangle\ldots\left|t_{h_k}^{(k+1)}\right\rangle\right\}.$$

We require $\bar{O}_g^{(k)}\left|v^{(k)}\right\rangle$ to be $\left|v'^{(k)}\right\rangle$. This is simply a permutation matrix given by $\left\{\left|t_{h_1}^{(k+1)}\right\rangle,\ldots\left|t_{h_k}^{(k+1)}\right\rangle\right\}\rightarrow\left\{\left|u_h^{(k)}\right\rangle,\left|t_{h_1}^{(k)}\right\rangle\ldots\left|t_{h_{k-1}}^{(k)}\right\rangle\right\}$. Note that this yields $\bar{O}_g^{(k)T}X_g'^{(k)}\bar{O}_g^{(k)}=X_g^{(k)}$ as $y_{h_j}^{(k)}=y_{g_{j'}}^{(k)}$. It remains to find $\bar{O}_h^{(k)}$ which we demand must

satisfy $\bar{O}_h^{(k)} \left|w^{(k)}\right\rangle = \left|w'^{(k)}\right\rangle$. Let us define $\bar{O}_h^{(k)} = \tilde{O}^{(k)} \left(\sum_{i=1}^{k} |a_i\rangle \left\langle t_{h_i}^{(k+1)}\right|\right)$. Observe that for $\tilde{O}^{(k)} = \mathbb{I}$ we have $\bar{O}_h^{(k)} \left|w^{(k)}\right\rangle = q_{h_k}^{(k)} \left|u_h^{(k)}\right\rangle + \sum_{i=1}^{k-1} q_{h_i}^{(k)} \left|t_{h_i}^{(k)}\right\rangle$ where

$$\{|a_1\rangle, \ldots |a_k\rangle\} \overset{\text{componentwise}}{=} \left\{\left|t_{h_1}^{(k)}\right\rangle, \left|t_{h_2}^{(k)}\right\rangle \ldots \left|t_{h_{k-1}}^{(k)}\right\rangle, \left|u_h^{(k)}\right\rangle\right\}.$$

If we define

$$\begin{aligned}\tilde{O}^{(k)} :=& \mathcal{N}\left[\sqrt{q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle + \sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}}\left|t_{h_j}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{h_k}^{(k)}}\left\langle u_h^{(k)}\right| + \sqrt{q_{g_j}^{(k)}}\left\langle t_{h_j}^{(k)}\right|\right] \\ &+ \mathcal{N}\left[\sqrt{q_{h_j}^{(k)} - q_{g_{j'}}^{(k)}}\left|u_h^{(k)}\right\rangle - \sqrt{q_{g_{j'}}^{(k)}}\left|t_{h_j}^{(k)}\right\rangle\right]\mathcal{N}\left[\sqrt{q_{g_j}^{(k)}}\left\langle u_h^{(k)}\right| - \sqrt{q_{h_k}^{(k)}}\left\langle t_{h_j}^{(k)}\right|\right] \\ &+ \sum_{i\in\{1,\ldots k\}\setminus j}\left|t_{h_i}^{(k)}\right\rangle\left\langle t_{h_i}^{(k)}\right|\end{aligned}$$

we get $\bar{O}_h^{(k)} \left|w^{(k)}\right\rangle = \left|w'^{(k)}\right\rangle$ as desired. We can also see that $\bar{O}_h^{(k)T} X_g'^{(k)} \bar{O}_h^{(k)}$ is essentially X_g with $\xi^{(k)}$ at $\left|t_{h_k}^{(k+1)}\right\rangle$ replaced by y_{h_j}. We therefore conclude that $X_g'^{(k)} \geq \bar{O}_g^{(k)} X_g^{(k)} \bar{O}_g'^{(k)}$. Following the substitution manoeuvre we have

$$\begin{aligned} & X_h'^{(k)} \geq O'^{(k)} X_g'^{(k)} O'^{(k)T} \geq O'^{(k)} \bar{O}_g^{(k)} X_g^{(k)} \bar{O}_g^{(k)T} O'^{(k)T} \\ \iff & \bar{O}_h^{(k)T} X_h'^{(k)} \bar{O}_h^{(k)} \geq \underbrace{\bar{O}_h^{(k)T} O'^{(k)} \bar{O}_g^{(k)}}_{:=O^{(k)}} X_g^{(k)} \bar{O}_g^{(k)T} O'^{(k)T} \bar{O}_h^{(k)} \\ \iff & X_h^{(k)} \geq O^{(k)} X_g^{(k)} O^{(k)T}\end{aligned}$$

and similarly

$$\begin{aligned} & O'^{(k)} \left|v'^{(k)}\right\rangle = \left|w'^{(k)}\right\rangle \\ \iff & O'^{(k)} \bar{O}_g^{(k)} \left|v^{(k)}\right\rangle = \bar{O}_h^{(k)} \left|w^{(k)}\right\rangle \\ \iff & O^{(k)} \left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle.\end{aligned}$$

This completes the proof.

– **Evaluate the Reverse Weingarten Map**:

1. Consider the point $\left|w^{(k)}\right\rangle / \sqrt{\left\langle w^{(k)}\right| X_h'^{(k)} \left|w^{(k)}\right\rangle}$ on the ellipsoid $X_h'^{(k)}$. Evaluate the normal at this point as $\left|u_h^{(k)}\right\rangle = \mathcal{N}\left(\sum_{i=1}^{n_h^{(k)}} \sqrt{p_{h_i}^{(k)}} x_{h_i}'^{(k)} \left|t_{h_i}^{(k+1)}\right\rangle\right)$. Similarly evaluate $\left|u_g^{(k)}\right\rangle$, the normal at the point $\left|v^{(k)}\right\rangle / \sqrt{\left\langle w^{(k)}\right| X_g'^{(k)} \left|w^{(k)}\right\rangle}$ on the ellipsoid $X_g'^{(k)}$.
2. Recall that for a given diagonal matrix $X = \sum_i y_i |i\rangle \langle i| > 0$ and normal vector $|u\rangle = \sum_i u_i |i\rangle$ the Reverse Weingarten map is given by $W_{ij} = \left(-\frac{y_j^{-1} y_i^{-1} u_i u_j}{r^2} + y_i^{-1}\delta_{ij}\right)$ where $r = \sqrt{\sum y_i^{-1} u_i^2}$. Evaluate the Reverse Weingarten maps $W_h'^{(k)}$ and $W_g'^{(k)}$ along $\left|u_h^{(k)}\right\rangle$ and $\left|u_g^{(k)}\right\rangle$ respectively.

3. Find the eigenvectors and eigenvalues of the Reverse Weingarten maps. The eigenvectors of W'_h form the h tangent (and normal) vectors $\left\{\left\{\left|t^{(k)}_{h_i}\right\rangle\right\}, \left|u^{(k)}_h\right\rangle\right\}$. The corresponding radii of curvature are obtained from the eigenvalues $\left\{\{r^{(k)}_{h_i}\}, 0\right\} = \left\{\{c^{(k)-1}_{h_i}\}, 0\right\}$ which are inverses of the curvature values. The tangents are labelled in the decreasing order of radii of curvature (increasing order of curvature). Similarly for the g tangent (and normal) vectors. Fix the sign freedom in the eigenvectors by requiring $\left\langle t^{(k)}_{h_i} | w^{(k)} \right\rangle \geq 0$ and $\left\langle t^{(k)}_{g_i} | v^{(k)} \right\rangle \geq 0$.

– **Finite Method: If** $\lambda \neq -\xi^{(k)}$ and $\lambda \neq -\chi^{(k)}$, i.e. if it is the finite case **then**

1. $\bar{O}^{(k)} := \left|u^{(k)}_h\right\rangle \left\langle u^{(k)}_g\right| + \sum_{i=1}^{k-1} \left|t^{(k)}_{h_i}\right\rangle \left\langle t^{(k)}_{g_i}\right|$
2. $\left|v^{(k-1)}\right\rangle := \bar{O}^{(k)} \left|v^{(k)}\right\rangle - \left\langle u^{(k)}_h\right| \bar{O}^{(k)} \left|v^{(k)}\right\rangle \left|u^{(k)}_h\right\rangle$ and $\left|w^{(k-1)}\right\rangle := \left|w^{(k)}\right\rangle - \left\langle u^{(k)}_h | w^{(k)}\right\rangle \left|u^{(k)}_h\right\rangle$.
3. Define $X^{(k-1)}_h := \mathrm{diag}\{c^{(k)}_{h_1}, c^{(k)}_{h_2} \ldots, c^{(k)}_{h_{k-1}}\}$, $X^{(k-1)}_g := \mathrm{diag}\{c^{(k)}_{g_1}, c^{(k)}_{g_2} \ldots c^{(k)}_{g_{k-1}}\}$.
4. **Jump** to **End**.

Our first burden is to prove that $O^{(k)}$ must have the form $\left(\left|u^{(k)}_h\right\rangle \left\langle u^{(k)}_h\right| + O^{(k-1)}\right) \bar{O}^{(k)}$ for $\bar{O}^{(k)} := \left|u^{(k)}_h\right\rangle \left\langle u^{(k)}_g\right| + \sum_{i=1}^{k-1} \left|t^{(k)}_{h_i}\right\rangle \left\langle t^{(k)}_{g_i}\right|$ if $O^{(k)}$ is to be a solution of the matrix instance $\underline{\mathrm{X}}^{(k)}$. This is best explained by imagining that Arthur is trying to find the orthogonal matrix and Merlin already knows the orthogonal matrix but has still been following the steps performed so far. Recall that we are now at a point where

$$\begin{aligned}\sum a'(x)x &= \langle w| X'_h |w\rangle - \langle v| X'_g |v\rangle \\ &= \langle w| X'_h |w\rangle - \langle w| O X'_g O^T |w\rangle \\ &= 0.\end{aligned}$$

From Merlin's point of view along the $|w\rangle$ direction the ellipsoids X'_h and OX'_gO^T touch. Suppose he started with the ellipsoids X'_h, X'_g and only subsequently rotated the second one. He can mark the point along the direction $|v\rangle$ on the X'_g ellipsoid as the point that would after rotation touch the X'_h ellipsoid because as $X'_g \to OX'_gO^T$ the point along the $|v\rangle$ direction would get mapped to the point along the direction $O|v\rangle = |w\rangle$. Now, since the ellipsoids touch it must be so, Merlin deduces, that the normal of the ellipsoid X'_g at the point $|v\rangle / \sqrt{\langle v| X'_g |v\rangle}$ is mapped to the normal of the ellipsoid X'_h at the point $|w\rangle / \sqrt{\langle w| X'_h |w\rangle}$ when X'_g is rotated to OX'_gO^T, i.e. $O|u_g\rangle = |u_h\rangle$.

From Arthur's point of view, who has been following Merlin's reasoning, in addition to knowing that O must satisfy $O|v\rangle = |w\rangle$ he now knows that it must also satisfy $O|u_g\rangle = |u_h\rangle$.

Merlin further concludes that the curvature of the X'_g ellipsoid at the point $|v\rangle / \sqrt{\langle v| X'_g |v\rangle}$ must be more than the curvature of the X'_h ellipsoid at the point $|w\rangle / \sqrt{\langle w| X'_h |w\rangle}$. To be precise, he needs to find a method for evaluating this curvature. He knows that the brute-force way of doing this is to find a coordinate system with its origin on the said point and then imagining the manifold, locally, as a function from $n-1$ coordinates to one coordinate, call it $x_n(x_1, x_2 \ldots x_{n-1})$ (think of a sphere centred at the origin; it can be thought of, locally, as a function from x

and y to z given by $z = \sqrt{x^2 + y^2}$). The curvature of this object is a generalisation of the second derivative which forms a matrix with its elements given by $\partial_{x_i}\partial_{x_j}x_n$. Since this matrix is symmetric he knows it can be diagonalised. The directions of the eigenvectors of this matrix he calls the principle directions of curvature where the curvature values are the corresponding eigenvalues. He recalls that there is a simpler way of evaluating these principle directions and curvatures which uses the Weingarten map. The eigenvectors of the Reverse Weingarten map W'_h, evaluated for X'_h at $|w\rangle$, yield the normal and tangent vectors with the corresponding eigenvalues zero and radii of curvature respectively. Curvature is the inverse of the radius of curvature. Similarly for the Reverse Weingarten map W'_g evaluated for X'_g at $|v\rangle$. With this knowledge Merlin deduces that he can write, for some $\tilde{O}_{ij} \in \mathbb{R}$ such that $\sum_j \tilde{O}_{ij}\tilde{O}_{jk} = \delta_{ik}$,

$$\begin{aligned} O^{(k)} &= |u_h\rangle\langle u_g| + \sum_{i,j} \tilde{O}_{ij} |t_{h_i}\rangle \left\langle t_{g_j}\right| \\ &= \left(|u_h\rangle\langle u_h| + \underbrace{\sum_{i,j} \tilde{O}_{ij} |t_{h_i}\rangle \left\langle t_{h_j}\right|}_{=O^{(k-1)}} \right) \left(\underbrace{|u_h\rangle\langle u_g| + \sum_i |t_{h_i}\rangle\langle t_{g_i}|}_{=\bar{O}^{(k)}} \right) \end{aligned}$$

where he re-introduced the superscript in the orthogonal operators. He then turns to his intuition about the curvature of the smaller ellipsoid being more than that of the larger ellipsoid. He observes that equivalently, the radius of curvature of the smaller ellipsoid must be smaller than that of the larger ellipsoid. To make this precise he first notes that the Weingarten map W'_g gets transformed to OW'_gO^T when X'_g is rotated as OX'_gO^T. He considers the point $|w\rangle / \sqrt{\langle w| X'_h |w\rangle}$, which is shared by both the X'_h and the OX'_gO^T ellipsoid. It must be so, he reasons, that along all directions in the tangent plane, the X'_h ellipsoid (the smaller one, remember larger X'_h means smaller ellipsoid) must have a smaller radius of curvature than the OX'_gO^T ellipsoid, i.e. for all $|t\rangle \in \text{span}\{|t_{h_i}\rangle\}$, $\langle t| W'_h |t\rangle \leq \langle t| OW'_gO^T |t\rangle$. Restricting his attention to the tangent space he deduces the statement is equivalent to $W'_h \leq OW'_gO^T$. He writes this out explicitly as $\sum c_{h_i}^{-1} |t_{h_i}\rangle\langle t_{h_i}| \leq \sum c_{g_i}^{-1} O |t_{g_i}\rangle\langle t_{g_i}| O^T$. Now he uses the form of O he had deduced to obtain $\sum c_{h_i}^{-1} |t_{h_i}\rangle\langle t_{h_i}| \leq \sum c_{g_i}^{-1} O^{(k-1)} |t_{h_i}\rangle\langle t_{h_i}| O^{(k-1)T}$. From this he is able to deduce that the inequality $X_h^{(k-1)} \geq O^{(k-1)} X_g^{(k-1)} O^{(k-1)T}$ must hold.
Merlin's reasoning entails, Arthur summarises, that $O^{(k)}$ must always have the form

$$O^{(k)} = \left(\left| u_h^{(k)} \right\rangle \left\langle u_h^{(k)} \right| + O^{(k-1)} \right) \bar{O}^{(k)}$$

and that $O^{(k-1)}$ must satisfy the constraint

$$X_h^{(k-1)} \geq O^{(k-1)} X_g^{(k-1)} O^{(k-1)T}.$$

Merlin, surprised by the similarity of the constraint he obtained with the one he started with, extends his reasoning to the vector itself. He knows that $O^{(k)} \left| v^{(k)} \right\rangle = \left| w^{(k)} \right\rangle$ but now he substitutes for $O^{(k)}$ to obtain

$\left(\left|u_h^{(k)}\right\rangle\left\langle u_h^{(k)}\right| + O^{(k-1)}\right)\bar{O}^{(k)}\left|v^{(k)}\right\rangle = \left|w^{(k)}\right\rangle$. He observes that $O^{(k-1)}$ can not influence the $\left|u_h^{(k)}\right\rangle$ component of the vector $\bar{O}^{(k)}\left|v^{(k)}\right\rangle$. He thus projects out the $\left|u_h^{(k)}\right\rangle$ component to obtain

$$O^{(k-1)}\underbrace{\left(\bar{O}^{(k)}\left|v^{(k)}\right\rangle - \langle u_h|\,\bar{O}^{(k)}\left|v^{(k)}\right\rangle|u_h\rangle\right)}_{=\left|v^{(k-1)}\right\rangle} = \underbrace{\left|w^{(k)}\right\rangle - \left\langle u_h^{(k)}|w^{(k)}\right\rangle\left|u_h^{(k)}\right\rangle}_{=\left|w^{(k-1)}\right\rangle}.$$

With this, Arthur realises, he can reduce his problem involving a k-dimensional orthogonal matrix into a smaller problem in $k-1$ dimensions with exactly the same form. Since Merlin's orthogonal matrix was any arbitrary solution, and since the constraints involved do not depend explicitly on the solution (only on the initial problem), Arthur concludes that this reduction must hold for all possible solutions.

– **Wiggle-v Method: If** $\lambda = -\xi^{(k)}$ or $\lambda = -\chi^{(k)}$ **then**

The aforesaid method relies on matching the normals. It works well so long as the correct operator monotone (the monotone that yields X'_h and X'_g for which $|w\rangle / \sqrt{\langle w|\,X'_h\,|w\rangle}$ is a point on both X'_h and OX'_gO^T) doesn't yield infinities. If the operator monotone yields infinities it means that one of the directions involved has infinite curvature which in turn means that the component of the normal along this direction can be arbitrary. To see this, imagine having a line contained inside an ellipsoid (both centred at the origin) touching its boundaries. The line can be thought of as an ellipse with infinite curvature along one of the directions. The normal of the line at its tip is arbitrary and therefore we can't require the usual condition that normals of the two curves must coincide. The solution is to consider the sequence leading to the aforesaid situation.

1. $\left|u_h^{(k)}\right\rangle$ is renamed to $\left|\bar{u}_h^{(k)}\right\rangle$, $\left|u_g^{(k)}\right\rangle$ remains the same.
2. Let $\tau = \cos\theta := \left\langle u_g^{(k)}|v^{(k)}\right\rangle / \left\langle \bar{u}_h^{(k)}|w^{(k)}\right\rangle$. Let $\left|\bar{t}_h^{(k)}\right\rangle$ be an eigenvector of $X_h'^{(k)-1}$ with zero eigenvalue (comment: this is also perpendicular to $\left|w^{(k)}\right\rangle$). Redefine

$$\left|u_h^{(k)}\right\rangle := \cos\theta\left|\bar{u}_h^{(k)}\right\rangle + \sin\theta\left|\bar{t}_h^{(k)}\right\rangle,$$

$$\left|t_{h_k}^{(k)}\right\rangle = s\left(-\sin\theta\left|\bar{u}_h^{(k)}\right\rangle + \cos\theta\left|\bar{t}_h^{(k)}\right\rangle\right)$$

   where the sign $s \in \{1, -1\}$ is fixed by demanding $\left\langle t_{h_k}^{(k)}|w^{(k)}\right\rangle \geq 0$.
3. $\bar{O}^{(k)}$ and $\left|v^{(k-1)}\right\rangle, \left|w^{(k-1)}\right\rangle$ are evaluated as step 1 and 2 of the finite case.
4. Define

$$X_h'^{(k-1)} := \mathrm{diag}\{c_{h_1}^{(k)}, c_{h_2}^{(k)}, \ldots, c_{h_{k-1}}^{(k)}\}, \qquad X_g'^{(k-1)} := \mathrm{diag}\{c_{g_1}^{(k)}, c_{g_2}^{(k)}, \ldots, c_{g_{k-1}}^{(k)}\}.$$

   Let $[\chi'^{(k-1)}, \xi'^{(k-1)}]$ denote the smallest interval containing $\mathrm{spec}[X_h'^{(k-1)} \oplus X_g'^{(k-1)}]$. Let $\lambda' = -\chi'^{(k-1)} + 1$ where instead of 1 any positive number would also work. Consider $f_{\lambda''}$ on $[\chi'^{(k-1)}, \xi'^{(k-1)}]$. Let $\eta = -f_{\lambda'}(\chi'^{(k-1)}) + 1$. Define

$$X_h^{(k-1)} := f_{\lambda'}(X_h'^{(k-1)}) + \eta, \qquad X_g^{(k-1)} := f_{\lambda'}(X_g'^{(k-1)}) + \eta.$$

5. **Jump** to **End**.

We start with the case $\lambda = -\xi^{(k)}$. The other case with $\lambda = -\chi^{(k)}$ follows analogously. For the moment just imagine $\eta = 0$ for simplicity; for $\eta \neq 0$ the argument goes through essentially unchanged. Note that because $\langle w| f_{-\xi}(X_h) |w\rangle - \langle v| f_{-\xi}(X_g) |v\rangle$ is zero we can conclude that $y_{h_i}^{(k)} = \xi$ implies $q_{h_i} = 0$. After the application of the map $f_{-\xi}$ these $y_{h_i}^{(k)}$s and $y_{g_i}^{(k)}$s would become infinities but $\left\langle t_{h_i}^{(k+1)}|w\right\rangle$ and $\left\langle t_{g_i}^{(k+1)}|v\right\rangle$ would be zero where we suppressed the superscripts for $\left|v^{(k)}\right\rangle$ and $\left|w^{(k)}\right\rangle$. Since the eigenvalues are arranged in the ascending order in $X_h^{(k)}$ (in the $\left\{\left|t_{h_i}^{(k+1)}\right\rangle\right\}$ basis) we have $y_{h_k}^{(k)} = \xi$ and the corresponding vector is $\left|t_{h_k}^{(k+1)}\right\rangle =: |\bar{t}_h\rangle$. It would be useful to define $|\tilde{t}_{h_i}\rangle = |t_{h_i}\rangle$ for $i = 1, 2, \ldots j-1$ and $|\bar{t}_{h_l}\rangle = |t_{h_i}\rangle$ for $i = j, j+1, \ldots k$, $l = (i-j)+1$ where j is the smallest i for which $x_{h_i} = \xi$ (their existence is a straight forward consequence of dimension counting, $k \geq n_g + n_h - 1$). This allows us to speak of the subspace with eigenvalue ξ of $X_h^{(k)}$ easily. We focus on the two dimensional plane spanned by $|w\rangle$ and $|\bar{t}_h\rangle$.

Consider the M-view (Merlin's point of view). Since Merlin has a solution $O^{(k)}$ to the matrix instance

$$\underline{\mathrm{X}}^{(k)} = \{X_h^{(k)}, X_g^{(k)}, \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle\}$$

his solution is also a solution to the matrix instance

$$\underline{\mathrm{X}}^{(k)}(\lambda) := \left\{f_\lambda(X_h^{(k)}), f_\lambda(X_g^{(k)}), \left|w^{(k)}\right\rangle, \left|v^{(k)}\right\rangle\right\}$$

for $\lambda \leq -\xi$ but close enough to $-\xi$ such that $f_\lambda(X_h), f_\lambda(X_g) > 0$. This is a consequence of f_λ being operator monotone. Using Corollary 94 and Lemma 95 we know that since the ellipsoids corresponding to the matrix instance $\underline{\mathrm{X}}(-\xi)$ touch along $|w\rangle$ (as we are given that $\langle w| f_{-\xi}(X_h) |w\rangle - \langle w| O f_{-\xi}(X_g) O^T |w\rangle = \langle w| f_{-\xi}(X_h) |w\rangle - \langle v| f_{-\xi}(X_g) |v\rangle = 0$) there must also exist some vector $|c(\lambda)\rangle$ such that $\langle c(\lambda)| f_\lambda(X_h) |c(\lambda)\rangle - \langle c(\lambda)| O f_\lambda(X_g) O^T |c(\lambda)\rangle = 0$ that is the ellipsoids corresponding to the matrix instance $\underline{\mathrm{X}}(\lambda)$ touch along the said direction. (Caution: Do not confuse $|c(\lambda)\rangle$ with c_{h_i}/c_{g_i}. The latter are used for curvature values and the former refers to the contact vector just defined.) Note that to match the other conditions of the lemma it suffices to assume that X_h and X_g do not have a common eigenvalue which in turn is guaranteed by the "remove spectral collision" part.

It is easy to convince oneself that $\lim_{\lambda \to -\xi} |c(\lambda)\rangle = |w\rangle$ (hint: argue along the lines $f_\lambda(X_h)$ is very close to $f_{-\xi}(X_h)$ and so the vectors should also be very close which satisfy the condition). Note that we can write

$$|w\rangle = \sum_{i=1}^{j-1} q_{h_i} |\tilde{t}_{h_i}\rangle$$

because $\langle \bar{t}_{h_i} | w\rangle = 0$. There is no such restriction on $|c(\lambda)\rangle$ which can have the more general form $|c(\lambda)\rangle = \sum_{i=1}^{j-1} c(\lambda)_i |\tilde{t}_{h_i}\rangle + \sum_{i=j}^{k} c(\lambda)_i |\bar{t}_{h_l}\rangle$ where $l = (i-j)+1$. Restating one of the limit conditions, for $i = j, j+1 \ldots k$, we must have the $\lim_{\lambda \to -\xi} c(\lambda)_i = 0$. At this point we use the fact that if O is a solution it entails that

$$\acute{O}(\lambda) := \left(\sum_{i=1}^{j-1} |\tilde{t}_{h_i}\rangle \langle \tilde{t}_{h_i}| + \sum_{i,m=1}^{k-j+1} Q(\lambda)_{im} |\bar{t}_{h_i}\rangle \langle \bar{t}_{h_m}|\right) O$$

is also a solution, where $Q(\lambda)$ is an orthogonal matrix in the space spanned by $\{|\bar{t}_{h_i}\rangle\}$. This is a consequence of the fact that $\{|\bar{t}_{h_i}\rangle\}$ spans an eigenspace (with the same eigenvalue, $f_\lambda(\xi)$,) of $f_\lambda(X_h)$. We can use this freedom to ensure that the point of contact always has the form

$$|c(\lambda)\rangle = \sum_{i=1}^{j-1} c(\lambda)_i \, |\tilde{t}_{h_i}\rangle + \bar{c}(\lambda) \, |\bar{t}_h\rangle$$

where $\bar{c}(\lambda) = \sqrt{\sum_{i=j}^{k} c(\lambda)_i^2}$ which must vanish in the limit $\lambda \to -\xi$ as its constituents disappear in the said limit. Similarly $\lim_{\lambda\to-\xi} c(\lambda)_i = q_{h_i}$.
Next we evaluate the normals $|u_h(\lambda)\rangle$ at $|c(\lambda)\rangle$ for the ellipsoid represented by $f_\lambda(X_h)$ and similarly the normal $|\bar{u}_h\rangle$ at $|w\rangle$ for the ellipsoid represented by $f_{-\xi}(X_h)$ to show that $\lim_{\lambda\to-\xi} |u_h(\lambda)\rangle \neq |\bar{u}_h\rangle$ (see Figure 11). Notice that the right-most term in $|u_h(\lambda)\rangle = \mathcal{N}\left[\sum_{i=1}^{j-1} f_\lambda(y_{h_i}) c(\lambda)_i \, |\tilde{t}_{h_i}\rangle + f_\lambda(\xi)\bar{c}(\lambda) \, |\bar{t}_h\rangle\right]$ has $f_\lambda(\xi)$ approaching infinity and $\bar{c}(\lambda)$ approaching zero as λ tends to $-\xi$. This is why it can have a finite component along $|\bar{t}_h\rangle$. On the other hand, $|\bar{u}_h\rangle = \mathcal{N}\left[\sum_{i=1}^{j-1} f_{-\xi}(y_{h_i}) q_{h_i} \, |\tilde{t}_{h_i}\rangle\right]$ which has no component along $|\bar{t}_h\rangle$. Since $\lim_{\lambda\to-\xi} f_\lambda(y_{h_i}) = f_{-\xi}(y_{h_i})$ and $\lim_{\lambda\to-\xi} c(\lambda)_i = q_{h_i}$ for $i \in \{1, 2 \ldots j-1\}$, we can write

$$\lim_{\lambda\to-\xi} |u_h(\lambda)\rangle = \cos\theta \, |\bar{u}_h\rangle + \sin\theta \, |\bar{t}_h\rangle := |u_h\rangle \, .$$

Evidently, we must use $|u_h\rangle$ instead of $|\bar{u}_h\rangle$ to be able to use the reasoning of the finite method. However, we do not know $\cos\theta$ yet.
Our strategy is to proceed as in the finite method with the assumption that $|c(\lambda)\rangle$ is known (which it isn't as we only know it exists and how it behaves in the limit of $\lambda \to -\xi$) and then use a consistency condition to find $\cos\theta$ in terms of known quantities. At this point we re-introduce the superscripts as we will reduce the dimension of the problem as we proceed. Let the normal and tangents at $O^T \, |c(\lambda)\rangle$ for $f_\lambda(X_g)$ be given by $\left\{\left|u_g^{(k)}(\lambda)\right\rangle, \left\{t_{g_i}^{(k)}(\lambda)\right\}\right\}$. Similarly at $|c(\lambda)\rangle$ for $f_\lambda(X_h)$ the normal and tangents are $\left\{\left|u_h^{(k)}(\lambda)\right\rangle, \left\{t_{h_i}^{(k)}(\lambda)\right\}\right\}$. From the finite method we know that $O^{(k)}(\lambda) := \left(|u_h(\lambda)\rangle\langle u_h(\lambda)| + O^{(k-1)}\right)\bar{O}^{(k)}$ where $\bar{O}^{(k)} = \left|u_h^{(k)}(\lambda)\right\rangle\left\langle u_g^{(k)}(\lambda)\right| + \sum_i \left|t_{h_i}^{(k)}\right\rangle\left\langle t_{g_i}^{(k)}\right|$ can be used to reduce the problem into a smaller instance of itself. In particular, we must have $\left\langle u_h^{(k)}(\lambda)|w\right\rangle = \left\langle u_h^{(k)}(\lambda)\right| O^{(k)}(\lambda) \, |v\rangle = \left\langle u_g^{(k)}(\lambda)|v\right\rangle$ because $O^{(k-1)}$ can influence only the subspace spanned by $\left\{\left|t_{h_i}^{(k)}\right\rangle\right\}$ and the component of the vectors $|w\rangle$ and $O^{(k)} \, |v\rangle$ along $\left|u_h^{(k)}(\lambda)\right\rangle$ must match for consistency.
We can determine $\cos\theta$ by taking the limit of the aforesaid condition as $\langle u_h|w\rangle = \langle u_g|v\rangle$ where we again suppressed the superscripts. Substituting $|u_h\rangle = \cos\theta \, |\bar{u}_h\rangle + \sin\theta \, |\bar{t}_h\rangle$ we obtain

$$\cos\theta = \frac{\langle u_g|v\rangle}{\langle \bar{u}_h|w\rangle}.$$

It now remains to find the limit of the reverse Weingarten maps. The reverse Weingarten map for $f_\lambda(X_g)$ along the normal $|u_g(\lambda)\rangle$ is not of concern because it has a well defined limit as $\lambda \to -\xi$. We consider the case for $f_\lambda(X_h)$ along the normal

$|u_h(\lambda)\rangle$. Note that the support function as defined in Equation (7) is finite in the limit $\lambda \to -\xi$ (use the definition of the normal to get $\sum x_i^{-1} u_i^2 = \sum x_i^{-1} x_i^2 c_i^2 = \sum x_i c_i^2 = \langle c| X |c\rangle$, plug in $|c\rangle = |w\rangle$, $X = f_{-\xi}(X_h)$ and then use the fact that $\langle w| f_{-\xi}(X_h) |w\rangle - \langle v| f_{-\xi}(X_g) |v\rangle = 0$ which means both must be finite by noting that we already dealt with the troublesome case of $\infty - \infty$ in the "remove spectral collision" part). Let us denote it by $h(\lambda)$. Now the reverse Weingarten map as defined in Equation (8) is given by

$$(W_h(\lambda))_{im} = -\frac{1}{h(\lambda)^2} \frac{u_{h_i}(\lambda) u_{h_m}(\lambda)}{f_\lambda(y_{h_i}) f_\lambda(y_{h_m})} + \frac{\delta_{im}}{f_\lambda(x_{h_i})}.$$

Since $\lim_{\lambda \to -\xi} |u(\lambda)\rangle$ is well defined, $\lim_{\lambda \to -\xi} h(\lambda)$ is finite, we only need to show that $\lim_{\lambda \to -\xi} 1/f_\lambda(y_{h_i})$ is well defined. (We assumed η is zero so $f_{-\xi}(y_{h_i}) \neq 0$. If η is not zero we must consider $f_{-\xi}(y_{h_i}) + \eta$ everywhere but that changes no argument.) For $i = 1, 2 \ldots j-1$, $f_{-\xi}(y_{h_i})$ is finite but for $i = j, j+1 \ldots k$, $f_{-\xi}(y_{h_i})$ is not well defined however $1/f_{-\xi}(y_{h_i}) = 0$. We therefore conclude that

$$\lim_{\lambda \to -\xi} (W_h(\lambda))_{im} = \begin{cases} -\frac{1}{h^2} \frac{u_{h_i} u_{h_m}}{f_{-\xi}(y_{h_i}) f_{-\xi}(y_{h_m})} + \frac{\delta_{im}}{f_{-\xi}(y_{h_i})} & i, m \in \{1, 2 \ldots j-1\} \\ 0 & i, m \in \{j, j+1 \ldots k\} \end{cases} := (W_h)_{im}$$

which is simply the reverse Weingarten map evaluated for $f_{-\xi}(X_h)$ along $|u_h\rangle = \cos\theta |\bar{u}_h\rangle + \sin\theta |\bar{t}_h\rangle$ and $\cos\theta = \langle u_g|v\rangle / \langle \bar{u}_h|w\rangle$. It remains to relate W_h with the reverse Weingarten map, $\bar{W}_h$, evaluated for $f_{-\xi}(X_h)$ along $|\bar{u}_h\rangle$. Surprisingly, it is easy to see that $W_h = \bar{W}_h$ because only the $\cos\theta |\bar{u}_h\rangle$ part contributes to the non-zero portion of W_h and the $\cos\theta$ factor gets cancelled due to the h^2 term. Further, recall that the normal vector is always an eigenvector of the reverse Weingarten map evaluated along it, with eigenvalue zero. This tells us that if there is(are) tangent(s) with zero radius of curvature then the normal is not uniquely defined. This confirms what we already knew. Now since both $|\bar{u}_h\rangle$, $|\bar{t}_h\rangle$ have zero eigenvalues for $\bar{W}_h (= W_h)$ and $|u\rangle = \cos\theta |\bar{u}_h\rangle + \sin\theta |\bar{t}_h\rangle$ we define $|t_h\rangle := s\left(\sin\theta |\bar{u}_h\rangle - \cos\theta |\bar{t}_h\rangle\right)$ to span the same space so that $|u\rangle$ is the correct normal vector (as we deduced earlier in our discussion) and $|t_h\rangle$ is the correct tangent vector corresponding to the point $|w\rangle$ of $f_{-\xi}(X_h)$.

The final step is to convert the condition on the reverse Weingarten map into a condition on the Weingarten map (inverse of the reverse Weingarten map). After extracting the tangent vectors appropriately, one simply needs to add a constant before inverting to obtain the Weingarten map condition. This is done in the last step. This completes the proof of the wiggle-v method for $\lambda = -\xi$.

To see how the same reasoning applies to the $\lambda = -\chi$ case first note that for $\lambda \geq -\chi$ we have $f_\lambda(X_h), f_\lambda(X_g) < 0$ (assuming $\eta = 0$ as before). The condition $f_\lambda(X_h) \geq O f_\lambda(X_g) O^T$ can then be expressed as $-f_\lambda(X_g) \geq -O^T f_\lambda(X_h) O$ with $O^T |w\rangle = |v\rangle$ which can now be reasoned analogous to the aforementioned analysis.

– **End**: Restart the current phase (phase 2) with the newly obtained $(k-1)$ sized objects. We end with giving the dimension argument. The dimension after every iteration is $k - 1 \geq n_g^{(k-1)} + n_h^{(k-1)} - 1$ if we start with the assumption that $k \geq n_g^{(k)} + n_h^{(k)} - 1$. The reason is that either $n_g^{(k-1)} = n_g^{(k)} - 1$ or $= n_g^{(k)}$. Similarly, either $n_h^{(k-1)} = n_h^{(k)} - 1$ or

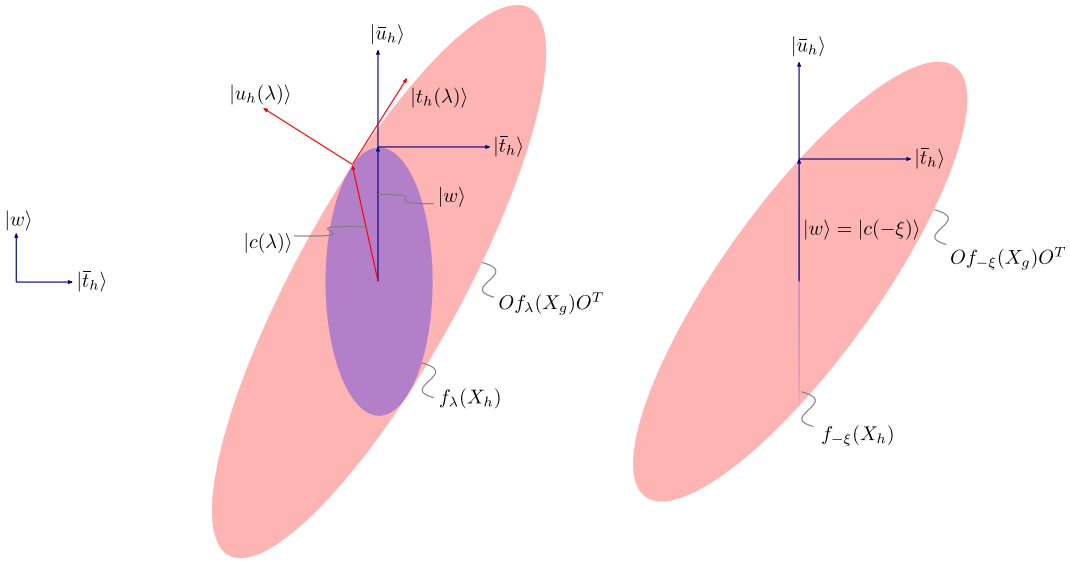


Figure 11: A sequence leading to infinite curvature.

$= n_h^{(k)}$. Justification of this is simply that we remove at least one component from the two vectors (from the $n_g^{(k)}$ for the usual wiggle-v). To see this, note that in the finite case we remove one from both as we write express the vector in a new basis. This new basis is the space where the vector has finite support. We then remove one of the components in the sub-problem. In the infinite case, it is possible that we remove one and add one for $n_h^{(k-1)}$, assuming it is the usual wiggle-v, but we necessarily reduce $n_g^{(k-1)}$ as this is similar to the finite case. For the other wiggle-v, g and h get swapped but the counting stays the same.

7.3.3 Phase 3: Reconstruction

Let k_0 be the iteration at which the algorithm stops. Using the relation

$$O^{(k)} = \bar{O}_g^{(k)} \left(\left| u_h^{(k)} \right\rangle \left\langle u_h^{(k)} \right| + O^{(k-1)} \right) \bar{O}_h^{(k)}$$

(or its transpose if $s^{(k)} = -1$), evaluate $O^{(k_1)}$ from $O^{(k_0)} := \mathbb{I}_{k_0}$, then $O^{(k_2)}$ from $O^{(k_1)}$, then $O^{(k_3)}$ from $O^{(k_2)}$ and so on until $O^{(n)}$ is obtained which solves the matrix instance $\underline{\mathrm{X}}^{(n)}$ we started with. In terms of EBRM matrices, the solution is given by $H = X_h^{(n)}$, $G = O^{(n)} X_g O^{(n)T}$, and $|w\rangle = \left| w^{(n)} \right\rangle$.

8 Conclusion

In the first part, we described a framework which we used to construct the unitaries required to implement the bias 1/10 protocol. In the second part, we described the EMA algorithm which allows us to find the unitaries corresponding to arbitrary Λ-valid moves, which combined with the framework, allows us to numerically find quantum WCF protocols with arbitrarily small bias.

A preliminary implementation of the EMA algorithm on python [20], which is usable but not automated enough for an end-user, has already yielded the following interesting results and future research directions.

1. *Mochon's denominator needs neither padding nor operator monotones.* For assignments given by Mochon's denominator, it is known that $\langle x_h \rangle = \langle x_g \rangle$ which means that for the first iteration of the algorithm, we need not use any operator monotone function. This was clear. The surprising result of the numerical implementation was that even for subsequent iterations one need not use operator monotones, which also explains why[6] we did not need padding, i.e. the (solution) orthogonal matrix has size $n = n_g = n_h$. An interesting open question is to analytically prove this as Mochon's denominator based assignment is very close to the kind of assignments used in Mochon's protocols. If analytic expressions for the unitaries corresponding to the former (Mochon's denominator) can be found, the unitaries corresponding to the assignments used in Mochon's protocols would be a small perturbation thereof. Hence, this might lead us to analytic expressions for the unitaries involved in games with arbitrarily low biases.

2. *Moves in the bias* 1/18 *protocol do not need padding (no wiggle-v).* We already know analytically that there are specific cases where padding is required. However, when we tried to numerically implement the moves involved in Mochon's protocols going as low as $\epsilon = 1/18$, to our surprise, we found that in no case was padding necessary (which means the wiggle-v method was never invoked). It would be interesting to see if this can be proven to be the case for all of Mochon's moves. There is another class of games that achieve arbitrarily low bias due to Pelchat and Høyer [21] whose moves are also worth investigating in this regard, and even otherwise.

3. *Trick to improve the precision of the EMA algorithm.* The algorithm tries to find a λ such that $\langle w| f_\lambda(X_h) |w\rangle - \langle v| f_\lambda(X_g) |v\rangle = 0$. In the finite case, one must also have for consistency, $\langle w|n_h(\lambda)\rangle = \langle v|n_g(\lambda)\rangle$ (this is because in subsequent steps, the space orthogonal is affected so if the component of the honest states along the normals is not mapped correctly, it would not get fixed later; this would mean there is no solution as we are only imposing necessary conditions). We observed that, numerically, we get a better precision if we use the latter condition for fine-tuning the result (after applying the former for obtaining a more course-grained solution). While analytically, the first condition implies the latter exactly, this ceases to be the case numerically due to the finiteness of precision. A careful error analysis of the EMA algorithm would be required to fully understand this behaviour. As a first step, we can understand this improvement as a direct consequence of the fact that the honest state is explicitly mapped correctly (up to the precision of the machine, which is about 16 floating points for most computers) if we use the method involving normals while in the latter, this should happen implicitly.

[6] To see this, note that the only time we spill over to the extra dimensions, is when we use the wiggle-v method. Otherwise, we stay inside the first $\max(n_g, n_h)$ dimensions.

Limitations of the current implementation.

1. *Limited wiggle-v.* We have not fully implemented the Wiggle-v method which means that it would be cumbersome to apply it to the general merge and split, for instance. However, for them we already give the explicit Blinkered Unitaries. For the rest, as we already saw, it does not even seem necessary.

2. *Minor pending issues.* Sometimes due to noise (arising from finiteness of the precision) our global minimiser gets trapped into local minima and has to be guided manually by looking at the graph. This means that a refined algorithm should also be able to solve the problem. Further, we did not implement the systematic method defined by the EMA algorithm for finding the spectrum of the matrices it uses but it appears that almost any guess works for Mochon's assignments.

An important open problem that remains is to account for noise in the quantum formalism itself. We assumed all along that the unitary is applied exactly. Now there are two possible effects of the unitary being noisy. The first effect is on the coordinates of the points. This can be accounted for by raising all the points a little bit (proportional to the noise). Doing this for each move would yield a cumulative effect on the final point, that is, the bias. This should not be very hard to evaluate. The second effect is on the weights of the points. This must be handled with care as most of the games rely on a cancellation of weight across different moves. One way to proceed would be to imagine, after a lossy operation, the remaining game/protocol is implemented with a slightly scaled down weight but with the same relative proportion. The rest of the weight should be collected in the end and merged with the final point to get a small raise. This would quantify the effect of this type of noise on the bias.

9 Acknowledgements

We are thankful to Nicolas Cerf, Mathieu Brandeho and Ognyan Oreshkov for various insightful discussions. We acknowledge support from the Belgian Fonds de la Recherche Scientifique – FNRS under grants no F.4515.16 (QUICTIME) and R.50.05.18.F (QuantAlgo). The QuantAlgo project has received funding from the QuantERA ERA-NET Cofund in Quantum Technologies implemented within the European Union's Horizon 2020 Programme. ASA further acknowledges the FNRS for support through the FRIA grant, 3/5/5 – MCF/XH/FC – 16754.

References

[1] C. Mochon, "Quantum weak coin flipping with arbitrarily small bias", arXiv:0711.4114 (2007).

[2] C. Mochon, "Large family of quantum weak coin-flipping protocols", Phys. Rev. A **72**, 022341 (2005).

[3] R. Cleve, "Limits on the security of coin flips when half the processors are faulty", in Proceedings of the eighteenth annual ACM symposium on theory of computing - STOC '86 (1986).

[4] A. Ambainis, "A new protocol and lower bounds for quantum coin flipping", Journal of Computer and System Sciences **68**, 398–416 (2004).

[5] A. Kitaev, "Quantum coin flipping", Talk at the 6th workshop on Quantum Information Processing, 2003.

[6] D. Aharonov, A. Chailloux, M. Ganz, I. Kerenidis and L. Magnin, "A simpler proof of existence of quantum weak coin flipping with arbitrarily small bias", SIAM Journal on Computing **45**, 633–679 (2014).

[7] A. Nayak, J. Sikora and L. Tunçel, "A search for quantum coin-flipping protocols using optimization techniques", Mathematical Programming **156**, 581–613 (2014).

[8] A. Nayak, J. Sikora and L. Tunçel, "Quantum and classical coin-flipping protocols based on bit-commitment and their point games", (2015).

[9] A. Chailloux and I. Kerenidis, "Optimal Quantum Strong Coin Flipping", in 50th focs (2009), pp. 527–533.

[10] A. Chailloux and I. Kerenidis, "Optimal Bounds for Quantum Bit Commitment", in 52nd focs (2011), pp. 354–362.

[11] A. Chailloux, G. Gutoski and J. Sikora, "Optimal bounds for semi-honest quantum oblivious transfer", Chicago Journal of Theoretical Computer Science, 2016 (2013).

[12] M. Ganz, "Quantum Leader Election", (2009).

[13] N. Aharon and J. Silman, "Quantum dice rolling: a multi-outcome generalization of quantum coin flipping", New Journal of Physics **12**, 033027 (2010).

[14] R. W. Spekkens and T. Rudolph, "A quantum protocol for cheat-sensitive weak coin flipping", Phys. Rev. Lett. vol 89, 227901 (2002) (2002).

[15] A. Nayak and P. Shor, "Bit-commitment-based quantum coin flipping", Phys. Rev. A **67**, 012304 (2003).

[16] I. Kerenidis and A. Nayak, "Weak coin flipping with small bias", Information Processing Letters **89**, 131–135 (2004).

[17] R. Bhatia, *Matrix analysis* (Springer New York, 1st Dec. 2013).

[18] T. Fritz, *Does the set of operator monotone functions become larger if we restrict ourselves to real symmetric matrices?*, (2018) https://mathoverflow.net/questions/298359/does-the-set-of-operator-monotone-functions-become-larger-if-we-restrict-ourselv.

[19] R. Schneider, *Convex Bodies: The Brunn-Minkowski Theory* (Cambridge University Press, 2009).

[20] A. S. Arora, J. Roland and S. Weis, *Weak Coin Flipping*, (2018) https://atulsingharora.github.io/WCF.

[21] P. Høyer and E. Pelchat, "Point Games in Quantum Weak Coin Flipping Protocols", MA thesis (University of Calgary, 2013).

A Blinkered $m \to n$ Transition

Recall that the unitary we 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, we need to complete a set containing one vector into a complete orthonormal basis. Let us 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 $|v\rangle = \frac{\sum_i \sqrt{p_i}|g_i\rangle}{\sqrt{\sum_i p_i}}$. We 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, we can do this inductively. Instead, we do this by examples for that makes it intuitive and demonstrates the generalisable argument right away. The first we 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 we 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 us 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 we can evaluate the vectors of interest

$$
\begin{aligned}
|v\rangle &= \frac{\sqrt{p_{g_1}}\,|g_{11}\rangle + \sqrt{p_{g_2}}\,|g_{22}\rangle + \sqrt{p_{g_3}}\,|g_{33}\rangle}{N_g} \\
|v_1\rangle &= \frac{\sqrt{p_{g_1}}\,|g_{11}\rangle - \frac{p_{g_1}}{\sqrt{p_{g_2}}}\,|g_{22}\rangle}{N_{g_1}} \\
|v_2\rangle &= \frac{\sqrt{p_{g_1}}\,|g_{11}\rangle + \sqrt{p_{g_2}}\,|g_{22}\rangle - \frac{\left(p_{g_1}+p_{g_2}\right)}{\sqrt{p_{g_3}}}\,|g_{33}\rangle}{N_{g_2}} \\
|w\rangle &= \frac{\sqrt{p_{h_1}}\,|h_{11}\rangle + \sqrt{p_{h_2}}\,|h_{22}\rangle}{N_h} \\
|w_1\rangle &= \frac{\sqrt{p_{h_2}}\,|h_{11}\rangle - \sqrt{p_{h_1}}\,|h_{22}\rangle}{N_h}
\end{aligned}
$$

where N_g, N_{g_1}, N_{g_2}, N_h are normalisations. In fact we want to express the constraints in this basis. To evaluate the first term we 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: We 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

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

We evaluate

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

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

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

For the second term, we 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 we 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 $U|g_{33}\rangle = \sqrt{p_{g_3}}\left(\frac{1}{N_g}|w\rangle + 0\,|v_1\rangle - \frac{1}{N'_{g_2}}|v_2\rangle\right)$ with

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

to get

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

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

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

Despite this appearing to be a complicated expression, we 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 we choose to write as

$$\begin{array}{c|cc|cc} & \langle v_1| & \langle v_2| & \langle w| & \langle w_1| \\ \hline |v_1\rangle & & & & \\ |v_2\rangle & \multicolumn{2}{c|}{C} & \multicolumn{2}{c}{B^T} \\ \hline |w\rangle & & & & \\ |w_1\rangle & \multicolumn{2}{c|}{B} & \multicolumn{2}{c}{A} \end{array} \geq 0.$$

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

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

We 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 $C^{-1} > 0$. We can therefore write the 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, we intend to use a more general technique (similar to the one used in the split analysis). Let us introduce

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

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

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

Note that this now has the exact same form as that of the split constraint with $x_{g_1} \to \langle x_g \rangle$. We 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 $\mathbb{I} \geq \langle x_g \rangle |w''\rangle \langle w''|$ where $|w''\rangle = \frac{\sqrt{\frac{p_{h_1}}{x_{h_1}}}|h_{11}\rangle + \sqrt{\frac{p_{h_2}}{x_{h_2}}}|h_{22}\rangle}{N_h}$. 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.

B 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. We prove this by induction (following Mochon's proof, just optimised for clarity instead of space). For $n = 2$

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

Now we show that if the result holds for $n - 1$ and it would also hold for n which would complete the inductive proof. We 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 we 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, we 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 we do this by induction on k. For $k = 0$ the result holds by the previous result. We assume it holds for order $k-1$ and show using this that it also holds for order k (this proof is also Mochon's). Let $g(x_i)$ be a polynomial of order $k-1$ s.t.

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

Notice that the first part of the sum disappears for all $1 \leq i \leq k$ because of the numerator. Consequently we 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. □

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 us define $d(n) := \sum_{i=1}^{n} \frac{x_i^{n-1}}{\prod_{j\neq i}(x_j - x_i)}$ to proceed inductively. We can then write

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

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

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

□

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. □