The extension problem in free harmonic analysis

Peter Burton and Kate Juschenko

March 10, 2020

Abstract

This paper studies certain aspects of harmonic analysis on the rank two free group. We focus on the concept of a positive definite function on the free group and our primary goal is to understand how such functions can be extended from balls of finite radius to the entire group. More specifically, we define a concept of 'relative energy' which measures the proximity between a pair of positive definite functions, and we ask whether a family of positive definite functions on a finite ball can be extended to the entire group with control on their relative energies. We find that the answer to this question depends on the configuration of relative energies that we seek to control, and that it has deep connections with classical harmonic analysis and with the recent refutation of Connes' embedding conjecture.

Contents

1 Introduction

In this paper we never consider free groups of rank higher than two, and we denote the free group of rank two by $\mathbb{F}$. We now present the definitions needed to formalize the discussion in the abstract and to state our main results.

1.1 Preliminary information

1.1.1 The fundamental inequality on the free group

Definition 1.1. *Let $d \in \mathbb{N}$ and let F be a finite subset of $\mathbb{F}$. We define a function $\mathsf{C} : F \to \mathrm{Mat}_{d\times d}(\mathbb{C})$ to be **positive definite** if we have the fundamental inequality*

$$\sum_{g,h\in E} \alpha(h)^* \mathsf{C}(h^{-1}g)\alpha(g) \geq 0 \tag{1.1}$$

for every subset E of $\mathbb{F}$ with $E^{-1}E \subseteq F$ and every function $\alpha : E \to \mathbb{C}^d$.

*We define C to be **strictly positive definite** if C is positive definite and the inequality in (1.1) is saturated only when α is identically 0. We define a function $\mathsf{C} : \mathbb{F} \to \mathrm{Mat}_{d\times d}(\mathbb{C})$ to be (strictly) positive definite if $\mathsf{C} \restriction F$ is (strictly) positive definite for every finite $F \subseteq \mathbb{F}$.*

A positive definite function on the free group can be thought of as a noncommutative analog of an infinite positive definite Toeplitz matrix.

1.1.2 The space of normalized strictly positive definite functions

We will always assume the following normalization condition.

Definition 1.2. *Let $\mathbf{I}_d$ denote the $d \times d$ identity matrix. If C is a positive definite function with values in $\mathrm{Mat}_{d\times d}(\mathbb{C})$ whose domain contains e, we define C to be **normalized** if $\mathsf{C}(e) = \mathbf{I}_d$.*

If C is normalized then for any fixed $g \in \mathbb{F}$ the vectors $\Phi_{\mathsf{C}}(g)_1, \ldots, \Phi_{\mathsf{C}}(g)_d$ are orthonormal. We denote the space of normalized strictly positive definite functions $\mathsf{C} : \mathbb{B}_r \to \mathrm{Mat}_{d\times d}(\mathbb{C})$ by $\mathrm{NSPD}(r, d)$. We endow the space of functions from $\mathbb{B}_r$ to $\mathrm{Mat}_{d\times d}(\mathbb{C})$ with the norm

$$||\mathsf{C}||_1 = \sum_{g\in\mathbb{B}_r} \sum_{j,k=1}^{d} |\mathsf{C}(g)_{j,k}|$$

1.1.3 Realizations of positive definite functions on balls

Note that $\mathbb{B}_r^{-1}\mathbb{B}_r = \mathbb{B}_{2r}$. Therefore if $\mathsf{C} \in \mathrm{NSPD}(2r, d)$ we can regard it as a positive definite kernel on the set $\mathbb{B}_r \times [d]$. By Theorem C.2.3 in [3] there exists a Hilbert space $\mathcal{X}(\mathsf{C})$ and a function $\Phi_{\mathsf{C}} : \mathbb{B}_r \to \mathcal{X}(\mathsf{C})^d$ such that

$$\langle \Phi_{\mathsf{C}}(g)_j, \Phi_{\mathsf{C}}(h)_k \rangle = \mathsf{C}(h^{-1}g)_{j,k} \tag{1.2}$$

for all $g, h \in \mathbb{B}_r$ and all $j, k \in [d]$. Moreover, we may and will assume that the coordinates of the range of Φ_{C} span $\mathcal{X}(\mathsf{C})$. The hypothesis that C is strictly positive definite ensures that the coordinates of the range of Φ_{C} will be linearly independent. We will refer to them as the canonical basis for $\mathcal{X}(\mathsf{C})$.

Definition 1.3. *We say that* $(\mathcal{X}(\mathsf{C}), \Phi_{\mathsf{C}})$ *as above is a* ***realization*** *of* C.

We can construct a realization of a positive definite function $\mathsf{C} : \mathbb{F} \to \mathrm{Mat}_{d\times d}(\mathbb{C})$ in the same way, obtaining a Hilbert space $\mathcal{X}(\mathsf{C})$ and a function $\Phi_{\mathsf{C}} : \mathbb{F} \to \mathcal{X}(\mathsf{C})^d$ such that the span of the coordinates of the range of Φ_{C} is dense in $\mathcal{X}(\mathsf{C})$.

It is clear that given two realizations of the same positive definite function there exists a natural unitary isomorphism from one realization Hilbert space to the other. This isomorphism transforms a canonical basis vector in one realization to the canonical basis vector in another realization having the same index. If $\mathsf{C} : \mathbb{F} \to \mathrm{Mat}_{d\times d}(\mathbb{C})$ is positive definite then the function $g \mapsto (\Phi_{\mathsf{C}}(hg)_1, \ldots, \Phi_{\mathsf{C}}(hg)_d)$ is a realization of Φ_{C} for any $h \in \mathbb{F}$. Thus we may make the following definition.

Definition 1.4. *Let* $d \in \mathbb{N}$ *and let* $\mathsf{C} : \mathbb{F} \to \mathrm{Mat}_{d\times d}(\mathbb{C})$ *be positive definite. Then any realization of* C *defines an* ***associated unitary representation*** *of* $\mathbb{F}$ *on* $\mathcal{X}(\mathsf{C})$ *denoted by* ρ_{C} *and given by the translation* $\rho_{\mathsf{C}}(h)\Phi_{\mathsf{C}}(g)_j = \Phi_{\mathsf{C}}(hg)_j$ *for* $g, h \in \mathbb{F}$ *and* $j \in [d]$.

1.1.4 Transport operators and relative energies

Definition 1.5. *Let* $\mathsf{C}, \mathsf{D} \in \mathrm{NSPD}(2r, d)$. *Let* $(\mathcal{X}(\mathsf{C}), \Phi_{\mathsf{C}})$ *and* $(\mathcal{X}(\mathsf{D}), \Phi_{\mathsf{D}})$ *be realizations of* C *and* D *respectively. Define the* ***transport operator*** $t[\mathsf{C}, \mathsf{D}] : \mathcal{X}(\mathsf{C}) \to \mathcal{X}(\mathsf{D})$ *by setting*

$$t[\mathsf{C}, \mathsf{D}] \sum_{g \in \mathbb{B}_r} \sum_{j=1}^{d} \alpha(g)_j \Phi_{\mathsf{C}}(g)_j = \sum_{g \in \mathbb{B}_r} \sum_{j=1}^{d} \alpha(g)_j \Phi_{\mathsf{D}}(g)_j$$

for functions $\alpha : \mathbb{B}_r \to \mathbb{C}^d$. *We refer to the square of the operator norm of* $t[\mathsf{C}, \mathsf{D}]$ *as the* ***relative energy*** *of the pair* (C, D) *and denote it by* $\mathfrak{e}(\mathsf{C}, \mathsf{D})$.

If $\mathsf{C}, \mathsf{D} : \mathbb{F} \to \mathrm{Mat}_{d\times d}(\mathbb{C})$ *are strictly positive definite we define the relative energy of the pair* (C, D) *to be* $\sup_{r\in\mathbb{N}} \mathfrak{e}(\mathsf{C} \restriction \mathbb{B}_r, \mathsf{D} \restriction \mathbb{B}_r)$. *We continue to denote it by* $\mathfrak{e}(\mathsf{C}, \mathsf{D})$. *In general we may have* $\mathfrak{e}(\mathsf{C}, \mathsf{D}) = \infty$. *If* $\mathfrak{e}(\mathsf{C}, \mathsf{D}) < \infty$ *then there is a naturally defined transport operator from* $\mathcal{X}(\mathsf{C})$ *to* $\mathcal{X}(\mathsf{D})$, *which we continue to denote by* $t[\mathsf{C}, \mathsf{D}]$.

The relevance of Definition 1.5 is that the transport operator between two strictly positive definite functions defined on all of $\mathbb{F}$ clearly intertwines the associated unitary representations. Thus transport operators will be useful in analyzing commuting representations of $\mathbb{F}$.

Note that the normalization hypotheses implies $\mathfrak{e}(\mathsf{C}, \mathsf{D}) \geq 1$ for all C, D and $\mathfrak{e}(\mathsf{C}, \mathsf{D}) = 1$ if and only if $\mathsf{C} = \mathsf{D}$.

1.2 Main results and discussion

Definition 1.6. *Let* $\Theta = (V, E)$ *be a finite directed graph and let* $r, d \in \mathbb{N}$. *We define a* ***d-dimensional radius-r configuration over*** Θ *to be a family* $(\mathsf{C}_v)_{v\in V}$ *of elements of* $\mathrm{NSPD}(r, d)$ *indexed by* V.

The following definitions are the central concepts of this paper, and our study indicates they are the appropriate way to formalize the extension problem.

Definition 1.7. *Given $\epsilon > 0$ and a d-dimensional radius-r configuration $\mathbf{C} = (\mathsf{C}_v)_{v\in V}$ over a finite directed graph Θ we define the* ***ϵ-malleable extension energy cost*** *of $\mathbf{C}$ to be the infimum of all numbers $M \in [1,\infty)$ such that there exists for each $v \in V$ a positive definite function $\widehat{\mathsf{C}}_v : \mathbb{F} \to \mathrm{Mat}_{d\times d}(\mathbb{C})$ with*

$$\max\Big(\mathfrak{e}(\mathsf{C}_v, \widehat{\mathsf{C}}_v \upharpoonright \mathbb{B}_r), \mathfrak{e}(\widehat{\mathsf{C}}_v \upharpoonright \mathbb{B}_r, \mathsf{C}_v)\Big) \leq 1 + \epsilon$$

and

$$\mathfrak{e}(\widehat{\mathsf{C}}_v, \widehat{\mathsf{C}}_w) - 1 \leq M\Big(\mathfrak{e}(\mathsf{C}_v, \mathsf{C}_w) - 1\Big) \tag{1.3}$$

for all $(v, w) \in E$. We denote this by $\mathrm{encost}_\epsilon(\mathbf{C})$. We adopt the convention that $\inf \emptyset = \infty$ in the sense that we set $\mathrm{encost}_\epsilon(\mathbf{C}) = \infty$ when there does not exist M as in (1.3).

We define the ***ϵ-malleable radius-r extension energy cost*** *of Θ to be $\sup_{\mathbf{C}} \mathrm{encost}_\epsilon(\mathbf{C})$, where the supremum is taken over all $d \in \mathbb{N}$ and all d-dimensional radius-r configurations $\mathbf{C}$ over Θ. We denote this by $\mathrm{encost}_{r,\epsilon}(\Theta)$. We define the* ***$(r, \epsilon)$-extension energy cost*** *of the free group to be $\sup_\Theta \mathrm{encost}_{r,\epsilon}(\Theta)$, where the supremum is taken over all finite directed graphs Θ. We denote this by $\mathrm{encost}_{r,\epsilon}(\mathbb{F})$.*

In order for Definition 1.7 to be reasonable we ought to know that any element of $\mathrm{NSPD}(r, d)$ admits an extension to a positive definite function defined on all of $\mathbb{F}$. This fact appears as Proposition 4.4 in [1] and as Lemma 25 in [13].

Our main results are the following.

Theorem 1.1. *If Θ is a finite tree directed toward a root or a finite directed cycle then $\mathrm{encost}_{r,\epsilon}(\Theta) = 1$ for all $r \in \mathbb{N}$ and all $\epsilon > 0$.*

Theorem 1.2. *There exist $r \in \mathbb{N}$ and $\epsilon > 0$ such that $\mathrm{encost}_{r,\epsilon}(\mathbb{F}) = \infty$. Indeed, for these r and ϵ we have*

$$\sup\Big\{\mathrm{encost}_{r,\epsilon}(\Theta) : \Theta \text{ is a finite 4-regular directed graph}\Big\} = \infty \tag{1.4}$$

We remark that the existence of a sequence of finitely supported positive definite functions on $\mathbb{Z}$ converging to the constant 1 from below ensures that after appropriately modifying Definition 1.7 we find $\mathrm{encost}_{r,\epsilon}(\mathbb{Z}) < \infty$ for all $r \in \mathbb{N}$ and $\epsilon > 0$.

Section 2 will be dedicated to the proof of Theorem 1.1. This proof proceeds by giving an explicit recursive procedure for constructing extensions of positive definite functions over a tree or a cycle while controlling relative energies. This procedure is based on an analysis of noncommutative Szegö parameters that builds on the work in [1].

Section 3 we will be dedicated to the proof of Theorem 1.2. The starting point for this part of the proof is the negation of Connes' embedding conjecture, which was recently established in the breakthrough paper [7].

Theorems 1.1 and 1.2 constitute the full extent of the authors' understanding of extension energy costs, aside from the minor fact that methods of the present paper can be easily modified to show that 4 in (1.4) can be replaced with 3. Beyond these results, energy extension costs seem somewhat mysterious. For example, if we fix the minimal r as in Theorem 1.2 then it is trivial that $\mathrm{encost}_{r,A}(\mathbb{F}) = 1$ for sufficiently large finite values of A. Therefore there exists a maximal positive value of ϵ such that $\mathrm{encost}_{r,\epsilon}(\mathbb{F}) = \infty$. The exact value of this minimal r and maximal ϵ can be thought of as a kind of phase transition point, and they seem extremely difficult to compute. We also observe that the examples in Theorem 1.1 show that the existence of cycles and vertexes of high degree are not in themselves obstructions a graph to having low extension energy cost, so it is unclear what to say about the graph theoretic implications of extension energy costs.

One way to gain insight into the last issue would be to obtain a more concrete understanding of the graphs constructed to prove Theorem 1.2. Our approach constructs these graphs directly from the negation of the statement of Connes' embedding conjecture. By using more information about a counterexample to Connes' embedding conjecture it might be possible to obtain more details about these graphs.

1.3 Notation

1.3.1 The free group

Fix a pair of free generators a and b for $\mathbb{F}$ and endow $\mathbb{F}$ with the standard Cayley graph structure corresponding to left multiplication by these generators. If Γ is a quotient of $\mathbb{F}$ we identify a and b with their images in Γ. We write e for the identity of $\mathbb{F}$. We will also use the symbol e for $2.718\ldots$

We consider the word length associated to a and b, which we denote by $|\cdot|$. For $r \in \mathbb{N}$ let $\mathbb{B}_r = \{g \in \mathbb{F} : |g| \leq r\}$ be the ball of radius r around e. Write K_r for the cardinality of $\mathbb{B}_r$.

We define an ordering $\preceq$ on the sphere of radius 1 in $\mathbb{F}$ by setting $a \preceq b \preceq a^{-1} \preceq b^{-1}$. From this we obtain a corresponding shortlex linear ordering on all of $\mathbb{F}$, which we continue to denote by $\preceq$. For $g \in \mathbb{F}$ define $\mathcal{I}_g = \bigcup\{\{h, h^{-1}\} : h \preceq g\}$. Define a generalized Cayley graph $\mathrm{Cay}(\mathbb{F}, g)$ with vertex set equal to $\mathbb{F}$ by placing an edge between distinct elements h and ℓ if and only if $\ell^{-1}h \in \mathcal{I}_g$. Write $g_\uparrow$ for the immediate predecessor of g in $\preceq$ and $g_\downarrow$ for the immediate successor of g in $\preceq$.

1.3.2 Miscellanea

If z and w are complex numbers and $\epsilon > 0$ we will sometimes write $z \approx[\epsilon] w$ to mean $|z - w| \leq \epsilon$.

We write $\mathbb{D}$ for the open unit disk in the complex plane.

If $n \in \mathbb{N}$ we write $[n]$ for $\{1, \ldots, n\}$.

It is customary in analysis to collect accumulating constants into a single notation, such as $O(\cdot)$ or a symbol such as C which can take on multiple values. We do employ $O(\cdot)$ and $o(\cdot)$ notation in Section 2. However, we choose to keep track explicitly of all constants in Section 3. Despite the fact that their specific values are unimportant, we feel that checking how these constants vary from line to line aids in verification of the proof.

1.4 Acknowledgements

We thank Lewis Bowen for several suggestions that improved the writing. We thank Rostyslav Kravchenko for suggesting the use of Proposition 2.11.

2 Proof of Theorem 1.1

In Section 2 we prove Theorem 1.1

2.1 Review of certain aspects of classical theory

In Subsection 2.1 we recall some information about classical harmonic analysis which will be relevant to our later arguments.

2.1.1 The fundamental inequality on the integers

Harmonic analysis typically begins with Fourier analysis. The theory of Fourier series on the unit circle $\mathbb{T}$ can be thought of as harmonic analysis on the additive group of integers $\mathbb{Z}$, which is the free group on one generator. We now discuss some aspects of this classical theory which are relevant to the method we use in Section 2.

We define a marked unitary representation of a countable discrete group G to be a unitary representation ρ of G on a Hilbert space $\mathcal{X}$ together with a distinguished vector $x \in \mathcal{X}$ which is cyclic in the sense that the span of the set $\{\rho(g)x : g \in G\}$ is dense in $\mathcal{X}$. Any unitary representation of a countable discrete group G can be decomposed into a countable orthogonal sum of subrepresentations, each of which admits a cyclic vector.

The family of marked unitary representations of $\mathbb{Z}$ is in canonical one to one correspondence with the the family of finite Borel measures on $\mathbb{T}$. This correspondence is given by placing both of these families of objects in canonical one to one correspondence with a third family of objects. This third family consists of so-called positive definite functions on $\mathbb{Z}$. A positive definite function C on $\mathbb{Z}$ is a function from $\mathbb{Z}$ to the complex numbers which satisfies the fundamental inequality

$$\sum_{m,n=-N}^{N} \alpha(m)\overline{\alpha(n)}C(m-n) \geq 0 \tag{2.1}$$

for every natural number N and every function $\alpha : \{-N, \dots N\} \to \mathbb{C}$. One can think of a positive definite function on $\mathbb{Z}$ as an infinite version of a positive definite Toeplitz matrix.

2.1.2 Fourier correspondence

The correspondence between a positive definite function C on $\mathbb{Z}$ and a finite Borel measure μ on $\mathbb{T}$ is given by the formula

$$C(n) = \widehat{\mu}(n) = \int_{\mathbb{T}} s^{-n} \, \mathrm{d}\mu(s) \tag{2.2}$$

Thus the values of C are the Fourier coefficients of μ. The fact that the sequence of Fourier coefficients of a measure is positive definite reflects the following fact, which can be verified with elementary computation.

$$\sum_{m,n=-N}^{N} \alpha(m)\overline{\alpha(n)}C(m-n) = \sum_{m,n=-N}^{N} \alpha(m)\overline{\alpha(n)} \int_{\mathbb{T}} s^{n-m} \, \mathrm{d}\mu(s)$$

$$= \int_{\mathbb{T}} \left| \sum_{m=-N}^{N} \alpha(m) s^{-m} \right|^2 \, \mathrm{d}\mu(s) \geq 0$$

The fact that a positive definite function on $\mathbb{Z}$ uniquely defines a finite Borel measure on $\mathbb{T}$ via (2.2) is known as Bôchner's theorem.

2.1.3 Spectral correspondence

The correspondence between a positive definite function C on $\mathbb{Z}$ and a marked unitary representation (ρ, x) of $\mathbb{Z}$ is given as follows. Let $u = \rho(1)$ be the unitary operator corresponding to the unique free generator of $\mathbb{Z}$. Then we set

$$C(n) = \langle u^n x, x \rangle \tag{2.3}$$

Thus the values of C are the matrix coefficients of the marked unitary representation. The fact that a sequence of unitary matrix coefficients is positive definite reflects the following fact, which can be verified with elementary computation.

$$\begin{aligned} \sum_{m,n=-N}^{N} \alpha(m)\overline{\alpha(n)} C(m-n) &= \sum_{m,n=-N}^{N} \alpha(m)\overline{\alpha(n)} \langle u^{m-n} x, x \rangle \\ &= \left\| \sum_{m=-N}^{N} \alpha(m) u^m x \right\|^2 \geq 0 \end{aligned}$$

The fact that a positive definite function on $\mathbb{Z}$ uniquely defines a marked unitary representation of $\mathbb{Z}$ via (2.3) is a version of the spectral theorem.

2.1.4 Constructing positive definite functions on the integers

In Subsection 2.1.4 we describe a method for recursively constructing positive definite functions on the integers. A noncommutative version of this construction is the main topic of Section 2

Note that if $m, n \in \{0, \ldots, N\}$ then we have $m - n \in \{-N, \ldots, N\}$. Suppose we have defined a function $C : \{-N, \ldots, N\} \to \mathbb{C}$ which satisfies the restricted fundamental inequality

$$\sum_{m,n=0}^{N} \alpha(m)\overline{\alpha(n)} C(m-n) \geq 0 \tag{2.4}$$

for every function $\alpha : \{0, \ldots, N\} \to \mathbb{C}$. We refer to C as a partially defined positive definite function. We wish to extend C to a partially defined positive definite function defined on $\{-N-1, \ldots, N+1\}$. In order to do so it suffices to specify the number $C(N+1)$, as then we must have $C(-N-1) = \overline{C(N+1)}$.

We will assume that C is strictly positive definite in the sense that the inequality in (2.4) is saturated only when α is identically zero. Standard theory then implies that we can find a Hilbert space $\mathcal{X}$ of dimension $N+1$ and canonical basis vectors $\Phi_0, \ldots, \Phi_N \in \mathcal{X}$ such that

$$\langle \Phi_m, \Phi_n \rangle = C(m-n) \tag{2.5}$$

for all $m, n \in \{0, \ldots, N\}$. We can also define a shifted Hilbert space $\mathcal{Y}$ of dimension $N+1$ with canonical basis vectors $\Phi_1, \ldots, \Phi_{N+1}$ satisfying (2.5) for all $m, n \in \{1, \ldots, N+1\}$. Let $\mathcal{Z}$ be the vector space consisting of the span of $\mathcal{X}$ and $\mathcal{Y}$, so that $\mathcal{Z}$ has a canonical basis $\Phi_0, \ldots, \Phi_{N+1}$.

In the vector space $\mathcal{Z}$ the inner product between Φ_0 and Φ_{N+1} is not defined. However, the inner products between all other pairs of elements of the canonical basis for $\mathcal{Z}$ are defined. This allows us to apply the Gram-Schmidt procedure to $\Phi_1, \ldots, \Phi_N$ to obtain an orthonormal basis $\mathcal{B}$ for the subspace

$$\mathcal{X} \cap \mathcal{Y} = \operatorname{span}(\Phi_1, \ldots, \Phi_N) \tag{2.6}$$

of $\mathcal{Z}$. Moreover, we can compute the inner products between Φ_0 and the elements of $\mathcal{B}$ and we can compute the inner products between Φ_{N+1} and the elements of $\mathcal{B}$. Therefore the orthogonal projection p from $\mathcal{Z}$ onto the subspace in (2.6) is well defined.

Write I for the identity operator on $\mathcal{Z}$. Then for any complex number ζ with $|\zeta| \leq 1$ we can set

$$\left\langle \frac{(I-p)\Phi_{N+1}}{||(I-p)\Phi_{N+1}||}, \frac{(I-p)\Phi_0}{||(I-p)\Phi_0||} \right\rangle = \zeta \tag{2.7}$$

The hypothesis that C is strictly positive definite ensures the denominators of the fractions in (2.7) are nonzero. The hypothesis that $|\zeta| \leq 1$ reflects to the need to satisfy the Cauchy-Schwartz inequality. Once we have chosen ζ, the space $\mathcal{Z}$ is promoted to a full Hilbert space. The fact that any choice of ζ with $|\zeta| \leq 1$ produces a valid positive definite extension follows from the observation that no matter the value of ζ we have a decomposition

$$\mathcal{Z} = (\mathcal{X} \cap \mathcal{Y}) \oplus \operatorname{span}\big((I-p)\Phi_{N+1}, (I-p)\Phi_0\big) \tag{2.8}$$

From (2.7) we can recover

$$\begin{aligned} C(N+1) &= \langle \Phi_{N+1}, \Phi_0 \rangle \\ &= \zeta ||(I-p)\Phi_{N+1}|| \, ||(I-p)\Phi_0|| + \langle p\Phi_{N+1}, \Phi_0 \rangle + \langle \Phi_{N+1}, p\Phi_0 \rangle - \langle p\Phi_{N+1}, p\Phi_0 \rangle \\ &= \zeta ||(I-p)\Phi_{N+1}|| \, ||(I-p)\Phi_0|| + \langle p\Phi_{N+1}, p\Phi_0 \rangle \end{aligned} \tag{2.9}$$

The norms and inner products in (2.9) are determined by C, so ζ is indeed the only free parameter. Thus the set of legal possibilities for $C(N+1)$ is a closed disk in $\mathbb{C}$ of radius $||(I-p)\Phi_{N+1}|| \, ||(I-p)\Phi_0||$ centered at the point $\langle p\Phi_{N+1}, p\Phi_0 \rangle$. The numbers $||(I-p)\Psi_0||$ and $||(I-p)\Phi_{N+1}||$ are bounded by $||\Phi_0|| = ||\Phi_{N+1}|| = C(0)$. Moreover, these numbers will be small if the numbers $||p\Phi_0||$ and $||p\Phi_{N+1}||$ are close to $C(0)$. Since p is the orthogonal projection onto $\mathcal{X} \cap \mathcal{Y}$, we can interpret this as indicating that if Φ_0 and Φ_{N+1} are close to $\mathcal{X} \cap \mathcal{Y}$ then the possibilities for $C(N+1)$ are more restricted. The number ζ is typically referred to as a Szegö parameter.

2.2 Constructing positive definite functions on the free group

In Subsection 2.2 we describe the general procedure for constructing positive definite functions on the free group. This will allow us to state Lemmas 2.1, 2.2 and 2.3 in Segment 2.2.9 below. These three lemmas easily imply Theorem 1.1, and will be proved in Subsections 2.3, 2.4 and 2.5 respectively.

2.2.1 Geometry of generalized Cayley graphs on the free group

In Segment 2.2.1 we describe certain geometric features of the graphs $\mathrm{Cay}(\mathbb{F}, g)$ which will be necessary for the construction of positive definite functions. We will use two known results about generalized Cayley graphs on $\mathbb{F}$.

Recall that a graph is said to be chordal if every induced cycle has length at most 3. The next fact appears as Proposition 3.2 in [1] and as Proposition 3.6.7 in [2].

Proposition 2.1 (Bakonyi, Timotin)**.** *Let $g \in \mathbb{F}$. Then $\mathrm{Cay}(\mathbb{F}, g)$ is chordal.*

Let $g \in \mathbb{F}$ and suppose $\mathcal{K}$ is a clique in $\mathrm{Cay}(\mathbb{F}, g)$ which is not a clique in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$. We note that $\mathcal{K}$ is maximal among all cliques in $\mathrm{Cay}(\mathbb{F}, g)$ if and only if $\mathcal{K}$ is maximal among those cliques in $\mathrm{Cay}(\mathbb{F}, g)$ which are not cliques in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$. The following appears as Corollary 3.3 in [1] and as Corollary 3.6.8 in [2].

Proposition 2.2 (Bakonyi, Timotin)**.** *Let $g \in \mathbb{F}$ and let $\mathcal{K}$ be a maximal clique in $\mathrm{Cay}(\mathbb{F}, g)$ which is not a clique in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$. Then there exists a unique edge in $\mathcal{K}$ which is not an edge in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$.*

The next two propositions are implicit in [1] but we include proofs for completeness.

Proposition 2.3. *Let $g \in \mathbb{F}$. Then there exists a unique maximal clique $\mathcal{K}_g$ in $\mathrm{Cay}(\mathbb{F}, g)$ which contains the edge (e, g).*

Proof of Proposition 2.3. Suppose toward a contradiction that Proposition 2.3 fails. Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two distinct maximal cliques in $\mathrm{Cay}(\mathbb{F}, g)$ which contain (e, g). Since $\mathcal{M}_1$ and $\mathcal{M}_2$ contain (e, g), they are not cliques in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$. Therefore Proposition 2.2 implies that for each $j \in \{1, 2\}$ the edge (e, g) is the unique edge in $\mathcal{M}_j$ which is not an edge in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$. Thus for each $j \in \{1, 2\}$ and every $m \in \mathcal{M}_j \setminus \{g, e\}$ we have that (m, g) and (m, e) are edges of $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$.

Since $\mathcal{M}_1$ and $\mathcal{M}_2$ are distinct and maximal, it must be the case that their union is not a clique in $\mathrm{Cay}(\mathbb{F}, g)$. Therefore we can choose $m_1 \in \mathcal{M}_1 \setminus \mathcal{M}_2$ and m_2 in $\mathcal{M}_2 \setminus \mathcal{M}_1$ such that (m_1, m_2) is not an edge in $\mathrm{Cay}(\mathbb{F}, g)$. Since $\{g, e\} \subseteq \mathcal{M}_1 \cap \mathcal{M}_2$ for each $j \in \{1, 2\}$ we have $m_j \notin \{g, e\}$. Again using the fact that (g, e) is the unique edge added to $\mathcal{M}_1$ and $\mathcal{M}_2$ when passing from $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$ to $\mathrm{Cay}(\mathbb{F}, g)$ we see that (m_1, m_2) is not an edge in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$.

Consider the path $e \to m_1 \to g \to m_2 \to e$. The previous paragraphs show that this is an induced cycle in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$, so we obtain a contradiction to Proposition 2.1. □

Proposition 2.4. *Suppose $\mathcal{M}$ is a maximal clique in $\mathrm{Cay}(\mathbb{F}, g)$ which is not a clique in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$. Then $\mathcal{M}$ is a translate of $\mathcal{K}_g$.*

Proof of Proposition 2.4. Since $\mathcal{M}$ is not a clique in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$ it must contain a pair of vertexes h, ℓ such that $\ell^{-1}h \notin \mathcal{I}_{g_\uparrow}$. Since g is the unique element of $\mathcal{I}_g \setminus \mathcal{I}_{g_\uparrow}$ and $\mathcal{M}$ is a clique in $\mathrm{Cay}(\mathbb{F}, g)$ this implies that $\ell^{-1}h = g$. Thus by translating we may assume that $h = g$ and $\ell = e$ and so Proposition 2.4 follows from Proposition 2.3. □

The next proposition is necessary because it may happen that $g_\uparrow \notin \mathcal{K}_g$.

Proposition 2.5. *Let $g \in \mathbb{F}$. Then there exists $h \in \mathcal{I}_g$ such that $\mathcal{K}_g \setminus \{g\}$ is contained in a translate of $\mathcal{K}_h$.*

Proof of Proposition 2.5. Proposition 2.1 implies that (e, g) is the unique edge in $\mathcal{K}_g$ which is not an edge in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$. Therefore we have that $\mathcal{K}_g \setminus \{g\}$ is a clique in $\mathrm{Cay}(\mathbb{F}, g_\uparrow)$. Let h be the $\preceq$-least element such that $\mathcal{K}_g \setminus \{g\}$ is a clique in $\mathrm{Cay}(\mathbb{F}, h)$. Let $\mathcal{M}$ be a maximal clique in $\mathrm{Cay}(\mathbb{F}, h)$ which contains $\mathcal{K}_g \setminus \{g\}$. If $\mathcal{M}$ were a clique in $\mathrm{Cay}(\mathbb{F}, h_\uparrow)$ then $\mathcal{K}_g \setminus \{g\}$ would be a clique in $\mathrm{Cay}(\mathbb{F}, h_\uparrow)$, contradicting our choice of h. Thus Proposition 2.3 implies $\mathcal{M}$ is contained in a translate of $\mathcal{K}_h$. □

2.2.2 Full Hilbert space realizations of partial positive definite functions

Write $\mathrm{NSPD}(g, d)$ for the space of normalized strictly positive definite functions from $\mathcal{I}_g$ to $\mathrm{Mat}_{d\times d}(\mathbb{C})$. We refer to such a function as a partial positive definite function. Let $\mathsf{C} \in \mathrm{NSPD}(g, d)$. If $\mathcal{J}$ is a clique in $\mathrm{Cay}(\mathbb{F}, g)$, there exists a Hilbert space $\mathcal{X}(\mathsf{C}, \mathcal{J})$ and a function $\Phi_\mathsf{C} : \mathcal{J} \to \mathcal{X}(\mathsf{C})^d$ such that

$$\langle \Phi_\mathsf{C}(h)_j, \Phi_\mathsf{C}(\ell)_k\rangle = \mathsf{C}(\ell^{-1}h)_{j,k} \tag{2.10}$$

for all $h, \ell \in \mathcal{J}$ and $j, k \in [d]$. The assertion that $\mathcal{J}$ is a clique in $\mathrm{Cay}(\mathbb{F}, g)$ is equivalent to the assertion that $\mathcal{J}^{-1}\mathcal{J} \subseteq \mathcal{I}_g$. The relevance of this hypothesis is that it ensures the inner products between all pairs of elements of the range of Φ are defined. If we require that the set

$$\{\Phi_\mathsf{C}(h)_j : h \in \mathcal{J}, j \in [d]\}$$

spans $\mathcal{X}(\mathsf{C}, \mathcal{J})$ then these data are unique up to a unique unitary isomorphism. We write $\mathcal{X}(\mathsf{C})$ for $\mathcal{X}(\mathsf{C}, \mathcal{K}_g)$.

2.2.3 Partial Hilbert space realizations of partial positive definite functions

The following definition makes precise the construction of the space $\mathcal{Z}$ from Subsection 2.1.4.

Definition 2.1. *A **partial Hilbert space** is a vector space $\mathcal{V}$ together with a pair of distinguished subspaces $\mathcal{V}_1$ and $\mathcal{V}_2$ of $\mathcal{V}$ having the following properties.*

- *The subspaces $\mathcal{V}_1$ and $\mathcal{V}_2$ span $\mathcal{V}$.*
- *Each $\mathcal{V}_m$ is a Hilbert space with inner product $\langle\cdot,\cdot\rangle_m$.*
- *The inner products on $\mathcal{V}_1$ and $\mathcal{V}_2$ are compatible in the sense that for any pair of vectors $x, y \in \mathcal{V}_1 \cap \mathcal{V}_2$ we have $\langle x, y\rangle_1 = \langle x, y\rangle_2$.*

*We refer to $\mathcal{V}_1 \cap \mathcal{V}_2$ as the **core** of $\mathcal{V}$ and denote it by* $\mathrm{core}(\mathcal{V})$.

The following proposition is the key to the extension procedure.

Proposition 2.6. *If $\mathcal{V}$ is a partial Hilbert space then the orthogonal projection from $\mathcal{V}$ onto* $\mathrm{core}(\mathcal{V})$ *is well-defined.*

Proof of Proposition 2.6. The third item in Definition 2.1 ensures that we can apply the Gram-Schmidt procedure to obtain an orthonormal basis for $\mathrm{core}(\mathcal{V})$. The first and second items in Definition 2.1 imply that we can compute the inner products between an arbitrary element of $\mathcal{V}$ and and element of $\mathrm{core}(\mathcal{V})$. Therefore Proposition 2.6 follows. □

We now describe how to associate a partial Hilbert space to a partial positive definite function. This partial Hilbert space will be used to calculate the set of legal possibilities for an extension of an element of $\text{NSPD}(g_\uparrow, d)$ to an element of $\text{NSPD}(g, d)$.

Let $\mathsf{C} \in \text{NSPD}(g_\uparrow, d)$. Proposition 2.2 implies the clique $\mathcal{K}_g$ in $\text{Cay}(\mathbb{F}, g)$ contains a unique edge which is not an edge in $\text{Cay}(\mathbb{F}, g_\uparrow)$. This must be the edge between g and e. Therefore the sets $\mathcal{K}_g \setminus \{g\}$ and $\mathcal{K}_g \setminus \{e\}$ are cliques in $\text{Cay}(\mathbb{F}, g_\uparrow)$. Hence the discussion in Segment 2.2.2 implies that we can construct two Hilbert spaces $\mathcal{X}(\mathsf{C}, \mathcal{K}_g \setminus \{g\})$ and $\mathcal{X}(\mathsf{C}, \mathcal{K}_g \setminus \{e\})$ together with functions

$$\Phi_{\mathsf{C}} : \mathcal{K}_g \setminus \{g\} \to \mathcal{X}(\mathsf{C}, \mathcal{K}_g \setminus \{g\})^d$$

and

$$\Psi_{\mathsf{C}} : \mathcal{K}_g \setminus \{e\} \to \mathcal{X}(\mathsf{C}, \mathcal{K}_g \setminus \{e\})^d$$

such that Φ_{C} and Ψ_{C} satisfy (2.10) on their domains.

Proposition 2.7. *Let $\mathsf{C} \in \text{NSPD}(g_\uparrow, d)$. Then there exists a partial Hilbert space $\mathcal{X}(\mathsf{C})_\bullet$ with distinguished subspaces that can be identified with $\mathcal{X}(\mathsf{C}, \mathcal{K}_g \setminus \{g\})$ and $\mathcal{X}(\mathsf{C}, \mathcal{K}_g \setminus \{e\})$. Moreover, if $h \in \mathcal{K}_g \setminus \{g, e\}$ then $\Psi_{\mathsf{C}}(h)_j = \Phi_{\mathsf{C}}(h)_j$ for all $j \in [d]$ and $\text{core}(\mathcal{X}(\mathsf{C})_\bullet)$ consists exactly of the span of these vectors.*

Proof of Proposition 2.7. We take $\mathcal{X}(\mathsf{C})_\bullet$ to be the quotient of the disjoint union of $\mathcal{X}(\mathsf{C}, \mathcal{K}_g \setminus \{g\})$ and $\mathcal{X}(\mathsf{C}, \mathcal{K}_g \setminus \{e\})$ by the equivalence relation which identifies $\Phi_{\mathsf{C}}(h)_j$ and $\Psi_{\mathsf{C}}(h)_j$ for $h \in \mathcal{K}_g \setminus \{g, e\}$ and $j \in [d]$. □

If $\mathsf{C} \in \text{NSPD}(g_\uparrow, d)$ we define

$$\Theta_{\mathsf{C}}(h)_m = \begin{cases} \Phi_{\mathsf{C}}(h)_j = \Psi_{\mathsf{C}}(h)_j \text{ if } h \in \mathcal{K}_g \setminus \{g, e\} \text{ and } j \in [d] \\ \Psi_{\mathsf{C}}(h)_j \text{ if } h = g \text{ and } j \in [d] \\ \Phi_{\mathsf{C}}(h)_j \text{ if } h = e \text{ and } j \in [d] \end{cases} \tag{2.11}$$

Thus $\{\Theta_{\mathsf{C}}(h)_m : h \in \mathcal{K}_g, j \in [d]\}$ forms a canonical basis for $\mathcal{X}(\mathsf{C})_\bullet$.

2.2.4 Partial positive definite functions with partial matrix specification

Segment 2.2.4 is purely notational. For $(j, k) \in [d]^2$ we adopt the notation $\text{NSPD}(g, d, j, k)$ for elements C of $\text{NSPD}(g_\uparrow, d)$ such that the entries $\mathsf{C}(g)_{l,m}$ have been specified for all pairs $(l, m) < (j, k)$. Write

$$\mathcal{D}(g, d, j, k) = (\mathcal{I}_{g_\uparrow} \times [d]^2) \cup \big(\{g\} \times \{(l, m) \in [d]^2 : (l, m) < (j, k)\}\big)$$

Thus we can regard an element of $\text{NSPD}(g, d, j, k)$ as a function from $\mathcal{D}(g, d, j, k)$ to $\mathbb{C}$. We endow the space of functions from $\mathcal{D}(g, d, j, k)$ to $\mathbb{C}$ with the norm

$$||\mathsf{C}||_1 = \sum_{(h,l,m) \in \mathcal{D}(g,d,j,k)} |\mathsf{C}(h)_{l,m}|$$

If $(j, k) \neq (d, d)$ we let $(j_\downarrow, k_\downarrow)$ be the immediate successor of (j, k) in the lexicographic order. We also let

$$\text{NSPD}(g, d, j, k)_\downarrow = \begin{cases} \text{NSPD}(g, d, j_\downarrow, k_\downarrow) & \text{if } (j, k) \neq (d, d) \\ \text{NSPD}(g_\downarrow, d, 1, 1) & \text{if } (j, k) = (d, d) \end{cases}$$

Write $\mathcal{P}(g, d, j, k)$ for the set of indexes $(h, m) \in \mathbb{F} \times [d]$ such that one of the following conditions holds.

- We have $h \in \mathcal{K}_g \setminus \{e, g\}$ and $m \in [d]$ is arbitrary.
- We have $h = g$ and $m \in [j-1]$.
- We have $h = e$ and $m \in [k-1]$.

Also let $\mathcal{Q}(g, d, j, k) = \mathcal{P}(g, d, j, k) \cup \{(e, k), (g, j)\}$.

2.2.5 Parameterizing extensions

In Segment 2.2.5 we describe the procedure for extending positive definite functions. This construction has its roots in [1] and in Section 3.6 of [2].

Let $g \in \mathbb{F}$ and fix $\mathsf{C} \in \mathrm{NSPD}(g_\uparrow, d)$. We wish to understand extensions of C to an element of $\mathrm{NSPD}(g, d)$. In order to describe such an extension it suffices to specify the matrix $\mathsf{C}(g)$ since then we must have $\mathsf{C}(g^{-1}) = \mathsf{C}(g)^*$.

We begin by constructing the partial Hilbert space $\mathcal{X}(\mathsf{C})_\bullet$ as in Proposition 2.7. Specifying the matrix $\mathsf{C}(g)$ amounts to specifying the inner products between $\Theta_\mathsf{C}(g)_j$ and $\Theta_\mathsf{C}(e)_k$ for $(j, k) \in [d]^2$. We do this by recursion on the lexicographic order on pairs $(j, k) \in [d]^2$. We use the notation $<$ for the lexicographic order on $[d]^2$.

Fix $(j, k) \in [d]^2$ and $\underline{\mathsf{C}} \in \mathrm{NSPD}(g, d, j, k)$. Since we have specified $\underline{\mathsf{C}}(g)_{l,m}$ for $(l, m) < (j, k)$ we have that the inner product between $\Theta_\mathsf{C}(g)_l$ and $\Theta_\mathsf{C}(e)_m$ is defined for all pairs $(l, m) < (j, k)$. Thus we have specified the inner products between $\Theta_\mathsf{C}(h)_m$ and $\Theta_\mathsf{C}(\ell)_q$ for all pairs (h, m) and (ℓ, q) of elements of $\mathcal{P}(g, d, j, k)$.

Using this data we can construct a partial Hilbert space $\mathcal{X}(\underline{\mathsf{C}})_\bullet$ which is the subspace of $\mathcal{X}(\mathsf{C})_\bullet$ spanned by $\{\Theta_\mathsf{C}(h)_m : (h, m) \in \mathcal{Q}(g, d, j, k)\}$. The distinguished subspaces of $\mathcal{X}(\underline{\mathsf{C}})_\bullet$ are given by

$$\mathcal{X}(\underline{\mathsf{C}})_g = \mathrm{span}\big(\{\Theta_\mathsf{C}(h)_m : (h, m) \in \mathcal{P}(g, d, j, k)\} \cup \{\Theta_\mathsf{C}(g)_j\}\big)$$

and

$$\mathcal{X}(\underline{\mathsf{C}})_e = \mathrm{span}\big(\{\Theta_\mathsf{C}(h)_m : (h, m) \in \mathcal{P}(g, d, j, k)\} \cup \{\Theta_\mathsf{C}(e)_k\}\big)$$

and $\mathrm{core}(\mathcal{X}(\underline{\mathsf{C}})_\bullet)$ is the span of $\{\Theta_\mathsf{C}(h)_m : (h, m) \in \mathcal{P}(g, d, j, k)\}$. The point of attempting to extend $\underline{\mathsf{C}}$ instead of directly extending C is that the pair $\{\Theta_\mathsf{C}(g)_j, \Theta_\mathsf{C}(e)_k\}$ is the only pair of canonical basis vectors in $\mathcal{X}(\underline{\mathsf{C}})_\bullet$ whose inner product is undefined.

Write I for the identity operator on $\mathcal{X}(\underline{\mathsf{C}})_\bullet$. By Proposition 2.6 we can consider the orthogonal projection p from $\mathcal{X}(\underline{\mathsf{C}})_\bullet$ onto $\mathrm{core}(\mathcal{X}(\underline{\mathsf{C}})_\bullet)$. Given a complex number ζ with $|\zeta| < 1$, we can make the following definition, which generalizes (2.7). Set

$$\left\langle \frac{(I-p)\Theta_\mathsf{C}(g)_j}{||(I-p)\Theta_\mathsf{C}(g)_j||}, \frac{(I-p)\Theta_\mathsf{C}(e)_k}{||(I-p)\Theta_\mathsf{C}(e)_k||} \right\rangle = \zeta \tag{2.12}$$

The hypothesis that $\underline{\mathsf{C}}$ is strictly positive definite implies that the denominators of the fractions in (2.12) are nonzero. As with (2.7), the requirement that $|\zeta| \leq 1$ is immediate from the need to satisfy the Cauchy-Schwartz inequality. The requirement that $|\zeta| \neq 1$ will be discussed in Segment 2.2.8. From (2.12) we have the analog of (2.9), whereby we recover $\mathsf{C}(g)_{j,k}$ as

$$\mathsf{C}(g)_{j,k} = \langle \Theta_\mathsf{C}(g)_j, \Theta_\mathsf{C}(e)_k \rangle$$

$$= \zeta ||(I-p)\Theta_{\mathsf{C}}(g)_j|| \, ||(I-p)\Theta_{\mathsf{C}}(e)_k|| + \langle p\Theta_{\mathsf{C}}(g)_j, p\Theta_{\mathsf{C}}(e)_k \rangle \tag{2.13}$$

The norms and inner product in (2.13) are determined by $\underline{\mathsf{C}}$, so that ζ is indeed the only free parameter. The fact that any value of ζ with $|\zeta| \leq 1$ produces a valid positive definite function follows by the analog of (2.8), which is the decomposition

$$\mathcal{X}(\underline{\mathsf{C}})_\bullet = \operatorname{core}(\mathcal{X}(\underline{\mathsf{C}})_\bullet) \oplus \operatorname{span}\big((I-p)\Theta_{\mathsf{C}}(g)_j, (I-p)\Theta_{\mathsf{C}}(e)_k\big)$$

valid for any value of ζ. Write $\underline{\mathsf{C}}^\zeta$ for the extension of $\underline{\mathsf{C}}$ by ζ. Thus after extending by ζ the partial Hilbert space $\mathcal{X}(\underline{\mathsf{C}})_\bullet$ is promoted to a full Hilbert space $\mathcal{X}(\underline{\mathsf{C}}^\zeta)$. The Hilbert space $\mathcal{X}(\underline{\mathsf{C}}^\zeta)$ has a canonical basis indexed by $\mathcal{Q}(g,d,j,k)$ and we can regard $\underline{\mathsf{C}}^\zeta$ as an element of $\mathrm{NSPD}(g,d,j,k)_\downarrow$. It is natural to think of ζ as a noncommutative Szegö parameter.

If we choose $\zeta = 0$ at every step of the extension procedure we obtain the so-called 'central extension', which corresponds to the construction of a higher-step Markov process on the free group. This is the construction given in Lemma 24 of [13]. However, the central extension does not have the properties required to prove Theorem 1.1 and we must choose the extension parameters more carefully.

The essential proofs in Section 2 are contained in Subsections 2.3 - 2.5. In these subsections we will always have g, d, j and k fixed and we will be considering elements of $\mathrm{NSPD}(g,d,j,k)$. Thus to ease notation we will omit the underline and denote elements of $\mathrm{NSPD}(g,d,j,k)$ by symbols such as C and D.

2.2.6 Transport operators of partially defined functions

We will need the following analog of Definition 1.5.

Definition 2.2. *Let $g \in \mathbb{F}, d \in \mathbb{N}$ and $(j,k) \in [d]^2$. Let $\mathsf{C}, \mathsf{D} \in \mathrm{NSPD}(g,d,j,k)$. The **transport operator** between the partial Hilbert spaces $\mathcal{X}(\mathsf{C})_\bullet$ and $\mathcal{X}(\mathsf{D})_\bullet$ is denoted $t[\mathsf{C},\mathsf{D}]$ and is given by setting*

$$t[\mathsf{C},\mathsf{D}] \sum_{(h,m)\in\mathcal{Q}(g,d,j,k)} \alpha(h)_m \Theta_{\mathsf{C}}(h)_m = \sum_{(h,m)\in\mathcal{Q}(g,d,j,k)} \alpha(h)_m \Theta_{\mathsf{D}}(h)_m$$

*for $\alpha : \mathcal{Q}(g,d,j,k) \to \mathbb{C}$. We define the **relative energy** of the pair (C,D) to be the maximum of the squares of the operator norms of the restrictions of $t[\mathsf{C},\mathsf{D}]$ to the distinguished subspaces $\mathcal{X}(\mathsf{C})_g$ and $\mathcal{X}(\mathsf{C})_e$. We denote the relative energy of the pair (C,D) by $\mathfrak{e}(\mathsf{C},\mathsf{D})$.*

There is a slight difference between Definitions 1.5 and 2.2. If $|g| = 2r$ and g is the $\preceq$ last element of its length then $\mathcal{K}_g$ is a translate of $\mathbb{B}_r$ different from $\mathbb{B}_r$.

2.2.7 Strong local free group extension graphs

Definition 2.3. *Let (V,E) be a finite directed graph. We define (V,E) to be a **strong local free group extension graph** if the following holds. Let $g, d \in \mathbb{N}$ and let $j,k \in [d]$ and let $\eta > 0$. Let $(\mathsf{D}_v)_{v\in V}$ be a family of elements of $\mathrm{NSPD}(g,d,j,k)$. Then there exist elements $(\mathsf{K}_v)_{v\in V}$ of $\mathrm{NSPD}(g,d,j,k)_\downarrow$ such that $||\mathsf{D}_v - (\mathsf{K}_v \restriction \mathcal{D}(g,d,j,k))||_1 \leq \eta$ for all $v \in V$ and such that $\mathfrak{e}(\mathsf{K}_v,\mathsf{K}_w) \leq \mathfrak{e}(\mathsf{D}_v,\mathsf{D}_w) + \eta$ whenever $(v,w) \in E$.*

2.2.8 Degenerate extensions

Observe that the construction described in Segment 2.2.5 makes sense if we choose ζ to be an element of the unit circle. However, in this case the resulting extension will not be strictly positive definite and so the denominators of the fractions in (2.12) will be zero at some later stage of the procedure. Thus we regard $|\zeta| = 1$ as an unacceptable degeneracy. Intuitively, such a degeneracy corresponds to a cycle in the time evolution represented by the extension procedure. The following definition will allow us to avoid this issue in the context of minimizing relative energies.

Definition 2.4. *Let $g \in \mathbb{F}, d \in \mathbb{N}$ and $(j,k) \in [d]^2$. Let $\mathsf{C}, \mathsf{D} \in \mathrm{NSPD}(g,d,j,k)$. We define the pair (C, D) to have **singular degeneracies** if there is a constant $c > 0$ depending only on previously introduced data such that*

$$\mathfrak{e}(\mathsf{C}^{\zeta}, \mathsf{D}^{\mu}) \geq \frac{c}{1 - |\zeta|^2}$$

*for all $\zeta, \mu \in \mathbb{D}$. If $N \in \mathbb{N}$, we say that a family $\mathsf{C}_1, \ldots, \mathsf{C}_N \in \mathrm{NSPD}(g,d,j,k)$ has **completely singular degeneracies** if the pair $(\mathsf{C}_l, \mathsf{C}_m)$ has singular degeneracies for all distinct pairs $l, m \in [N]$.*

2.2.9 Three lemmas

We now formulate three lemmas that will imply Theorem 1.1. We will prove them below in Subsections 2.3, 2.4 and 2.5 respectively.

Lemma 2.1 (Small perturbations give singular degeneracies)**.** *Let $g \in \mathbb{F}$ be such that $|g| \geq 5$, let $d, N \in \mathbb{N}$, let $j, k \in [d]$ and let $\eta > 0$. Let also $\mathsf{C}_1, \ldots, \mathsf{C}_N \in \mathrm{NSPD}(g,d,j,k)$. Then there exist $\mathsf{D}_1, \ldots, \mathsf{D}_N \in \mathrm{NSPD}(g,d,j,k)$ such that $||\mathsf{C}_m - \mathsf{D}_m||_1 \leq \eta$ for all $m \in [N]$ and such the family $\mathsf{D}_1, \ldots, \mathsf{D}_N$ has completely singular degeneracies.*

Lemma 2.2. *If Θ is a strong local free group extension graph then $\mathrm{encost}_{r,\epsilon}(\Theta) = 1$ for all $r \in \mathbb{N}$ and $\epsilon > 0$.*

Lemma 2.3. *A tree directed toward a root and a single directed cycle are strong local free group extension graphs.*

2.3 Proof of Lemma 2.1

In Subsection 2.3 we prove Lemma 2.1.

2.3.1 Smoothness of Gram-Schmidt

In Segment 2.3.1 we establish Proposition 2.8. This is a general result about the Gram-Schmidt procedure which is likely well-known.

Definition 2.5. *Let $n \in \mathbb{N}$ and let $M \in \mathrm{Mat}_{n \times n}(\mathbb{C})$ be strictly positive definite with ones on the diagonal. Let $\mathcal{Z}$ be a Hilbert space and let $y_1, \ldots, y_n \in \mathcal{Z}$ be a basis such that $\langle y_j, y_k \rangle = M_{j,k}$ for all $j, k \in [n]$. Let $z_1, \ldots, z_n$ be the orthogonal basis obtained by applying the Gram-Schmidt procedure to $y_1, \ldots, y_n$. We define the **orthogonalization matrix** of M to be the $n \times n$ matrix which changes $y_1, \ldots, y_n$ coordinates to $z_1, \ldots, z_n$ coordinates. We denote the orthogonalization matrix by $\mathcal{G}(M)$. Also define the **orthonormalization matrix** of M to be the matrix which changes $y_1, \ldots, y_n$ coordinates to $z_1||z_1||^{-1}, \ldots, z_n||z_n||^{-1}$ coordinates. We denote the orthonormalization matrix by $\mathcal{N}(M)$.*

Proposition 2.8. *The entries of $\mathcal{G}(M)$ is an analytic function of the entries of M. The entries of $\mathcal{N}(M)$ are differentiable functions of M.*

Proof of Proposition 2.8. We establish the proposition by induction on n. The case $n = 1$ is trivial, so assume we have established the result for n with the goal of establishing it for $n+1$. Fix an $n \times n$ strictly positive definite matrix $M_\bullet$ with ones on the diagonal. If M is an $(n+1) \times (n+1)$ matrix we have that the first n columns of $\mathcal{G}(M)$ depend only on the $n \times n$ upper left corner of M. Therefore we will express M in terms of a variable $\mathbf{x} = (x_1, \dots, x_n, 1) \in \mathbb{C}^{n+1}$ representing a column to be augmented on the right of $M_\bullet$ to form a $(n+1) \times (n+1)$ strictly positive definite matrix $M(\mathbf{x})$. Note that for a given strictly positive definite $M_\bullet$ the set of $\mathbf{x} \in \mathbb{C}^{n+1}$ such that $M(\mathbf{x})$ is strictly positive definite is open. Moreover, it is nonempty since it contains the vector $(0, \dots, 0, 1)$.

Let $\mathbf{q}(\mathbf{x}) = (q_1(\mathbf{x}), \dots, q_n(\mathbf{x}), 1)$ be the $k+1$ column of $\mathcal{G}(M(\mathbf{x}))$. For $k \in [n]$ let $M_\bullet(\check{k})$ be $M_\bullet$ with the k^{th} column removed. We have the following expression for $\mathbf{q}(\mathbf{x})$.

$$q_k(\mathbf{x}) = (-1)^{k+n} \frac{\det(M_\bullet(\check{k})\,|\mathbf{x})}{\det(M_\bullet)}. \tag{2.14}$$

This appears, for example, as (35) in Section 6 of Chapter IX of [6]. The first clause in Proposition 2.8 is clear from (2.14). Writing $(r_1(\mathbf{x}), \dots, r_n(\mathbf{x}), r_{n+1}(\mathbf{x}))$ for the $k+1$ column of $\mathcal{N}(M(\mathbf{x}))$ from the same reference we have the formula

$$r_k(\mathbf{x}) = (-1)^{k+n} \frac{\det(M_\bullet(\check{k})\,|\mathbf{x})}{\sqrt{\det(M_\bullet)\det(M(\mathbf{x}))}}. \tag{2.15}$$

The second clause in Proposition 2.8 is clear from (2.15). ☐

2.3.2 First perturbation of the configuration

Let g, d, N, j, k, η and $\mathsf{C}_1, \dots, \mathsf{C}_N$ be as in the statement of Lemma 2.1.

For $s \in \{1, \dots, |g|\}$ let g_s be the word consisting of the first s letters of g. Given a 4-tuple $\Lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ of complex numbers, let $\mathsf{C}_{m,\Lambda}$ denote the modification of C_m given by setting

$$\begin{aligned}
(\mathsf{C}_{m,\Lambda})(g_1^{-1}g)_{j,1} &= \overline{(\mathsf{C}_{m,\Lambda})(g^{-1}g_1)_{j,1}} = (\mathsf{C}_m)(g_1^{-1}g)_{j,1} + \lambda_1 \\
(\mathsf{C}_{m,\Lambda})(g_2^{-1}g)_{j,1} &= \overline{\mathsf{C}_{m,\Lambda}(g^{-1}g_2)_{j,1}} = (\mathsf{C}_m)(g_2^{-1}g)_{j,1} + \lambda_2 \\
(\mathsf{C}_{m,\Lambda})(g_1^{-1})_{k,1} &= \overline{\mathsf{C}_{m,\Lambda}(g_1)_{k,1}} = (\mathsf{C}_m)(g_1^{-1})_{k,1} + \lambda_3 \\
(\mathsf{C}_{m,\Lambda})(g_2^{-1})_{k,1} &= \overline{\mathsf{C}_{m,\Lambda}(g_2)_{k,1}} = (\mathsf{C}_m)(g_2^{-1})_{k,1} + \lambda_4
\end{aligned}$$

and leaving all other entries of C_m unchanged. Since the space of strictly positive definite functions is open, if $||\Lambda||_1$ is small enough we will have $\mathsf{C}_{m,\Lambda} \in \mathrm{NSPD}(g, d, j, k)$.

Let $w_m(\Lambda)$ be the orthogonal projection from $\mathcal{X}(\mathsf{C}_{m,\Lambda})_\bullet$ onto the span of $\Theta_{\mathsf{C}_{m,\Lambda}}(g_1)_1$ and $\Theta_{\mathsf{C}_{m,\Lambda}}(g_2)_1$. Let $W_m(\Lambda)$ be the matrix of $w_m(\Lambda)$ constructed with respect to the canonical basis

$$\{\Theta_{\mathsf{C}_{m,\Lambda}}(h)_n : (h, n) \in \mathcal{Q}(g, d, j, k)\} \tag{2.16}$$

Note that this matrix does not depend on the correlation between $\Theta_{\mathsf{C}_m}(g)_j$ and $\Theta_{\mathsf{C}_m}(e)_k$ and so it is well-defined on the partial Hilbert space $\mathcal{X}(\mathsf{C}_{m,\Lambda})_\bullet$. Let $W'_m(\Lambda)$ be the restriction of $W_m(\Lambda)$ to the span of $\{\Theta_{\mathsf{C}_{m,\Lambda}}(g)_j, \Theta_{\mathsf{C}_{m,\Lambda}}(e)_k\}$. Thus $W'_m(\Lambda)$ is a 2×2 square matrix. Since $W'_m(\Lambda)$ is an invertible affine function of Λ we obtain the following.

Proposition 2.9. *For each $m \in [N]$ there exists Λ with $||\Lambda||_1 \leq \frac{\eta}{2}$ such that the matrix $W'_m(\Lambda)$ is invertible.*

Using Proposition 2.9 for each $m \in [N]$ we can fix $\Lambda_m \in \mathbb{C}^4$ such that $W'_m(\Lambda_m)$ is invertible, $||\mathsf{C}_{m,\Lambda_m} - \mathsf{C}_m||_1 \leq \frac{\eta}{2}$ and C_{m,Λ_m} remains strictly positive definite. Write C'_m for C_{m,Λ_m}.

2.3.3 Second perturbation of the configuration

As before, let $\Delta \in \mathrm{NSPD}(g,d,j,k)$ be given by setting $\Delta(e) = \mathbf{I}_d$ and letting all other entries of Δ be equal to 0. For $s \in [0,1]$ and $m \in [N]$ let $\mathsf{C}_{m,s} = (1-s)\mathsf{C}'_m + s\Delta$.

We specify that when performing the Gram-Schmidt orthogonalization procedure on the canonical basis (2.16) the vectors $\{\Theta_{\mathsf{C}'_m}(g)_j, \Theta_{\mathsf{C}'_m}(e)_k\}$ should be the last to be orthogonalized. Let $A_{m,s} = (\mathbf{I} \oplus \mathbf{0})\mathcal{G}(\mathsf{C}'_{m,s})$, where $\mathbf{I}$ is a copy of the identity matrix corresponding to the indexes in $\mathcal{P}(g,d,j,k)$ and $\mathbf{0}$ is a copy of the zero matrix corresponding to the indexes $\{(g,j),(e,k)\}$. Note that since the last two rows of $A_{m,s}$ are zero, this matrix is well-defined even though the inner product $\langle \Theta_{\mathsf{C}'_m}(g)_j, \Theta_{\mathsf{C}'_m}(e)_k \rangle$ is not yet specified.

Proposition 2.10. *There exist $s_1, \ldots, s_N$ such that for each distinct pair $l, m \in [N]$ the kernel of A_{l,s_l} has trivial intersection with the kernel of A_{m,s_m} and such that $||\mathsf{C}'_{m,s_m} - \mathsf{C}'_m||_1 \leq \frac{\eta}{2}$ for all $m \in [N]$.*

Proof of Proposition 2.10. Proposition 2.8 implies that the entries of $A_{m,s}$ are real analytic functions of s. Let $Q_{m,s}$ be the span of $\{\Theta_{\mathsf{C}_{m,s}}(g)_j, \Theta_{\mathsf{C}_{m,s}}(e)_k\}$. The kernel of $A_{m,s}$ is equal to the orthogonal complement of $\mathrm{core}(\mathcal{X}(\mathsf{C}_{m,s})_\bullet)$ in $\mathcal{X}(\mathsf{C}_{m,s})_\bullet$. Proposition 2.9 implies that the restriction to Q_m of the orthogonal projection from $\mathcal{X}(\mathsf{C}'_m)_\bullet$ onto $\mathrm{core}(\mathcal{X}(\mathsf{C}'_m)_\bullet)$ has trivial kernel. Therefore the kernel of the matrices $A_{m,0}$ have trivial intersection with Q_m. On the other hand, it is clear that the kernel of $A_{m,1}$ is equal to Q_m. Hence if we write

$$D_{l,m}(s,u) = \det\big(A^*_{l,s}A_{l,s} + A^*_{m,u}A_{m,u}\big)$$

then $D_{l,m}(0,1) \neq 0$ for all $l, m \in [N]$. Since $D_{l,m}(s,u)$ is a real analytic function of s and u we see that the set of pairs (s,u) such that $D_{l,m}(s,u) = 0$ has Lebesgue measure 0. Hence we can choose $s_1, \ldots, s_N$ such that $D_{l,m}(s_l, s_m) \neq 0$ for all $l, m \in [N]$ and such that $||\mathsf{C}'_{m,s_m} - \mathsf{C}'_m||_1 \leq \frac{\eta}{2}$. □

Let $s_1, \ldots, s_N$ be as in Proposition 2.10. Write D_m for C_{m,s_m} and $\widehat{A}_m$ for A_{m,s_m}. Let $\theta > 0$ be such that if $\alpha : \mathcal{P}(g,d,j,k) \to \mathbb{C}$ satisfies $||\alpha||_2 \geq 1$ and $\widehat{A}_m\alpha = 0$ for some $m \in [N]$ then $||\widehat{A}_l\alpha||_2 \geq \theta$ for all $l \in [N]$ different from m. Let also $B_{m,\circ}$ be the orthogonal basis for $\mathrm{core}(\mathcal{X}(\mathsf{D}_m)_\bullet)$ obtained by applying the Gram-Schmidt orthogonalization procedure to $\{\Theta_{\mathsf{D}_m}(h)_n : (h,n) \in \mathcal{P}(g,d,j,k)\}$. Let κ be the minimal norm of among all elements of $B_{m,\circ}$ for $m \in [N]$.

2.3.4 Establishing the existence of energy singularities

Fix $l, m \in [N]$ and $\zeta, \mu \in \mathbb{D}$. Consider the extensions $\mathsf{D}^\zeta_m, \mathsf{D}^\mu_l \in \mathrm{NSPD}(g,d,j,k)$. Let B be the orthogonal basis produced obtained applying the Gram-Schmidt orthogonalization procedure to the canonical basis

$$\{\Theta_{\mathsf{D}_l}(h)_n : (h,n) \in \mathcal{P}(g,d,j,k)\} \cup \{\Theta_{\mathsf{D}_l}(g)_j, \Theta_{\mathsf{D}_l}(e)_k\}$$

Note that this orthogonalization procedure can be completed in its entirety because we have specified $\langle \Theta_{\mathsf{D}_l}(g)_j, \Theta_{\mathsf{D}_l}(e)_k \rangle$. Moreover, we stipulate that when performing this procedure the vectors $\Theta_{\mathsf{D}_l}(g)_j$ and $\Theta_{\mathsf{D}_l}(e)_k$ are orthogonalized at the last stage. Then B extends $B_{l,\circ}$.

Consider a function $\alpha : \mathcal{Q}(g,d,j,k) \to \mathbb{C}$, which defines a vector

$$x = \alpha(g)_j \Theta_{\mathsf{D}_m}(g)_j + \alpha(e)_k \Theta_{\mathsf{D}_m} + \sum_{(h,n)\in\mathcal{P}(g,d,j,k)} \alpha(h)_n \Theta_{\mathsf{D}_m}(h)_n$$

in the space $\mathcal{X}(\mathsf{D}_l^\mu)$. If we rewrite x in the basis B, the resulting coordinates are given by $\widehat{A}_l\alpha$. Therefore $||x|| \geq \kappa ||\widehat{A}_l \alpha||_2$. From our choice of θ we see that if $\widehat{A}_m\alpha = 0$ then $||\widehat{A}_l\alpha||_2 \geq \theta||\alpha||_2$ so that $||x|| \geq \kappa\theta||\alpha||_2$.

Let p be the orthogonal projection from the extended Hilbert space $\mathcal{X}(\mathsf{D}_m^\zeta)$ onto $\mathrm{core}(\mathcal{X}(\mathsf{D}_m)_\bullet)$. Then we have

$$\left\langle \overline{\zeta}\frac{(I-p)\Theta_{\mathsf{D}_m}(g)_j}{||(I-p)\Theta_{\mathsf{D}_m}(g)_j||}, \frac{(I-p)\Theta_{\mathsf{D}_m}(e)_k}{||(I-p)\Theta_{\mathsf{D}_m}(e)_k||} \right\rangle = |\zeta|^2 \tag{2.17}$$

Let

$$y = \overline{\zeta}\frac{(I-p)\Theta_{\mathsf{D}_m}(g)_j}{||(I-p)\Theta_{\mathsf{D}_m}(g)_j||} - \frac{(I-p)\Theta_{\mathsf{D}_m}(e)_k}{||(I-p)\Theta_{\mathsf{D}_m}(e)_k||}$$

If we rewrite

$$y = \alpha(g)_j \Theta_{\mathsf{D}_m}(g)_j + \alpha(e)_k \Theta_{\mathsf{D}_m} + \sum_{(h,n)\in\mathcal{P}(g,d,j,k)} \alpha(h)_n \Theta_{\mathsf{D}_m}(h)_n$$

then we must have

$$\alpha(e)_k = \frac{1}{||(I-p)\Theta_{\mathsf{D}}(e)_k||}$$

so that $||\alpha||_2 \geq 1$. Moreover, we have $\widehat{A}_m(\alpha) = 0$ since y lies in the orthogonal complement of $\mathrm{core}(\mathcal{X}(\mathsf{D}_m)_\bullet)$. It follows that $||t[\mathsf{D}_m^\zeta, \mathsf{D}_l^\mu]y|| \geq \kappa\theta$.

On the other hand, from (2.17) we see $||y||^2 \leq 2 - 2|\zeta|^2$. Therefore $\mathfrak{c}(\mathsf{D}_m^\zeta, \mathsf{D}_l^\mu]) \geq \kappa^2\theta^2(2-2|\zeta|^2)^{-1}$. Since κ and θ are determined by $\mathsf{D}_1, \ldots, \mathsf{D}_N$ this completes the proof of Lemma 2.1.

2.4 Proof of Lemma 2.2

In Subsection 2.4 we prove Lemma 2.2.

2.4.1 Bounding transport operators by ℓ^1 distance

We will need the following preliminary result. Given invertible $n\times n$ matrices L and M, let $t[L,M] : \mathbb{C}^n \to \mathbb{C}^n$ be the operator which maps the j^{th} column of L to the j^{th} column of M.

Proposition 2.11. *Let L be an invertible $n \times n$ matrix. Then for every $\sigma > 0$ there exists $\eta > 0$ such that if M is an $n \times n$ matrix and $||L^*L - M^*M||_1 \leq \eta$ then*

$$\max(||t[L,M]||_{\mathrm{op}}, ||t[M,L]||_{\mathrm{op}}) \leq 1 + \sigma$$

Proof of Proposition 2.11. Let $\sigma \in (0,1)$ and choose $\eta > 0$ such that $2\eta||L^{-1}||_{\text{op}}^2 \leq \sigma$. Assume $||L^*L - M^*M||_1 \leq \eta$. We compute

$$\begin{aligned}
||M^{-1}||_{\text{op}}^2 &= ||(M^*M)^{-1}||_{\text{op}} \\
&= ||(M^*M - L^*L + L^*L)^{-1}||_{\text{op}} \\
&= \left|\left|L^{-1}\big(\mathbf{I}_n - (L^*)^{-1}(L^*L - M^*M)L^{-1}\big)^{-1}(L^*)^{-1}\right|\right|_{\text{op}} \\
&\leq ||L^{-1}||_{\text{op}}^2 \left|\left|\big(\mathbf{I}_n - (L^*)^{-1}(L^*L - M^*M)L^{-1}\big)^{-1}\right|\right|_{\text{op}} && (2.18)\\
&\leq ||L^{-1}||_{\text{op}}^2\big(1 - ||L^{-1}||_{\text{op}}^2||M^*M - L^*L||_{\text{op}}\big)^{-1} && (2.19)\\
&\leq ||L^{-1}||_{\text{op}}^2\big(1 - ||L^{-1}||_{\text{op}}^2||M^*M - L^*L||_1\big)^{-1} \\
&\leq ||L^{-1}||_{\text{op}}^2(1 - \eta||L^{-1}||_{\text{op}}^2)^{-1} \\
&\leq 2||L^{-1}||_{\text{op}}^2 && (2.20)
\end{aligned}$$

Here, (2.19) follows from (2.18) since $||(\mathbf{I}_n - A)^{-1}||_{\text{op}} \leq (1 - ||A||_{\text{op}})^{-1}$ whenever $||A||_{\text{op}} < 1$.

Now, the matrix of t is given by ML^{-1}. We compute

$$\begin{aligned}
||ML^{-1}||_{\text{op}}^2 &= ||(ML^{-1})^*ML^{-1}||_{\text{op}} \\
&= ||(L^{-1})^*M^*ML^{-1}||_{\text{op}} \\
&= ||(L^{-1})^*(M^*M - L^*L + L^*L)L^{-1}||_{\text{op}} \\
&\leq 1 + ||(L^{-1})^*(M^*M - L^*L)L^{-1}||_{\text{op}} \\
&\leq 1 + ||L^{-1}||_{\text{op}}^2||M^*M - L^*L||_{\text{op}} \\
&\leq 1 + \sigma
\end{aligned}$$

Similarly, we find

$$\begin{aligned}
||LM^{-1}||_{\text{op}} &\leq 1 + \eta||M^{-1}|| && (2.21)\\
&\leq 1 + 2\eta||L^{-1}||_{\text{op}} && (2.22)\\
&\leq 1 + \sigma
\end{aligned}$$

Here, (2.22) follows from (2.21) by (2.20). ☐

2.4.2 Carrying out the recursive procedure

Let (V, E) be a strong local free group extension graph. Let r, d, ϵ and $(\mathsf{C}_v)_{v\in V}$ be as in Definition 1.7. We perform a recursive construction, first over $\mathbb{F}$ along the ordering $\preceq$ and secondly along $[d]^2$ according to the lexicographic order $\leq$. We will continue to use the notation $\preceq$ to refer to this ordering among triples (g, j, k) where $g \in \mathbb{F}$ and $j, k \in [d]$.

Let $(\sigma_{g,j,k})_{g\in\mathbb{F},j,k\in[d]}$ be a sequence of positive numbers such that

$$2\sum_{g\in\mathbb{F}}\sum_{j,k=1}^{d}\sigma_{g,j,k} \leq \frac{\epsilon}{2}$$

Let $g_\circ$ be the $\preceq$ first element of $\mathbb{F} \setminus \mathbb{B}_r$. Let $\mathsf{C}_{v,g_\circ,1,1} = \mathsf{C}_v$. Fix $g \in \mathbb{F} \setminus \mathbb{B}_r$ and $j, k \in [d]$.

Suppose that for all $(h, l, m) \preceq (g, j, k)$ we have constructed elements $(\mathsf{C}_{v,h,l,m})_{v \in V}$ of $\mathrm{NSPD}(h, d, l, m)$ such that

$$||\mathsf{C}_{v,h,l_\uparrow,m_\uparrow} - (\mathsf{C}_{v,h,l,m} \restriction \mathcal{D}(h, d, l_\uparrow, m_\uparrow))||_1 \leq \sigma_{h,l,m} \tag{2.23}$$

for all $(h, l, m) \preceq (g, j, k)$ and such that

$$\mathfrak{e}\big(\mathsf{C}_{v,h,l,m}, \mathsf{C}_{w,h,l,m}\big) \leq \mathfrak{e}(\mathsf{C}_v, \mathsf{C}_w) + \sum_{(\ell,n,p) \preceq (h,l,m)} \sigma_{\ell,n,p} \tag{2.24}$$

for all $(h, l, m) \preceq (g, j, k)$ and all $(v, w) \in E$. Write $\widetilde{\mathsf{C}}_v$ for $\mathsf{C}_{v,g,j,k}$.

Choose $\eta' > 0$ such that

$$(1 + \eta')\mathfrak{e}(\mathsf{C}_v, \mathsf{C}_w) \leq \mathfrak{e}(\mathsf{C}_v, \mathsf{C}_w) + \sigma_{g,j_\downarrow,k_\downarrow} \tag{2.25}$$

for all $(v, w) \in E$.

Given a positive definite matrix J we can write $J = U^{-1}QU$ for a unitary matrix U and a positive diagonal matrix Q. Let $\sqrt{Q}$ be the diagonal matrix whose entries are the square roots of the entries of Q. Then the columns of $\sqrt{Q}U$ form a realization of J. Moreover, Section II.6.2 of [9] guarantees that U and Q can be chosen to be continuous functions of J in the ℓ^1-norm.

We can regard each $\widetilde{\mathsf{C}}_v$ as a square matrix indexed by $\mathcal{P}(g, d, j, k)$. By applying Proposition 2.11 to the construction in the previous paragraph we can find $\eta > 0$ such that if $\mathsf{J} \in \mathrm{NSPD}(g, d, j, k)$ and $||\widetilde{\mathsf{C}}_v - \mathsf{J}||_1 \leq \eta$ then

$$\max(\mathfrak{e}(\widetilde{\mathsf{C}}_v, \mathsf{J}), \mathfrak{e}(\mathsf{J}, \widetilde{\mathsf{C}}_v)) \leq 1 + \eta' \tag{2.26}$$

Apply Lemma 2.1 to this choice of η and the family $(\widetilde{\mathsf{C}}_v)_{v \in V}$ to obtain a family $(\mathsf{D}_v)_{v \in V}$ of elements of $\mathrm{NSPD}(g, d, j, k)$ with completely singular degeneracies satisfying

$$\max(\mathfrak{e}(\widetilde{\mathsf{C}}_v, \mathsf{D}_v), \mathfrak{e}(\mathsf{D}_v, \widetilde{\mathsf{C}}_v)) \leq 1 + \sigma_{g,j_\downarrow,k_\downarrow} \tag{2.27}$$

for all $v \in V$. Then for $(v, w) \in E$ we have

$$\mathfrak{e}(\mathsf{D}_v, \mathsf{D}_w) \leq \mathfrak{e}(\mathsf{D}_v, \widetilde{\mathsf{C}}_v)\mathfrak{e}(\widetilde{\mathsf{C}}_v, \widetilde{\mathsf{C}}_w)\mathfrak{e}(\widetilde{\mathsf{C}}_w, \mathsf{D}_w) \tag{2.28}$$

$$\leq (1 + \eta')^2\mathfrak{e}(\widetilde{\mathsf{C}}_v, \widetilde{\mathsf{C}}_w) \tag{2.29}$$

$$\leq \mathfrak{e}(\widetilde{\mathsf{C}}_v, \widetilde{\mathsf{C}}_w) + \sigma_{g,j_\downarrow,k_\downarrow} \tag{2.30}$$

Here, (2.29) follows from (2.28) by (2.26) and (2.30) follows from (2.29) by (2.25). Apply Definition 2.3 to the family $(\mathsf{D}_v)_{v \in V}$ with $\eta = \sigma_{g,j_\downarrow,k_\downarrow}$ to obtain a family $(\mathsf{K}_v)_{v \in V}$ of elements of $\mathrm{NSPD}(g, d, j, k)_\downarrow$ such that

$$\mathfrak{e}(\mathsf{K}_v, \mathsf{K}_w) \leq \mathfrak{e}(\widetilde{\mathsf{C}}_v, \widetilde{\mathsf{C}}_w) + \sigma_{g,j_\downarrow,k_\downarrow} \tag{2.31}$$

for all $(v, w) \in E$. Set $\mathsf{K}_v = \mathsf{C}_{g,d,j_\downarrow,k_\downarrow}$. Then (2.31) is the next step in the recursion for (2.24).

Thus we may assume that we have constructed $(\mathsf{C}_{v,g,j,k})_{v \in V}$ for all $g \in \mathbb{F}$ and all $j, k \in [d]$. Write $\mathsf{C}_{v,g}$ for $\mathsf{C}_{v,g,d,d}$. From (2.23) we see that the matrices $(\mathsf{C}_{v,g}(h))_{h \preceq g}$ form a Cauchy sequence for each $h \in \mathbb{F}$. Setting $\widehat{\mathsf{C}}_v(h)$ to be the limit of these matrices completes the proof.

2.5 Proof of Lemma 2.3

In Segments 2.5.1 - 2.5.7 of Subsection 2.5 we establish Propositions 2.12 - 2.18, which will be used below in Segments 2.5.8 and 2.5.9 to prove Lemma 2.3.

2.5.1 Introducing initial data

Let $g \in \mathbb{F}$, let $d \in \mathbb{N}$, let $j, k \in [d]$ and let $\mathsf{C}, \mathsf{D} \in \mathrm{NSPD}(g, d, j, k)$. Observe that for any $\zeta \in \mathbb{D}$, the space $\mathcal{X}(\mathsf{C}^\zeta)$ has a canonical basis as in (2.11) indexed by $\mathcal{Q}(g, d, j, k)$. We will adopt the convention that for $\lambda \in \mathbb{D}$ a vector $x \in \mathcal{X}(\mathsf{C}^\zeta)$ is identified with the vector in $\mathcal{X}(\mathsf{C}^\lambda)$ having the same coordinates with respect to the canonical basis.

Let $\zeta, \mu \in \mathbb{D}$ and consider the extensions C^ζ and D^μ. We consider an additive perturbation $\chi\varsigma$ to the parameters ζ and μ where $\varsigma \in \mathbb{D}$ and $\chi \in \mathbb{R}$ is sufficiently small that $\max(|\zeta + \chi\varsigma|, |\mu + \chi\varsigma|) < 1$. It will be convenient to introduce the asymptotic notations $O(\cdot)$ and $o(\cdot)$ with respect to the limit $\chi \to 0$.

Let p be the orthogonal projection from $\mathcal{X}(\mathsf{C})_\bullet$ onto $\mathrm{core}(\mathcal{X}(\mathsf{C})_\bullet)$ and let q be the orthogonal projection from $\mathcal{X}(\mathsf{D})_\bullet$ onto $\mathrm{core}(\mathcal{X}(\mathsf{D})_\bullet)$. We introduce the following notations.

$$S = \frac{(I-p)\Theta_{\mathsf{C}}(g)_j}{||(I-p)\Theta_{\mathsf{C}}(g)_j||} \tag{2.32}$$

$$S' = \frac{(I-p)\Theta_{\mathsf{C}}(e)_k}{||(I-p)\Theta_{\mathsf{C}}(e)_k||} \tag{2.33}$$

$$T = \frac{(I-q)\Theta_{\mathsf{D}}(g)_j}{||(I-p)\Theta_{\mathsf{D}}(g)_j||} \tag{2.34}$$

$$T' = \frac{(I-q)\Theta_{\mathsf{D}}(e)_k}{||(I-q)\Theta_{\mathsf{D}}(e)_k||} \tag{2.35}$$

2.5.2 Energy increases require extension components

Proposition 2.12. *Let $x \in \mathcal{X}(\mathsf{C})_\bullet$ and write $x = \alpha S + \alpha' S' + x'$ for $x' \in \mathrm{core}(\mathcal{X}(\mathsf{C})_\bullet)$. If $||t[\mathsf{C}^\zeta, \mathsf{D}^\mu]x||^2 > \mathfrak{e}(\mathsf{C}, \mathsf{D})||x||^2$ then both α and α' are nonzero.*

Proof of Proposition 2.12. This is immediate from the observation that if one of α and α' is zero then x lies in one of the distinguished subspaces of the partial Hilbert space $\mathcal{X}(\mathsf{C})_\bullet$. □

2.5.3 Energy increases give one dimensional norm achievers

Proposition 2.13. *Suppose that $\zeta, \mu \in \mathbb{D}$ are such that*

$$\mathfrak{e}(\mathsf{C}^\zeta, \mathsf{D}^\mu) > \mathfrak{e}(\mathsf{C}, \mathsf{D}) \tag{2.36}$$

Then the space of vectors which achieve the norm of $t[\mathsf{C}^\zeta, \mathsf{D}^\mu]$ is one-dimensional.

Proof of Proposition 2.13. Suppose toward a contradiction that x and y are orthogonal unit vectors in $\mathcal{X}(\mathsf{C}^\zeta)$ which achieve the norm of $t[\mathsf{C}^\zeta, \mathsf{D}^\mu]$. Write $x = \alpha S + \alpha' S' + x'$ for $x' \in \mathrm{core}(\mathcal{X}(\mathsf{C})_\bullet)$ and write $y = \beta S + \alpha' S' + y'$ for $y' \in \mathrm{core}(\mathcal{X}(\mathsf{C})_\bullet)$. Proposition 2.12 and the hypothesis (2.36) implies that $\alpha \neq 0$. Consider the vector $z = \beta\alpha^{-1}x - y$. This vector z also achieves the norm of $t[\mathsf{C}, \mathsf{D}]$. Since z has no S component using Proposition 2.12 we obtain a contradiction to the hypothesis (2.36). □

2.5.4 Initial bounds on coefficients and energies

Proposition 2.14. *Let $x \in \mathcal{X}(\mathsf{C}^\zeta)$ and write $x = \alpha S + \alpha' S' + x'$ for $x' \in \mathrm{core}(\mathcal{X}(\mathsf{C})_\bullet)$. Then we have $|\alpha||\alpha'| \le (1-|\zeta|)^{-1}||x||^2$.*

Proof of Proposition 2.14. We have

$$\begin{aligned}
||x||^2 &= ||\alpha S + \alpha' S' + x'||^2 && (2.37)\\
&\ge ||\alpha S + \alpha' S'||^2 && (2.38)\\
&= |\alpha|^2 + |\alpha'|^2 + 2\,\mathrm{Re}(\zeta\alpha\overline{\alpha'}) && (2.39)\\
&\ge |\alpha|^2 + |\alpha'|^2 - 2|\alpha||\alpha'||\zeta| \\
&= (1-|\zeta|)(|\alpha|^2 + |\alpha'|^2) + |\zeta|(|\alpha|^2 + |\alpha'|^2 - 2|\alpha||\alpha'|) && (2.40)\\
&\ge (1-|\zeta|)(|\alpha|^2 + |\alpha'|^2) && (2.41)
\end{aligned}$$

Here,

- (2.38) follows from (2.37) since $\alpha S + \alpha' S' \perp x'$
- (2.39) follows from (2.38) since by construction we have $\langle S, S'\rangle = \zeta$,
- and (2.41) follows from (2.40) since

$$|\alpha|^2 + |\alpha'|^2 - 2|\alpha||\alpha'| = (|\alpha| - |\alpha'|)^2 \ge 0$$

Therefore we have $\max(|\alpha|^2, |\alpha'|^2) \le (1-|\zeta|)^{-1}$ and so the proof of Proposition 2.14 is complete. □

Proposition 2.15. *We have $\mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu) = \mathfrak{e}(\mathsf{C}^\zeta, \mathsf{D}^\mu) + O(\chi)$ and $\mathfrak{e}(\mathsf{C}^\zeta, \mathsf{D}^{\mu+\chi\varsigma}) = \mathfrak{e}(\mathsf{C}^\zeta, \mathsf{D}^\mu) + O(\chi)$.*

Proof of Proposition 2.15. Let $x_\chi \in \mathcal{X}(\mathsf{C}^{\zeta+\chi\varsigma})$ be a unit vector which achieves the norm of $t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu]$. Write $||\cdot||_\chi$ for the norm on $\mathcal{X}(\mathsf{C}^{\zeta+\chi\varsigma})$ and write $x_\chi = \alpha_\chi S + \alpha'_\chi S' + x'_\chi$ for $x'_\chi \in \mathrm{core}(\mathcal{X}(\mathsf{C})_\bullet)$. We have $(1-|\zeta+\chi\varsigma|)^{-1} = O(1)$, so that Proposition 2.14 implies $|\alpha_\chi||\alpha'_\chi| = O(1)$. We have

$$\begin{aligned}
||x_\chi||_0^2 &= ||x_\chi||_\chi^2 - 2\chi\,\mathrm{Re}(\varsigma\alpha_\chi\overline{\alpha'_\chi}) && (2.42)\\
&= 1 - 2\chi\,\mathrm{Re}(\varsigma\alpha_\chi\overline{\alpha'_\chi}) && (2.43)\\
&\le 1 + 2\chi|\alpha_\chi||\alpha'_\chi| \\
&\le 1 + O(\chi) && (2.44)
\end{aligned}$$

Here, (2.43) follows from (2.42) since we assumed that x_χ was a unit vector in $\mathcal{X}(\mathsf{C}^{\zeta+\chi\varsigma})$. Therefore we have

$$\begin{aligned}
\mathfrak{e}(\mathsf{C}^\zeta, \mathsf{D}^\mu) &\ge ||t[\mathsf{C}^\zeta, \mathsf{D}^\mu]x_\chi||^2\,||x_\chi||_0^{-2} && (2.45)\\
&= ||t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu]x_\chi||^2\,||x_\chi||_0^{-2} && (2.46)\\
&= \mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu)||x_\chi||_0^{-2} && (2.47)\\
&\ge \mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu)(1+O(\chi))^{-1} && (2.48)\\
&= \mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu)(1-O(\chi)) && (2.49)\\
&= \mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu) - O(\chi) && (2.50)
\end{aligned}$$

Here,

- (2.46) follows from (2.45) since $||t[\mathsf{C}^{\zeta+\chi\varsigma},\mathsf{D}^{\mu}]x_\chi||$ is computed in $\mathfrak{X}(\mathsf{D}^\mu)$ and hence does not depend on χ,
- (2.47) follows from (2.46) since we assumed x_χ is a unit vector achieving the norm of $t[\mathsf{C}^{\zeta+\chi\varsigma},\mathsf{D}^{\mu}]$
- (2.48) follows from (2.47) by (2.44)
- (2.50) follows from (2.49) since Proposition 2.14 implies $\mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma},\mathsf{D}^{\mu}) = O(1)$.

Now, let $x_\chi \in \mathfrak{X}(\mathsf{C}^\zeta)$ be a unit vector which achieves the norm of $t[\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\chi\varsigma}]$. We modify the notation $||\cdot||_\chi$ to now refer to the norm of $\mathfrak{X}(\mathsf{D}^{\mu+\chi\varsigma})$ and write $t[\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\chi\varsigma}]x_\chi = \beta_\chi T + \beta'_\chi T' + x'_\chi$ for $x' \in \mathrm{core}(\mathfrak{X}(\mathsf{D})_\bullet)$. We have

$$\begin{aligned}
\mathfrak{e}(\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\chi\varsigma}) &= ||t[\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\chi\varsigma}]x_\chi||_\chi^2 \\
&= ||t[\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\chi\varsigma}]x_\chi||_0^2 + 2\chi\,\mathrm{Re}(\varsigma\beta_\chi\overline{\beta'_\chi}) && (2.51)\\
&\leq \mathfrak{e}(\mathsf{C}^{\zeta},\mathsf{D}^{\mu}) + 2\chi\,\mathrm{Re}(\varsigma\beta_\chi\overline{\beta'_\chi}) + O(\chi) && (2.52)\\
&\leq \mathfrak{e}(\mathsf{C}^{\zeta},\mathsf{D}^{\mu}) + O(\chi) && (2.53)
\end{aligned}$$

Here, (2.52) follows from (2.51) since x_χ is a unit vector in $\mathfrak{X}(\mathsf{C}^\zeta)$ and (2.53) follows from (2.52) since $||t[\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\varsigma\chi}]x_\chi|| = O(1)$ and therefore Proposition 2.14 shows $\max(|\beta_\chi|,|\beta'_\chi|) = O(1)$. □

2.5.5 Differentiability of the energy

We emphasize that for any extension parameter λ, the space $\mathfrak{X}(\mathsf{C}^\lambda)$ has a canonical basis as in (2.11).

Proposition 2.16. *Suppose the space of vectors which achieve the norm of $t[\mathsf{C},\mathsf{D}]$ is one dimensional and let $\varsigma \in \partial\mathbb{D}$. Then for sufficiently small χ the quantity $\mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma},\mathsf{D}^{\mu})$ is a differentiable function of χ. Moreover, for every such χ there is a vector $x_\chi \in \mathfrak{X}(\mathsf{C}^{\zeta+\chi\varsigma})$ which achieves the norm of $t[\mathsf{C}^{\zeta+\chi\varsigma},\mathsf{D}^{\mu}]$ and such that the coordinates of x_χ with respect to the canonical basis are differentiable functions of χ.*

Similarly, for sufficiently small χ the quantity $\mathfrak{e}(\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\chi\varsigma})$ is a differentiable function of χ. Moreover, for every such χ there is a vector $y_\chi \in \mathfrak{X}(\mathsf{D}^{\mu+\chi\varsigma})$ which is the image of a vector achieving the norm of $t[\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\chi\varsigma}]$ and such that the coordinates of y_χ with respect to the canonical basis are differentiable functions of χ. We may assume that y_0 is the image of a unit vector.

Proof of Proposition 2.16. We establish Proposition 2.16 for perturbations of ζ. The case of perturbations of μ can be established using a similar method. Let B_χ be the orthonormal basis for $\mathfrak{X}(\mathsf{C}^{\zeta+\chi\varsigma})$ obtained by applying the Gram-Schmidt orthonormalization procedure to the basis

$$\{\Theta_{\mathsf{C}}(h)_m : (h,m) \in \mathcal{P}(g,d,j,k)\} \cup \{\Theta_{\mathsf{C}}(e)_k, \overline{\varsigma}\Theta_{\mathsf{C}}(g)_j\} \tag{2.54}$$

Let $\mathcal{N}_\chi$ be the matrix which changes coordinates from the basis in (2.54) to B_χ. Since we have introduced the phase $\overline{\varsigma}$ to these bases, the matrix $\mathcal{N}_\chi$ is a real perturbation of the matrix $\mathcal{N}_0$.

Also let D be the orthonormal basis for $\mathfrak{X}(\mathsf{D}^\mu)$ obtained by applying the Gram-Schmidt orthonormalization procedure to the basis $\{\Theta_{\mathsf{D}}(h)_m : (h,m) \in \mathcal{Q}(g,d,j,k)\}$ and let Ω be the matrix which changes coordinates from the basis in (2.54) to D. With respect to the bases B_χ and D the matrix of the transport operator

$t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^{\mu}]$ is given by $\Omega\mathcal{N}_{\chi}^{-1}$. The advantage of writing the matrix with respect to these orthonormal bases is that the matrix of the adjoint $t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^{\mu}]^{*}$ is given by the conjugate transpose of the matrix $\Omega\mathcal{N}_{\chi}^{-1}$, which we denote simply by $(\Omega\mathcal{N}_{\chi}^{-1})^{*}$. Let $\xi_{\chi} = (\Omega\mathcal{N}_{\chi}^{-1})^{*}\Omega\mathcal{N}_{\chi}^{-1}$

The second clause in 2.8 implies the entries of ξ_{χ} are differentiable functions of χ. Therefore Theorem 6.1 in [9] implies that there differentiable functions $\kappa_1(\chi), \ldots, \kappa_n(\chi)$ which represent the eigenvalues of ξ_{χ} for sufficiently small values of χ. We may assume that $\kappa_m(0) \geq \kappa_{m+1}(0)$ for all $m \in [n-1]$. Since $\kappa_1(0) > \kappa_2(0)$, for all sufficiently small χ the function $\kappa_1(\chi)$ represents the norm of ξ_{χ}. This complete the proof of the first claim in Proposition 2.16.

Recall that the index set for the matrix ξ_{χ} is $\mathcal{Q}(g, d, j, k)$. Theorem 6.1 in [9] also implies that for each χ there are vectors $\varphi_1(\chi), \ldots, \varphi_n(\chi) \in \ell^2(\mathcal{Q}(g, d, j, k))$ such that $\varphi_m(\chi)$ is an eigenvector of ξ_{χ} with eigenvalue $\kappa_m(\chi)$ and such that the coordinates of $\varphi_m(\chi)$ are differentiable functions of χ. These numerical vectors represent the coordinates of the singular vectors for $t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^{\mu}]$ in the basis B_{χ}. In order to change the coordinates back to the basis in (2.54) we need to multiply $\varphi_m(\chi)$ by $\mathcal{N}_{\chi}^{-1}$. Since the entries of $\mathcal{N}_{\chi}^{-1}$ are differentiable functions of χ the second claim in Proposition 2.16 follows. □

2.5.6 Calculation of derivatives

Proposition 2.17. *Suppose that for all sufficiently small χ the space of vectors which achieve the norm of $t[\mathsf{C}^{\zeta+\varsigma\chi}, \mathsf{D}^{\mu}]$ is one-dimensional. Let $x \in \mathfrak{X}(\mathsf{C}^{\zeta})$ be a unit vector which achieves the norm of $t[\mathsf{C}^{\zeta}, \mathsf{D}^{\mu}]$ and write $x = \alpha S + \alpha' S' + x'$ for $x' \in \operatorname{core}(\mathfrak{X}(\mathsf{C})_{\bullet})$. Then we have*

$$\frac{\mathrm{d}}{\mathrm{d}\chi}\mathfrak{e}(\mathsf{C}^{\zeta+\varsigma\chi}, \mathsf{D}^{\mu})\bigg|_{\chi=0} = -2\,\mathfrak{e}(\mathsf{C}^{\zeta}, \mathsf{D}^{\mu})\operatorname{Re}(\varsigma\alpha\overline{\alpha'})$$

Also let $y \in \mathfrak{X}(\mathsf{D}^{\mu})$ be given by $y = t[\mathsf{C}^{\zeta}, \mathsf{D}^{\mu}]x$ and write $y = \beta T + \beta' T' + y'$ for $y' \in \mathfrak{X}(\mathsf{D})$. Then we have

$$\frac{\mathrm{d}}{\mathrm{d}\chi}\mathfrak{e}(\mathsf{C}^{\zeta}, \mathsf{D}^{\mu+\varsigma\chi})\bigg|_{\chi=0} = 2\operatorname{Re}(\varsigma\beta\overline{\beta'})$$

Proof of Proposition 2.17. Let $x_{\chi} \in \mathfrak{X}(\mathsf{C}^{\zeta+\chi\varsigma})$ be as in Proposition 2.16. Write $x_{\chi} = \alpha_{\chi} S + \alpha'_{\chi} S' + x'_{\chi}$ for $x'_{\chi} \in \operatorname{core}(\mathfrak{X}(\mathsf{C})_{\bullet})$.

In order to distinguish between the norms on different spaces, we will write $||\cdot||_{\chi}$ for the norm on $\mathfrak{X}(\mathsf{C}^{\zeta+\chi\varsigma})$. We have

$$\begin{aligned}||x_{\chi}||_{\chi}^{2} &= ||x_{\chi}||_{0}^{2} + 2\chi\operatorname{Re}(\varsigma\alpha_{\chi}\overline{\alpha'_{\chi}}) \\ &= ||x_{\chi}||_{0}^{2} + 2\chi\operatorname{Re}(\varsigma(\alpha_{\chi}-\alpha)\overline{\alpha'_{\chi}}) + 2\chi\operatorname{Re}(\varsigma\alpha(\overline{\alpha'_{\chi}} - \overline{\alpha'})) + 2\chi\operatorname{Re}(\varsigma\alpha\overline{\alpha'})\end{aligned} \tag{2.55}$$

By Proposition 2.16 we see that $\alpha_{\chi} = \alpha + o(1)$ and $\alpha'_{\chi} = \alpha' + o(1)$. Therefore in (2.55) we have that the terms

$$2\chi\operatorname{Re}(\varsigma(\alpha_{\chi}-\alpha)\overline{\alpha'_{\chi}})$$

and

$$2\chi\operatorname{Re}(\varsigma\alpha(\overline{\alpha'_{\chi}} - \overline{\alpha'}))$$

are $o(\chi)$. It follows that

$$||x_\chi||_\chi^2 = ||x_\chi||_0^2 + 2\chi \mathrm{Re}(\varsigma\alpha\overline{\alpha'}) + o(\chi) \tag{2.56}$$

We compute

$$\mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu) = ||t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu]x_\chi||^2\, ||x_\chi||_\chi^{-2} \tag{2.57}$$
$$= ||t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu]x_\chi||^2(||x_\chi||_0^2 + 2\chi\, \mathrm{Re}(\varsigma\alpha\overline{\alpha'}) + o(\chi))^{-1} \tag{2.58}$$
$$= ||t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu]x_\chi||^2\, ||x_\chi||_0^{-2}(1 + 2\chi||x_\chi||_0^{-2}\, \mathrm{Re}(\varsigma\alpha\overline{\alpha'}) + o(\chi))^{-1} \tag{2.59}$$
$$= ||t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu]x_\chi||^2\, ||x_\chi||_0^{-2}(1 - 2\chi||x_\chi||_0^{-2}\, \mathrm{Re}(\varsigma\alpha\overline{\alpha'}) + o(\chi)) \tag{2.60}$$
$$= ||t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu]x_\chi||^2\, ||x_\chi||_0^{-2}(1 - 2\chi\, \mathrm{Re}(\varsigma\alpha\overline{\alpha'}) + o(\chi)) \tag{2.61}$$
$$= ||t[\mathsf{C}^{\zeta}, \mathsf{D}^\mu]x_\chi||^2\, ||x_\chi||_0^{-2}(1 - 2\chi\, \mathrm{Re}(\varsigma\alpha\overline{\alpha'}) + o(\chi)) \tag{2.62}$$
$$\leq \mathfrak{e}(\mathsf{C}^{\zeta}, \mathsf{D}^\mu)(1 - 2\chi\, \mathrm{Re}(\varsigma\alpha\overline{\alpha'}) + o(\chi)) \tag{2.63}$$

This computation can be justified as follows.

- (2.58) follows from (2.57) by (2.56)
- (2.59) follows from (2.58) since Proposition 2.16 together with the hypothesis that $||x||_0 = 1$ implies $||x_\chi||_0 = 1 + o(1)$, so that $||x_\chi||_0$ can be absorbed into the $o(\chi)$ term.
- (2.60) follows from (2.59) since the higher terms in the geometric series can be absorbed into the $o(\chi)$ term
- (2.61) follows from (2.60) again since $||x_\chi||_0 = 1 + o(1)$
- (2.62) follows from (2.61) since $||t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu]x_\chi||$ is computed in $\mathcal{X}(\mathsf{D}^\mu)$

From the above computation we see that

$$\frac{\mathrm{d}}{\mathrm{d}\chi}\mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma})\bigg|_{\chi=0} \leq -2\,\mathfrak{e}(\mathsf{C}^{\zeta}, \mathsf{D}^\mu)\,\mathrm{Re}(\varsigma\alpha\overline{\alpha'}) \tag{2.64}$$

On the other hand, we have

$$\mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu) \geq ||t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu]x||^2\, ||x||_\chi^{-2} \tag{2.65}$$
$$= ||t[\mathsf{C}^{\zeta}, \mathsf{D}^\mu]x||^2\, ||x||_\chi^{-2} \tag{2.66}$$
$$= \mathfrak{e}(\mathsf{C}^{\zeta}, \mathsf{D}^\mu)||x||_\chi^{-2} \tag{2.67}$$
$$= \mathfrak{e}(\mathsf{C}^{\zeta}, \mathsf{D}^\mu)(1 + 2\chi\, \mathrm{Re}(\varsigma\alpha\overline{\alpha'}))^{-1} \tag{2.68}$$
$$= \mathfrak{e}(\mathsf{C}^{\zeta}, \mathsf{D}^\mu)(1 - 2\chi\, \mathrm{Re}(\varsigma\alpha\overline{\alpha'}) + o(\chi)) \tag{2.69}$$

The above computation can be justified as follows.

- (2.66) follows from (2.65) since $||t[\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu]x||$ is computed in $\mathcal{X}(\mathsf{D}^\mu)$ and hence is independent of χ
- (2.67) follows from (2.66) since we assumed x was a unit vector in $\mathcal{X}(\mathsf{C}^\zeta)$ which achieves the norm of $t[\mathsf{C}^\zeta, \mathsf{D}^\mu]$,

- (2.69) follows from (2.68) since the higher terms in the geometric series can be absorbed into the $o(\chi)$ term.

From the above computation and (2.64) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\chi}\mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma})\bigg|_{\chi=0} = -2\,\mathfrak{e}(\mathsf{C}^{\zeta},\mathsf{D}^{\mu})\,\mathrm{Re}(\varsigma\alpha\overline{\alpha'})$$

This establishes Proposition 2.17 for perturbations of ζ.

Now, let $y_\chi \in \mathcal{X}(\mathsf{D}^{\mu+\chi\varsigma})$ be as in Proposition 2.16, so that y_χ is the image of a vector which achieves the norm of $t[\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\chi\varsigma}]$ and the coordinates of y_χ with respect to the canonical basis (2.11) are continuous functions of χ. Write $y_\chi = \beta_\chi T + \beta'_\chi T' + y'_\chi$ for $y' \in \mathrm{core}(\mathcal{X}(\mathsf{D})_\bullet)$. We will alter the notation $||\cdot||_\chi$ to now refer to the norm on $\mathsf{D}^{\mu+\chi\varsigma}$. We have

$$||y_\chi||_\chi = ||y_\chi||_0 + 2\chi\,\mathrm{Re}(\varsigma\beta_\chi\overline{\beta'_\chi})$$

so that the same argument used to establish (2.56) shows that

$$||y_\chi||_\chi = ||y_\chi||_0 + 2\chi\,\mathrm{Re}(\varsigma\beta\overline{\beta'}) + o(\chi) \tag{2.70}$$

We compute

$$\begin{aligned}
\mathfrak{e}(\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\chi\varsigma}) &= ||y_\chi||_\chi^2||t[\mathsf{D}^{\mu+\chi\varsigma},\mathsf{C}]y_\chi||^{-2} &(2.71)\\
&= ||y_\chi||_\chi^2||t[\mathsf{D}^{\mu},\mathsf{C}^{\zeta}]y_\chi||^{-2} &(2.72)\\
&= (||y_\chi||_0^2 + 2\chi\,\mathrm{Re}(\varsigma\beta\overline{\beta}) + o(\chi))||t[\mathsf{D}^{\mu},\mathsf{C}^{\zeta}]y_\chi||^{-2} &(2.73)\\
&= ||y_\chi||_0^2||t[\mathsf{D}^{\mu},\mathsf{C}^{\zeta}]y_\chi||^{-2} + 2\chi\,\mathrm{Re}(\varsigma\beta\overline{\beta})||t[\mathsf{D}^{\mu},\mathsf{C}^{\zeta}]y_\chi||^{-2} + o(\chi) &(2.74)\\
&\le \mathfrak{e}(\mathsf{C}^{\zeta},\mathsf{D}^{\mu}) + 2\chi\,\mathrm{Re}(\varsigma\beta\overline{\beta})||t[\mathsf{D}^{\mu},\mathsf{C}^{\zeta}]y_\chi||^{-2} + o(\chi) &(2.75)\\
&= \mathfrak{e}(\mathsf{C}^{\zeta},\mathsf{D}^{\mu}) + 2\chi\,\mathrm{Re}(\varsigma\beta\overline{\beta})(1+o(1)) + o(\chi) &(2.76)\\
&= \mathfrak{e}(\mathsf{C}^{\zeta},\mathsf{D}^{\mu}) + 2\chi\,\mathrm{Re}(\varsigma\beta\overline{\beta}) + o(\chi)
\end{aligned}$$

This computation can be justified as follows.

- (2.72) follows from (2.71) since $||t[\mathsf{D}^{\mu},\mathsf{C}^{\zeta}]y_\chi||$ is computed in $\mathcal{X}(\mathsf{C}^{\zeta})$
- (2.73) follows from (2.72) by (2.70)
- (2.74) follows from (2.73) since y_0 was assumed to be the image of a unit vector in $\mathcal{X}(\mathsf{C}^{\zeta})$ and therefore Proposition 2.16 implies
$$||t[\mathsf{D}^{\mu},\mathsf{C}^{\zeta}]y_\chi||^{-2} = 1 + o(1)$$
- (2.76) follows from (2.75) by the same reasoning as in the previous bullet.

The above computation shows that

$$\frac{\mathrm{d}}{\mathrm{d}\chi}\mathfrak{e}(\mathsf{C}^{\zeta},\mathsf{D}^{\mu+\chi\varsigma})\bigg|_{\chi=0} \le 2\,\mathrm{Re}(\varsigma\beta\overline{\beta}) \tag{2.77}$$

On the other hand, we have

$$\mathfrak{e}(\mathsf{C}^\zeta, \mathsf{D}^{\mu+\chi\varsigma}) \geq ||y||_\chi^2 ||t[\mathsf{D}^{\mu+\varsigma\chi}, \mathsf{C}^\zeta]y||^{-2} \tag{2.78}$$
$$= ||y||_\chi^2 \tag{2.79}$$
$$= ||y||_0^2 + 2\chi \operatorname{Re}(\varsigma\beta\overline{\beta'}) \tag{2.80}$$
$$= \mathfrak{e}(\mathsf{C}^\zeta, \mathsf{D}^\mu) + 2\chi \operatorname{Re}(\varsigma\beta\overline{\beta'}) \tag{2.81}$$

Here, (2.79) follows from (2.78) since y was assumed to be the image of a unit vector in $\mathcal{X}(\mathsf{C}^\zeta)$ and (2.81) follows from (2.80) since y was assumed to achieve the norm of $t[\mathsf{C}^\zeta, \mathsf{D}^\mu]$. □

2.5.7 Relationship between extension scalars before and after transport

Proposition 2.18. *Let $x \in \mathcal{X}(\mathsf{C})_\bullet$ and write $x = \alpha S + \alpha' S' + x'$ for $x' \in \operatorname{core}(\mathcal{X}(\mathsf{C})_\bullet)$. Let $\zeta, \mu \in \mathbb{D}$ be arbitrary and write $t[\mathsf{C}^\zeta, \mathsf{D}^\mu]x = \beta T + \beta' T' + y'$ for $y' \in \mathcal{X}(\mathsf{D})$. Then we have*

$$\beta\overline{\beta'} = \alpha\overline{\alpha'} \frac{||(I-p)\Theta_\mathsf{C}(g)_j||\, ||(I-p)\Theta_\mathsf{C}(e)_k||}{||(I-p)\Theta_\mathsf{D}(g)_j||\, ||(I-q)\Theta_\mathsf{D}(e)_k||} \tag{2.82}$$

Proof of Proposition 2.18. Inspecting the definition (2.32) we see that the quantity $\alpha||(I-p)\Theta_\mathsf{C}(g)_j||$ is the coefficient of $\Theta_\mathsf{C}(g)_j$ in the expression of x with respect to the canonical basis (2.11) for $\mathcal{X}(\mathsf{C})_\bullet$. From Definition 2.2 we see that $\alpha||(I-p)\Theta_\mathsf{C}(g)_j||$ is the coefficient of $\Theta_\mathsf{D}(g)_j$ in the expression of $t[\mathsf{C}^\zeta, \mathsf{D}^\mu]x$ with respect to the canonical basis (2.11) for $\mathcal{X}(\mathsf{D})_\bullet$. Thus from (2.34) we see that

$$\beta = \alpha \frac{||(I-p)\Theta_\mathsf{C}(g)_j||}{||(I-p)\Theta_\mathsf{D}(g)_j||}$$

Similarly, we have

$$\beta' = \alpha' \frac{||(I-p)\Theta_\mathsf{C}(e)_k||}{||(I-p)\Theta_\mathsf{D}(e)_k||}$$

and Proposition 2.18 follows. □

2.5.8 Proof of Lemma 2.3 for a tree

We can now obtain the following strong form of Lemma 2.3 for a single edge.

Proposition 2.19. *Let $g \in \mathbb{F}$, let $d \in \mathbb{N}$ and let $j, k \in [d]$. Let C and D be elements of $\operatorname{NSPD}(g, d, j, k)$ such that the pair (C, D) has singular degeneracies. Then for any $\mu \in \mathbb{D}$ there exists $\zeta \in \mathbb{D}$ such that $\mathfrak{e}(\mathsf{C}^\zeta, \mathsf{D}^\mu) = \mathfrak{e}(\mathsf{C}, \mathsf{D})$.*

Proof of Proposition 2.19. Let g, d, j, k, C and D be as in Statement 2.19. Fix $\mu \in \mathbb{D}$. The hypothesis of completely singular degeneracies implies that the function $\zeta \mapsto \mathfrak{e}(\mathsf{C}^\zeta, \mathsf{D}^\mu)$ attains its minimum on $\mathbb{D}$. Fix ζ at which this minimum is attained and suppose toward a contradiction that $\mathfrak{e}(\mathsf{C}^\zeta, \mathsf{D}^\mu) > \mathfrak{e}(\mathsf{C}, \mathsf{D})$. Let $x \in \mathcal{X}(\mathsf{C}^\zeta)$ be a unit vector which achieves the norm of $t[\mathsf{C}^\zeta, \mathsf{D}^\mu]$ and write $x = \alpha S + \alpha' S' + x'$ for $x' \in \operatorname{core}(\mathcal{X}(\mathsf{C})_\bullet)$. Since we must have

$$\frac{\mathrm{d}}{\mathrm{d}\chi} \mathfrak{e}(\mathsf{C}^{\zeta+\chi\varsigma}, \mathsf{D}^\mu)\bigg|_{\chi=0} = 0$$

for all $\varsigma \in \partial\mathbb{D}$, the first clause in Proposition 2.17 implies $\operatorname{Re}(\varsigma\alpha\overline{\alpha'}) = 0$ for all $\varsigma \in \partial\mathbb{D}$. Therefore $\alpha\overline{\alpha'} = 0$. Using Proposition 2.12 this contradicts the hypothesis that $||t[\mathsf{C}^\zeta, \mathsf{D}^\mu]x||^2 > \mathfrak{e}(\mathsf{C}, \mathsf{D})$. □

By applying Proposition 2.19 to each edge moving backwards from the root, we obtain the statement of Lemma 2.3 for a tree directed toward a root.

2.5.9 Proof of Lemma 2.3 for a cycle

Let $g \in \mathbb{F}$, let $d, N \in \mathbb{N}$ and let $j, k \in [d]$. To simplify notation, in Segment 2.5.9 we assume that all indexes in $[N]$ are taken modulo N. Let $\mathsf{D}_1, \ldots, \mathsf{D}_N \in \mathrm{NSPD}(g, d, j, k)$ have completely singular degeneracies. This hypothesis ensures that the function

$$f(\zeta_1, \ldots, \zeta_N) = \sum_{n=1}^{N} \left(\mathfrak{e}(\mathsf{D}_n^{\zeta_n}, \mathsf{D}_{n+1}^{\zeta_{n+1}}) - \mathfrak{e}(\mathsf{D}_n, \mathsf{D}_{n+1})\right)^2$$

from $\mathbb{D}^N$ to $[0, \infty)$ attains its minimum. Fix $\zeta_1, \ldots, \zeta_n$ at which this minimum is attained.

For $n \in [N]$ we define S_n and S'_n as in (2.32) and (2.33) relativized to $\mathsf{D}_1, \ldots, \mathsf{D}_n$. If the space of vectors achieving the norm of $t[\mathsf{D}_n^{\zeta_n}, \mathsf{D}_{n+1}^{\zeta_{n+1}}]$ is one-dimensional we let x be a unit vector achieving this norm and write $x = \alpha_n S_n + \alpha'_n S'_n + x'$ for $x' \in \mathrm{core}(\mathcal{X}(\mathsf{D}_n)_\bullet)$. The scalars α_n and α'_n may depend on x but the quantity $\alpha_n \overline{\alpha'_n}$ is well-defined. We also write $t[\mathsf{D}_n^{\zeta_n}, \mathsf{D}_{n+1}^{\zeta_{n+1}}]x = \beta_n S_n + \beta'_n S'_n + y$ for $y \in \mathrm{core}(\mathcal{X}(\mathsf{D}_{n+1})_\bullet)$. If the space of vectors achieving the norm of $t[\mathsf{D}_n, \mathsf{D}_{n+1}]$ has dimension greater than one we set $\alpha_n = \alpha'_n = \beta_n = \beta'_n = 0$.

Write $\mathfrak{f}_n = \mathfrak{e}(\mathsf{D}_n^{\zeta_n}, \mathsf{D}_{n+1}^{\zeta_{n+1}} - \mathfrak{e}(\mathsf{D}_n, \mathsf{D}_{n+1})$ and $\mathfrak{e}_n = \mathfrak{e}(\mathsf{D}_n, \mathsf{D}_{n+1})$. From Proposition 2.17 we find

$$\frac{\mathrm{d}}{\mathrm{d}\chi} f(\zeta_1, \ldots, \zeta_n + \chi\varsigma, \ldots, \zeta_N)\bigg|_{\chi=0} = 4\mathrm{Re}\Big(\varsigma\big(\mathfrak{f}_{n-1}\beta_{n-1}\overline{\beta'_{n-1}} - \mathfrak{f}_n\mathfrak{e}_n\alpha_n\overline{\alpha'_n}\big)\Big)$$

Thus from the minimization hypothesis we find

$$\mathfrak{f}_{n-1}\beta_{n-1}\overline{\beta'_{n-1}} = \mathfrak{f}_n\mathfrak{e}_n\alpha_n\overline{\alpha'_n} \tag{2.83}$$

for all $n \in [N]$. Define

$$\mathfrak{p}_n = ||(I-p)\Theta_{\mathsf{D}_n}(g)_j||\,||(I-p)\Theta_{\mathsf{D}_n}(e)_k||$$

From Proposition 2.18 we find

$$\beta_n\overline{\beta'_n} = \mathfrak{p}_n\mathfrak{p}_{n+1}^{-1}\alpha_n\overline{\alpha'_n} \tag{2.84}$$

From (2.83) and (2.84) we have

$$\mathfrak{f}_{n-1}\mathfrak{p}_{n-1}\mathfrak{p}_n^{-1}\alpha_{n-1}\overline{\alpha'_{n-1}} = \mathfrak{f}_n\mathfrak{e}_n\alpha_n\overline{\alpha'_n}$$

By induction we obtain

$$\mathfrak{f}_{n-m}\mathfrak{p}_{n-m}\mathfrak{p}_n^{-1}\alpha_{n-m}\overline{\alpha'_{n-m}} = \left(\prod_{l=0}^{m-1}\mathfrak{e}_{n-l}\right)\mathfrak{f}_n\alpha_n\overline{\alpha'_n}$$

and therefore choosing $m = N$ we obtain

$$\alpha_n\overline{\alpha'_n} = \left(\prod_{l=0}^{N=1}\mathfrak{e}_{n-l}\right)\alpha_n\overline{\alpha'_n}$$

It is clear that the hypothesis of completely singular degeneracies implies $\mathfrak{e}_m > 1$ for all $m \in [N]$. Therefore we obtain $\alpha_n\overline{\alpha'_n} = 0$ and the proof of Lemma 2.3 is complete.

3 Proof of Theorem 1.2

In Section 3 we prove Theorem 1.2.

3.1 Representation theory of $\mathbb{F} \times \mathbb{F}$

In Subsection 3.1 we conduct an analysis of unitary representations of $\mathbb{F} \times \mathbb{F}$

3.1.1 Unitary approximate conjugacy of representations

We will use the theory of weak containment of unitary representations of countable discrete groups, for which we refer the reader to Appendix H of [10]. We will say that a unitary representation of a countable discrete group G is maximal if it weakly contains every other unitary representation of G.

If G is a countable discrete group, $\mathcal{X}$ is a Hilbert space and $\rho : G \to \mathrm{U}(\mathcal{X})$ is a unitary representation, there is a unique extension of ρ to a $*$-homomorphism from $C^*(G)$ to the algebra $\mathrm{B}(\mathcal{X})$ of bounded operators on $\mathcal{X}$. We denote this extension by $\widetilde{\rho}$. Let $\xi : G \to \mathrm{U}(\mathcal{Y})$ be another unitary representation, potentially on a different Hilbert space. By Theorem F.4.4 in [3], if ξ is weakly contained in ρ then $||\widetilde{\xi}(s)||_{\mathrm{op}} \leq ||\widetilde{\rho}(s)||_{\mathrm{op}}$ for all $s \in C^*(G)$. It follows that if ρ is a maximal unitary representation then $\widetilde{\rho}$ is injective. We now recall a different notion of approximation for representations.

Definition 3.1. *Unitary representations $\rho : G \to \mathrm{U}(\mathcal{X})$ and $\xi : G \to \mathrm{U}(\mathcal{Y})$ are said to be **unitarily approximately conjugate** if there is a sequence of unitary operators $u_n : \mathcal{X} \to \mathcal{Y}$ such that for each $g \in G$ we have*

$$\lim_{n\to\infty} ||u_n^{-1}\xi(g)u_n - \rho(g)||_{\mathrm{op}} = 0.$$

The following is a special case of Corollary 1.7.5 in [4].

Theorem 3.1 (Voiculescu)**.** *Let G be a countable discrete group. Suppose ξ and ρ are unitary representations of G such that $\widetilde{\xi}$ and $\widetilde{\rho}$ are injective and such that $\widetilde{\xi}(C^*(G))$ and $\widetilde{\rho}(C^*(G))$ contain no nonzero compact operators. Then ξ and ρ are unitarily approximately conjugate.*

We can now connect weak containment and unitary approximate conjugacy.

Proposition 3.1. *Suppose ξ and ρ are maximal unitary representations of $\mathbb{F}$. Then ξ and ρ are unitarily approximately conjugate.*

Proof of Proposition 3.1. By Corollary VII.6.7 in [5] the image of an injective representation of $C^*(\mathbb{F})$ contains no nonzero compact operators. Thus Proposition 3.1 follows from Theorem 3.1. □

3.1.2 The negation of Connes' embedding conjecture

We now translate the negation of Connes' embedding conjecture into a representation theoretic statement.

Definition 3.2. *Let $\mathcal{X}$ be a Hilbert space and let G, H be countable discrete groups. We define a linear representation $\zeta : G \times H \to \mathrm{GL}(\mathcal{X})$ to be **half finite** if there exist a finite quotient Γ of G and a linear representation $\zeta_\bullet : \Gamma \times H \to \mathrm{GL}(\mathcal{X})$ such that ζ factors as $\zeta_\bullet$ precomposed with $\Pi \times \iota$, where $\Pi : G \twoheadrightarrow \Gamma$ is the quotient map and ι is the identity map on H.*

Definition 3.3. *Let $\mathcal{X}$ be a Hilbert space, let $\rho : G \times H \to \mathrm{U}(\mathcal{X})$ be a unitary representation and let $x \in \mathcal{X}$ be a unit vector. Let E be a finite subset of G and F be a finite subset of H and let $\epsilon > 0$. We define a **half-finite approximation** to $(\rho, x, E, F, \epsilon)$ to be a Hilbert space $\mathcal{Y}$, a finite quotient Γ of G, a half finite unitary representation $\zeta : G \times H \to \Gamma \times H \to \mathrm{U}(\mathcal{Y})$ and a unit vector $y \in \mathcal{Y}$ such that*

$$|\langle \rho(g, g')x, x\rangle - \langle \zeta(g, g')y, y\rangle| \leq \epsilon$$

*for all $g \in E$ and $g' \in F$. We define the pair (G, H) to have the **half-finite approximation property** if there exists a half-finite approximation for every $(\rho, x, E, F, \epsilon)$ as above.*

If A and B are C^*-algebras, we will write $A \otimes_{\max} B$ for the maximal tensor product and $A \otimes_{\min} B$ for the minimal tensor product. For information about tensor products of operator algebras we refer the reader to Chapter 11 of [8]. If G is a countable discrete group, we will write $C^*(G)$ for the full group C^*-algebra of G. For information about group C^*-algebras we refer the reader to Chapter VII of [5].

Proposition 3.2. *Suppose (G, H) has the half-finite approximation property. Then*

$$C^*(G) \otimes_{\max} C^*(H) = C^*(G) \otimes_{\min} C^*(H)$$

Proof of Proposition 3.2. Write $||\cdot||_{\max}$ for the norm on $C^*(G) \otimes_{\max} C^*(H)$ and $||\cdot||_{\min}$ for the norm on $C^*(G) \otimes_{\min} C^*(H)$. Fix an element ϕ in the group ring $\mathbb{C}[G \times H]$ such that $||\phi||_{\max} = 1$. In order to prove Theorem **??** suffices to show that $||\phi||_{\min} = 1$. To this end, let $\sigma > 0$.

Since $||\phi||_{\max} = 1$ we can find a Hilbert space $\mathcal{X}$, a unitary representation $\rho : G \times H \to \mathrm{U}(\mathcal{X})$ and a unit vector $x \in \mathcal{X}$ such that $||\rho(\phi)x||^2 \geq 1 - \sigma$. Write

$$\phi = \sum_{g \in E} \sum_{h \in F} \alpha_{g,h}(g, h)$$

for finite sets $E \subseteq G$ and $F \subseteq H$ and complex numbers $(\alpha_{g,h})_{g \in E, h \in F}$. Let $\epsilon > 0$ be such that

$$\epsilon \left(\sum_{g \in E} \sum_{h \in F} |\alpha_{g,h}| \right)^2 \leq \sigma$$

Apply the half-finite approximation property to find a half-finite approximation to $(\rho, x, E^{-1}E, F^{-1}F, \epsilon)$. We obtain $\mathcal{Y}, \Gamma, \zeta$ and y. We have

$$\begin{aligned}
&|\langle \rho(\phi)x, \rho(\phi)x\rangle - \langle \zeta(\phi)y, \zeta(\phi)y\rangle| \\
&\quad = \Bigg| \left\langle \left(\sum_{g \in E} \sum_{h \in F} \alpha_{g,h}\rho(g, h) \right) x, \left(\sum_{g' \in E} \sum_{h' \in F} \alpha_{g',h'}\rho(g', h') \right) x \right\rangle \\
&\qquad\qquad - \left\langle \left(\sum_{g \in E} \sum_{h \in F} \alpha_{g,h}\zeta(g, h) \right) y, \left(\sum_{g' \in E} \sum_{h' \in F} \alpha_{g',h'}\zeta(g', h') \right) y \right\rangle \Bigg| \\
&\quad = \Bigg| \sum_{g,g' \in E} \sum_{h,h' \in F} \alpha_{g,h}\overline{\alpha_{g',h'}}\langle \rho(g, h)x, \rho(g', h')x \ \rangle
\end{aligned}$$

$$- \sum_{g,g' \in E} \sum_{h,h' \in F} \alpha_{g,h} \overline{\alpha_{g',h'}} \langle \zeta(g,h)y, \zeta(g',h')y \rangle \Bigg|$$

$$= \Bigg| \sum_{g,g' \in E} \sum_{h,h' \in F} \alpha_{g,h} \overline{\alpha_{g',h'}} \langle \rho\big((g')^{-1}g, (h')^{-1}h\big)x, x \rangle$$

$$- \sum_{g,g' \in E} \sum_{h,h' \in F} \alpha_{g,h} \overline{\alpha_{g',h'}} \langle \zeta\big((g')^{-1}g, (h')^{-1}h\big)y, y \rangle \Bigg|$$

$$\leq \sum_{g,g' \in E} \sum_{h,h' \in F} |\alpha_{g,h}||\alpha_{g',h'}| \big| \langle \rho\big((g')^{-1}g, (h')^{-1}h\big)x, x \rangle - \langle \zeta\big((g')^{-1}g, (h')^{-1}h\big)y, y \rangle \big|$$

$$\leq \epsilon \sum_{g,g' \in E} \sum_{h,h' \in F} |\alpha_{g,h}||\alpha_{g',h'}|$$

$$= \epsilon \left(\sum_{g \in E} \sum_{h \in F} |\alpha_{g,h}| \right)^2 \leq \sigma$$

Therefore we obtain $||\zeta(\phi)y||^2 \geq ||\rho(\phi)x||^2 - \sigma$ so that $||\zeta(\phi)||^2_{\mathrm{op}} \geq 1 - 2\sigma$.

There is a natural commutative diagram

$$\begin{array}{ccc} C^*(G) \otimes_{\max} C^*(H) & \longrightarrow & C^*(G) \otimes_{\min} C^*(H) \\ \downarrow & & \downarrow \\ C^*(G) \otimes_{\max} C^*(H) & \longrightarrow & C^*(\Gamma) \otimes_{\min} C^*(H) \end{array}$$

where all the arrows represent surjective $*$-homomorphisms. Moreover, there are canonical copies of ϕ in each of the above algebras. Since ζ factors through $\Gamma \times H$, we see that the norm of ϕ in the bottom left corner is at least $\sqrt{1 - 2\sigma}$. Since $C^*(\Gamma)$ is finite dimensional, Lemma 11.3.11 in [8] implies the arrow across the bottom of the above diagram is an isomorphism. It follows that the norm of ϕ in the bottom right corner is at least $\sqrt{1 - 2\sigma}$ and so $||\phi||_{\min} \geq \sqrt{1 - 2\sigma}$. Since $\sigma > 0$ was arbitrary we obtain $||\phi||_{\min} = 1$ as required. □

In [12] Kirchberg showed that Connes' embedding conjecture is equivalent to the statement

$$C^*(\mathbb{F}) \otimes_{\max} C^*(\mathbb{F}) = C^*(\mathbb{F}) \otimes_{\min} C^*(\mathbb{F})$$

Thus from Proposition 3.2 and the negation of Connes' embedding conjecture we find that the pair $(\mathbb{F}, \mathbb{F})$ fails to have the half-finite approximation property. For the remainder of Section 3 we fix a Hilbert space $\mathcal{X}$, a unitary representation $\rho : \mathbb{F} \times \mathbb{F} \to \mathrm{U}(\mathcal{X})$, a unit vector $x \in X$, $\epsilon > 0$ and finite sets $E, F \subseteq \mathbb{F}$ which witness this failure. It is clear that we may assume ρ is a maximal representation of $\mathbb{F} \times \mathbb{F}$, and indeed we make this assumption.

3.1.3 Choosing initial parameters

Consider the group $\mathbb{F} \times \mathbb{F}$. In order to keep a distinction between the factors, we will write $\mathbb{F}_{\triangleleft}$ for the left copy and $\mathbb{F}_{\triangleright}$ for the right copy. We again fix free generators for each copy and endow them with the corresponding word lengths. We will consistently use the letters g, h for elements of $\mathbb{F}_\triangleleft$ and g', h' for elements of $\mathbb{F}_\triangleright$. If $\mathcal{X}$ is a Hilbert space, $\rho : \mathbb{F}_\triangleleft \times \mathbb{F}_\triangleright \to \mathrm{GL}(\mathcal{X})$ is a linear representation and $\jmath \in \{\triangleleft, \triangleright\}$ we will write $\rho_\jmath$ for the restriction of ρ to $\mathbb{F}_\jmath$. We will also write $\mathbb{B}_{r,\jmath}$ for the ball of radius r around the identity in $\mathbb{F}_\jmath$.

Let $r \in \mathbb{N}$ be such that $E \subseteq \mathbb{B}_{r,\triangleleft}$ and $F \subseteq \mathbb{B}_{r,\triangleright}$. We may assume that $r \geq 5$. We suppose toward a contradiction that $\mathrm{encost}_{r,\epsilon}(\mathbb{F}) = M < \infty$ for our chosen values of r and ϵ.

Choose $R \in \mathbb{N}$ such that

$$\frac{45K_r}{R} \leq \epsilon \tag{3.1}$$

Write

$$L_{r,R} = \left(1 + \exp\left(8R^2(4R+1)^3(10R+1)\log(2M)\right)\right)R \tag{3.2}$$

Choose $\delta > 0$ such that if we write

$$s_\delta = 1536 r K_r (4R+1)^3 R^3 \delta \tag{3.3}$$

then we have

$$160 L_{r,R}^{10}(e^{s_\delta} - 1) \leq \frac{1}{R} \tag{3.4}$$

The rest of the proof of Theorem 1.2 is structured as follows. In the remainder of Subsection 3.1 we continue to analyze unitary representations of $\mathbb{F} \times \mathbb{F}$. In Subsection 3.2 we construct a 4-regular directed graph Θ. In Subsections 3.3 - 3.5 we will use the hypothesis that $\mathrm{encost}_{r,\epsilon}(\Theta) \leq M$ produce a half-finite approximation to $(\rho, x, E, F, \epsilon)$, thereby contradicting our choice of these data.

3.1.4 Approximate conjugacy with the profinite completion

Let $\overline{\mathbb{F}}$ denote the profinite completion of $\mathbb{F}$ and let μ be its Haar probability measure. For each finite quotient Λ of $\mathbb{F}$, there exists a canonical projection $\Pi_\Lambda : \overline{\mathbb{F}} \twoheadrightarrow \Lambda$. Writing $\mathbf{1}_B$ for the indicator function of a subset B of $\overline{\mathbb{F}}$, for each $\lambda \in \Lambda$ we have

$$\begin{aligned}
\left|\left|\mathbf{1}_{\Pi_\Lambda^{-1}(\lambda)}\right|\right|_2 &= \left(\int_{\overline{\mathbb{F}}} \left|\mathbf{1}_{\Pi_\Lambda^{-1}(\lambda)}(\omega)\right|^2 \, \mathrm{d}\mu(\omega)\right)^{\frac{1}{2}} \\
&= \sqrt{\mu\left(\Pi_\Lambda^{-1}(\lambda)\right)} \\
&= \frac{1}{\sqrt{|\Lambda|}}
\end{aligned}$$

Moreover, if λ and λ' are distinct elements of Λ then the sets $\Pi_\Lambda^{-1}(\lambda)$ and $\Pi_\Lambda^{-1}(\lambda')$ are disjoint, so that $\mathbf{1}_{\Pi_\Lambda^{-1}(\lambda)}$ and $\mathbf{1}_{\Pi_\Lambda^{-1}(\lambda')}$ are orthogonal in $L^2(\overline{\mathbb{F}}, \mu)$. Therefore the set of functions

$$\left\{\sqrt{|\Lambda|}\mathbf{1}_{\Pi_{\Lambda^{-1}(\lambda)}} : \lambda \in \Lambda\right\}$$

is orthonormal.

The profinite structure of $\overline{\mathbb{F}}$ guarantees that the collection of sets

$$\left\{\Pi_\Lambda^{-1}(\lambda) : \lambda \in \Lambda,\ \Lambda \text{ is a finite quotient of } \mathbb{F}\right\}$$

generates the Borel σ-algebra of $\overline{\mathbb{F}}$. Therefore we have that the span of the functions

$$\left\{\sqrt{|\Lambda|}\mathbf{1}_{\Pi_\Lambda^{-1}(\lambda)} : \lambda \in \Lambda,\ \Lambda \text{ is a finite quotient of } \mathbb{F}\right\} \tag{3.5}$$

is dense in $L^2(\overline{\mathbb{F}},\mu)$. Choose a sequence $(\Lambda_n)_{n=1}^\infty$ of finite quotients of $\mathbb{F}$ such Λ_n is a quotient of Λ_{n+1} and such that any finite quotient Λ of $\mathbb{F}$ is a quotient of Λ_n for some $n \in \mathbb{N}$. Write Π_n for Π_{Λ_n}. Then the span of the set of functions

$$\bigcup_{n=1}^\infty \left\{\sqrt{|\Lambda_n|}\mathbf{1}_{\Pi_n^{-1}(\lambda)} : \lambda \in \Lambda_n\right\} \tag{3.6}$$

is equal to the span of the set of functions in (3.5). Hence the span of the set of functions in (3.6) is dense in $L^2(\overline{\mathbb{F}},\mu)$. Moreover, the spans of each of the sets inside the union in (3.6) are increasing.

By considering induced representations, we see that since ρ is a maximal representation of $\mathbb{F}_\triangleleft \times \mathbb{F}_\triangleright$ we must have that $\rho_\triangleleft$ is a maximal representation of $\mathbb{F}_\triangleleft$. By Theorem 3.1 in [11], the left translation action of $\mathbb{F}$ on $(\overline{\mathbb{F}},\mu)$ is maximal in the order of weak containment among measure preserving actions of $\mathbb{F}$. We refer the reader to Chapter 10 of [10] for information on this variant of weak containment, but all we will need to know about it is that Proposition 10.5 and Theorem E.1 in [10] imply that the Koopman representation of a maximal action is a maximal representation. Write $\overline{\kappa} : \mathbb{F} \to \mathrm{U}(L^2(\overline{\mathbb{F}},\mu))$ for the Koopman representation of the left translation action, so that Proposition 3.1 implies $\rho_\triangleleft$ and $\overline{\kappa}$ are unitarily approximately conjugate. Let $u : \mathcal{X} \to L^2(\overline{\mathbb{F}},\mu)$ be a unitary operator such that

$$||u^{-1}\overline{\kappa}(g)u - \rho_\triangleleft(g)||_{\mathrm{op}} \leq \delta$$

for all $g \in \mathbb{B}_{1,\triangleleft}$.

Consider the vector ux. Our previous discussion of (3.6) implies that we can find $n \in \mathbb{N}$ and a function $\alpha : \Lambda_n \to \mathbb{C}$ with

$$\left|\left|ux - \sum_{\lambda\in\Lambda_n} \alpha(\lambda)\sqrt{|\Lambda_n|}\mathbf{1}_{\Pi_n^{-1}(\lambda)}\right|\right|_2 \leq \delta$$

We may assume n is large enough that the balls of radius $4R$ in $\mathbb{F}/\Lambda_n$ are isomorphic to the balls of radius $4R$ in $\mathbb{F}$.

The partition $\{\Pi_n^{-1}(\lambda) : \lambda \in \Lambda_n\}$ of $\overline{\mathbb{F}}$ is permuted by the left translation action of $\mathbb{F}$ on $\overline{\mathbb{F}}$ so that $g\Pi_n^{-1}(\lambda) = \Pi_n^{-1}(g\lambda)$ for all $g \in \mathbb{F}$ and $\lambda \in \Lambda_n$. Thus we have $\kappa(g)\mathbf{1}_{\Pi_n^{-1}(\lambda)} = \mathbf{1}_{\Pi_n^{-1}(g\lambda)}$. Since the vectors

$$\left\{\sqrt{|\Lambda_n|}\mathbf{1}_{\Pi_n^{-1}(\lambda)} : \lambda \in \Lambda_n\right\}$$

are orthonormal and x is a unit vector, we may assume that

$$\sum_{\lambda\in\Lambda_n} |\alpha(\lambda)|^2 = 1$$

Enumerate $\Lambda_n = \{\lambda_1, \ldots, \lambda_d\}$ and for $j \in [d]$ let

$$x_j = u^{-1}\sqrt{d}\,\mathbf{1}_{\Pi_n^{-1}(\lambda_j)}$$

We may assume without loss of generality that $d \geq R$. Write $\kappa = u^{-1}\overline{\kappa}u$. We summarize the objects we have just constructed.

- An orthonormal set of vectors $x_1, \ldots, x_d$ in $\mathcal{X}$ and an element $\alpha \in \mathbb{C}^d$ with

$$\sum_{j=1}^{d} |\alpha_j|^2 = 1 \tag{3.7}$$

  such that

$$||x - \alpha_1 x_1 - \cdots - \alpha_d x_d|| \leq \delta \tag{3.8}$$

- An action $\sigma : \mathbb{F}_{\triangleleft} \to \mathrm{Sym}(d)$.
- A unitary representation $\kappa : \mathbb{F}_{\triangleleft} \to \mathrm{U}(\mathcal{X})$ such that

$$||\kappa(g) - \rho_{\triangleleft}(g)||_{\mathrm{op}} \leq \delta \tag{3.9}$$

  for all $g \in \mathbb{B}_{1,\triangleleft}$ and such that

$$\kappa(g)x_j = x_{\sigma(g)j} \tag{3.10}$$

  for all $g \in \mathbb{F}_{\triangleleft}$ and all $j \in [d]$.

3.1.5 Proximity between inner products at individual nodes

Proposition 3.3. *Let $g', h' \in \mathbb{B}_{r,\triangleright}$ and let $\beta, \eta \in \mathbb{C}^d$. Also let $g \in \mathbb{B}_{1,\triangleleft}$ and let $\varsigma \in \mathrm{Sym}(d)$. We have*

$$\left| \left\langle \rho_{\triangleright}(g') \sum_{j=1}^{d} \beta_j x_{\varsigma j}, \rho_{\triangleright}(h') \sum_{k=1}^{d} \eta_k x_{\varsigma k} \right\rangle - \left\langle \rho_{\triangleright}(g') \sum_{j=1}^{d} \beta_j x_{\sigma(g)\varsigma j}, \rho_{\triangleright}(h') \sum_{k=1}^{d} \eta_k x_{\sigma(g)\varsigma k} \right\rangle \right| \leq 2\delta ||\beta||_2 ||\eta||_2$$

Proof of Proposition 3.3. All norms and inner products in the proof of Proposition 3.3 will be in $\mathcal{X}$. From (3.10) we have

$$\begin{aligned} &\left| \left\langle \rho_{\triangleright}(g') \sum_{j=1}^{d} \beta_j x_{\varsigma j}, \rho_{\triangleright}(h') \sum_{k=1}^{d} \eta_k x_{\varsigma k} \right\rangle - \left\langle \rho_{\triangleright}(g') \sum_{j=1}^{d} \beta_j x_{\sigma(g)\varsigma j}, \rho_{\triangleright}(h') \sum_{k=1}^{d} \eta_k x_{\sigma(g)\varsigma k} \right\rangle \right| \\ &= \left| \left\langle \rho_{\triangleright}(g') \sum_{j=1}^{d} \beta_j x_{\varsigma j}, \rho_{\triangleright}(h') \sum_{k=1}^{d} \eta_k x_{\varsigma k} \right\rangle - \left\langle \rho_{\triangleright}(g')\kappa(g) \sum_{j=1}^{d} \beta_j x_{\varsigma j}, \rho_{\triangleright}(h')\kappa(g) \sum_{k=1}^{d} \eta_k x_{\varsigma k} \right\rangle \right| \qquad (3.11) \end{aligned}$$

Write x_β for $\sum_{j=1}^{d} \beta_j x_{\varsigma j}$ and x_η for $\sum_{k=1}^{d} \eta_k x_{\varsigma k}$. We compute

$$\begin{aligned} (3.11) \leq &\left| \left\langle \rho_{\triangleright}(g')x_\beta, \rho_{\triangleright}(h')x_\eta \right\rangle - \left\langle \rho_{\triangleright}(g')\rho_{\triangleleft}(g)x_\beta, \rho_{\triangleright}(h')\rho_{\triangleleft}(g)x_\eta \right\rangle \right| & (3.12) \\ &+ \left| \left\langle \rho_{\triangleright}(g')\kappa(g)x_\beta, \rho_{\triangleright}(h')\kappa(g)x_\eta \right\rangle - \left\langle \rho_{\triangleright}(g')\rho_{\triangleleft}(g)x_\beta, \rho_{\triangleright}(h')\rho_{\triangleleft}(g)x_\eta \right\rangle \right| \\ = &\left| \left\langle \rho_{\triangleright}(g')x_\beta, \rho_{\triangleright}(h')x_\eta \right\rangle - \left\langle \rho_{\triangleleft}(g)\rho_{\triangleright}(g')x_\beta, \rho_{\triangleleft}(g)\rho_{\triangleright}(h')x_\eta \right\rangle \right| & (3.13) \\ &+ \left| \left\langle \rho_{\triangleright}(g')\kappa(g)x_\beta, \rho_{\triangleright}(h')\kappa(g)x_\eta \right\rangle - \left\langle \rho_{\triangleright}(g')\rho_{\triangleleft}(g)x_\beta, \rho_{\triangleright}(h')\rho_{\triangleleft}(g)x_\eta \right\rangle \right| & (3.14) \\ = &\left| \left\langle \rho_{\triangleright}(g')\kappa(g)x_\beta, \rho_{\triangleright}(h')\kappa(g)x_\eta \right\rangle - \left\langle \rho_{\triangleright}(g')\rho_{\triangleleft}(g)x_\beta, \rho_{\triangleright}(h')\rho_{\triangleleft}(g)x_\eta \right\rangle \right| & (3.15) \\ \leq &\left| \left\langle \rho_{\triangleright}(g')\kappa(g)x_\beta, \rho_{\triangleright}(h')\kappa(g)x_\eta \right\rangle - \left\langle \rho_{\triangleright}(g')\kappa(g)x_\beta, \rho_{\triangleright}(h')\rho_{\triangleleft}(g)x_\eta \right\rangle \right| \end{aligned}$$

$$
\begin{aligned}
&\qquad + \left|\left\langle \rho_{\triangleright}(g')\kappa(g)x_\beta, \rho_{\triangleright}(h')\rho_{\triangleleft}(g)x_\eta\right\rangle - \left\langle \rho_{\triangleright}(g')\rho_{\triangleleft}(g)x_\beta, \rho_{\triangleright}(h')\rho_{\triangleleft}(g)x_\eta\right\rangle\right| \\
&= \left|\left\langle \rho_{\triangleright}(g')\kappa(g)x_\beta, \rho_{\triangleright}(h')(\kappa(g)-\rho_{\triangleleft}(g))x_\eta\right\rangle\right| \\
&\qquad + \left|\left\langle \rho_{\triangleright}(g')(\kappa(g)-\rho_{\triangleleft}(g))x_\beta, \rho_{\triangleright}(h')\rho_{\triangleleft}(g)x_\eta\right\rangle\right| \\
&\le ||\rho_{\triangleright}(g')\kappa(g)x_\beta||\,||\rho_{\triangleright}(h')(\kappa(g)-\rho_{\triangleleft}(g))x_\eta|| && (3.16)\\
&\qquad + ||\rho_{\triangleright}(g')(\kappa(g)-\rho_{\triangleleft}(g))x_\beta||\,||\rho_{\triangleright}(h')\rho_{\triangleleft}(g)x_\eta|| && (3.17)\\
&= ||x_\beta||\,||(\kappa(g)-\rho_{\triangleleft}(g))x_\eta|| + ||(\kappa(g)-\rho_{\triangleleft}(g))x_\beta||\,||x_\eta|| && (3.18)\\
&\le 2\,||\kappa(g)-\rho_{\triangleleft}(g)||_{\mathrm{op}}||x_\beta||\,||x_\eta|| && (3.19)\\
&= 2\,||\kappa(g)-\rho_{\triangleleft}(g)||_{\mathrm{op}}||\beta||_2||\eta||_2 && (3.20)\\
&\le 2\delta||\beta||_2||\eta||_2 && (3.21)
\end{aligned}
$$

Here,

- (3.12) is equal to (3.13) since $\rho_{\triangleleft}$ and $\rho_{\triangleright}$ commute,
- (3.15) follows from (3.13) - (3.14) since $\rho_{\triangleleft}$ is unitary and therefore (3.13) is 0,
- (3.18) follows from (3.16) - (3.17) since $\rho_{\triangleright}$ and κ are unitary,
- (3.20) follows from (3.19) since $x_1, \dots, x_d$ is orthonormal,
- and (3.21) follows from (3.20) by (3.9)

Proposition 3.3 follows by combining (3.11) with (3.21). □

3.1.6 Constructing a family of permuted positive definite functions

For $\varsigma \in \mathrm{Sym}(d)$ define a positive definite function $\mathsf{C}_\varsigma : \mathbb{B}_{2r,\triangleright} \to \mathrm{Mat}_{d\times d}(\mathbb{C})$ by setting

$$\mathsf{C}_\varsigma\big((h')^{-1}g'\big)_{j,k} = \langle \rho_{\triangleright}(g')x_{\varsigma j}, \rho_{\triangleright}(h')x_{\varsigma k}\rangle \tag{3.22}$$

for $g', h' \in \mathbb{B}_{r,\triangleright}$. Also define $\Delta_r : \mathbb{B}_{2r,\triangleright} \to \mathrm{Mat}_{d\times d}(\mathbb{C})$ by setting

$$\Delta_r(g') = \begin{cases} \mathbf{I}_d & \text{if } g' = e \\ \mathbf{0}_d & \text{if } g' \in \mathbb{B}_{2r}\setminus\{e\} \end{cases}$$

where $\mathbf{0}_d$ denotes the $d\times d$ zero matrix. Let $\mathsf{D}_\varsigma = (1-R^{-1})\mathsf{C}_\varsigma + R^{-1}\Delta_r$. Note that for any function $\beta : \mathbb{B}_{r,\triangleright} \to \mathbb{C}^d$ we have

$$
\begin{aligned}
&\left|\left|\sum_{g'\in\mathbb{B}_{r,\triangleright}}\sum_{j=1}^{d}\beta(g')_j\Phi_{\mathsf{D}_\varsigma}(g')_j\right|\right|^2 && (3.23)\\
&\quad= \left(1-\frac{1}{R}\right)\left|\left|\sum_{g'\in\mathbb{B}_{r,\triangleright}}\sum_{j=1}^{d}\beta(g')_j\Phi_{\mathsf{C}_\varsigma}(g')_j\right|\right|^2 + \frac{1}{R}\left|\left|\sum_{g'\in\mathbb{B}_{r,\triangleright}}\sum_{j=1}^{d}\beta(g')_j\Phi_{\Delta_r}(g')_j\right|\right|^2 && (3.24)\\
&\quad\ge \frac{1}{R}\left|\left|\sum_{g'\in\mathbb{B}_{r,\triangleright}}\sum_{j=1}^{d}\beta(g')_j\Phi_{\Delta_r}(g')_j\right|\right|^2 && (3.25)
\end{aligned}
$$

$$= \frac{1}{R}\left(\sum_{g'\in\mathbb{B}_{r,\triangleright}}\sum_{j=1}^{d}|\beta(g')_j|^2\right) \tag{3.26}$$

Here, (3.26) follows from (3.25) since the set

$$\{\Phi_{\Delta_r}(g')_j : g'\in\mathbb{B}_{r,\triangleright}, j\in[d]\}$$

is orthonormal.

3.1.7 Establishing bounds on transport operators

Proposition 3.4. *We have* $\mathfrak{e}(\mathsf{D}_\varsigma, \mathsf{D}_{\sigma(g)\varsigma}) \le 1 + 2K_r R\delta$ *for all* $\varsigma \in \mathrm{Sym}(d)$ *and all* $g \in \mathbb{B}_{1,\triangleleft}$.

Proof of Proposition 3.4. Let $\varsigma \in \mathrm{Sym}(d)$ and $g \in \mathbb{B}_{1,\triangleleft}$ Let $\beta : \mathbb{B}_{r,\triangleright} \to \mathbb{C}^d$ be such that if we write

$$y = \sum_{g'\in\mathbb{B}_{r,\triangleright}}\sum_{j=1}^{d}\beta(g')_j\Phi_{\mathsf{D}_\varsigma}(g')_j$$

then y is a unit vector in $\mathcal{X}(\mathsf{D}_\varsigma)$. Thus from (3.26) we have

$$1 \ge \frac{1}{R}\left(\sum_{g'\in\mathbb{B}_{r,\triangleright}}\sum_{j=1}^{d}|\beta(g')_j|^2\right) \tag{3.27}$$

We compute

$$\begin{aligned}
&\left|\|t[\mathsf{D}_\varsigma, \mathsf{D}_{\sigma(g)\varsigma}]y\|^2 - 1\right| \\
&\quad = \Bigg|\left\langle \sum_{g'\in\mathbb{B}_{r,\triangleright}}\sum_{j=1}^{d}\beta(g')_j\Phi_{\mathsf{D}_{\sigma(g)\varsigma}}(g')_j, \sum_{h'\in\mathbb{B}_{r,\triangleright}}\sum_{k=1}^{d}\beta(h')_k\Phi_{\mathsf{D}_{\sigma(g)\varsigma}}(h')_k\right\rangle \\
&\qquad\qquad - \left\langle \sum_{g'\in\mathbb{B}_{r,\triangleright}}\sum_{j=1}^{d}\beta(g')_j\Phi_{\mathsf{D}_\varsigma}(g')_j, \sum_{h'\in\mathbb{B}_{r,\triangleright}}\sum_{k=1}^{d}\beta(h')_k\Phi_{\mathsf{D}_\varsigma}(h')_k\right\rangle\Bigg| &(3.28)\\
&\quad = \Bigg|\sum_{g',h'\in\mathbb{B}_{r,\triangleright}}\sum_{j,k=1}^{d}\beta(g')_j\overline{\beta(h')_k}\mathsf{D}_{\sigma(g)\varsigma}\big((h')^{-1}g'\big)_{j,k} \\
&\qquad\qquad - \sum_{g',h'\in\mathbb{B}_{r,\triangleright}}\sum_{j,k=1}^{d}\beta(g')_j\overline{\beta(h')_k}\mathsf{D}_\varsigma\big((h')^{-1}g'\big)_{j,k}\Bigg| &(3.29)\\
&\quad \le \sum_{g',h'\in\mathbb{B}_{r,\triangleright}}\left|\sum_{j,k=1}^{d}\beta(g')_j\overline{\beta(h')_k}\left(\mathsf{D}_{\sigma(g)\varsigma}\big((h')^{-1}g'\big)_{j,k} - \mathsf{D}_\varsigma\big((h')^{-1}g'\big)_{j,k}\right)\right| &(3.30)\\
&\quad = \left(1-\frac{1}{R}\right)\sum_{g',h'\in\mathbb{B}_{r,\triangleright}}\left|\sum_{j,k=1}^{d}\beta(g')_j\overline{\beta(h')_k}\left(\mathsf{C}_{\sigma(g)\varsigma}\big((h')^{-1}g'\big)_{j,k} - \mathsf{C}_\varsigma\big((h')^{-1}g'\big)_{j,k}\right)\right| &(3.31)
\end{aligned}$$

$$\leq \sum_{g',h'\in\mathbb{B}_{r,\triangleright}} \left| \sum_{j,k=1}^{d} \beta(g')_j \overline{\beta(h')_k} \left(\mathsf{C}_{\sigma(g)\varsigma}\big((h')^{-1}g'\big)_{j,k} - \mathsf{C}_{\varsigma}\big((h')^{-1}g'\big)_{j,k} \right) \right| \tag{3.32}$$

$$= \sum_{g',h'\in\mathbb{B}_{r,\triangleright}} \left| \sum_{j,k=1}^{d} \beta(g')_j \overline{\beta(h')_k} \big(\langle \rho_\triangleright(g') x_{\sigma(g)\varsigma j}, \rho_\triangleright(h') x_{\sigma(g)\varsigma k} \rangle - \langle \rho_\triangleright(g') x_{\varsigma j}, \rho_\triangleright(h') x_{\varsigma k} \rangle \big) \right| \tag{3.33}$$

$$= \sum_{g',h'\in\mathbb{B}_{r,\triangleright}} \Bigg| \Bigg(\Bigg\langle \rho_\triangleright(g') \sum_{j=1}^{d} \beta(g')_j x_{\sigma(g)\varsigma j}, \rho_\triangleright(h') \sum_{k=1}^{d} \beta(h')_k x_{\sigma(g)\varsigma k} \Bigg\rangle$$

$$- \Bigg\langle \rho_\triangleright(g') \sum_{j=1}^{d} \beta(g')_j x_{\varsigma j}, \rho_\triangleright(h') \sum_{k=1}^{d} \beta(h')_k x_{\varsigma k} \Bigg\rangle \Bigg) \Bigg| \tag{3.34}$$

$$\leq 2\delta \sum_{g',h'\in\mathbb{B}_{r,\triangleright}} \left(\sum_{j=1}^{d} |\beta(g')_j|^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^{d} |\beta(h')_k|^2 \right)^{\frac{1}{2}} \tag{3.35}$$

$$\leq 2K_r\delta \left(\sum_{g'\in\mathbb{B}_{r,\triangleright}} \sum_{j=1}^{d} |\beta(g')_j|^2 \right) \tag{3.36}$$

$$\leq 2K_r R\delta \tag{3.37}$$

Here,

- (3.29) follows from (3.28) using (1.2)
- (3.31) follows from (3.30) since the Δ_r components in the definitions of D_ς and $\mathsf{D}_{\sigma(g)\varsigma}$ cancel,
- (3.33) follows from (3.32) by (3.22)
- (3.35) follows from (3.34) by Proposition 3.3
- and (3.37) follows from (3.36) by (3.27).

□

3.2 Constructing the graph

Let τ_0 be the action of $\mathbb{F}_\triangleleft$ on $\mathrm{Sym}(d)$ given by letting g according to left multiplication by $\sigma(g)$. Let Θ_0 be the directed graph on $\mathrm{Sym}(d)$ with directed edges corresponding to left multiplication by $\tau_0(a)$ and $\tau_0(b)$. By our choice of Λ_n in Segment 3.1.4 we have that the balls of radius $4R$ in Θ_0 are isomorphic to the balls of radius $4R$ in $\mathbb{F}$. In particular, this implies that every cycle in Θ_0 has length at least $4R$. We will now modify τ_0 to obtain a new action of $\mathbb{F}_\triangleleft$ on a finite set.

If V is the vertex set of a graph Θ and $J \subseteq V$ we write $\Theta \restriction J$ for the induced subgraph of Θ on J.

Proposition 3.5. *There exists a finite superset V of* $\mathrm{Sym}(d)$ *with* $|V \setminus \mathrm{Sym}(d)| \leq 9R^{-1}d!$ *and an action* $\tau : \mathbb{F}_\triangleleft \to \mathrm{Sym}(V)$ *with the following properties. Below Θ denotes the directed graph on V corresponding to the actions of $\tau(a)$ and $\tau(b)$.*

(G − 1) *There exists a subset W of V with $|W| \geq (1-20K_rR^{-1})d!$ such that if $w \in W$ then the r-ball around w in Θ can be identified with the r-ball around w in Θ_0.*

(G − 2) *There exists a subset B of V with $|B| \leq R^{-1}d!$ which forms a single $\tau(b)$-cycle. With the exception of the b-cycle corresponding to the vertexes in B, all a and b cycles in Θ have length at most $2(4R+1)$.*

(G − 3) *If $v, w \in \mathrm{Sym}(d)$ then the undirected distance from v to w in Θ_0 is at most $(4R+1)R^2$ times the undirected distance from v to w in $\Theta \upharpoonright (V \setminus B)$.*

(G − 4) *For every $v \in V$ there exists $w \in B$ and a directed path in Θ from w to v with length at most $8(4R+1)^2(10R+1)$.*

(G − 5) *If $g \in \mathbb{B}_r$ and $v, \tau(g)v \in \mathrm{Sym}(d)$ there exists a directed path in Θ from v to $\tau(g)v$ with length at most $256r(4R+1)^2$ which does not pass through an element of B.*

(G − 6) *Each element of $V \setminus \mathrm{Sym}(d)$ is adjacent in Θ to an element of $\mathrm{Sym}(d)$.*

(G − 7) *The $\tau(a)$ and $\tau(b)$ cycles in Θ have length at least 4.*

Proof of Proposition 3.5. We will construct Θ with three stages of modifications to Θ_0. The first stage of modifications will be designed to shorten the length of a-cycles, and its results are summarized in the conditions ($\mathrm{G}_2 - 1$) through ($\mathrm{G}_2 - 5$). The second stage of modifications will be designed to shorten the length of b-cycles, and its results are summarized in the conditions ($\mathrm{G}_3 - 1$) through ($\mathrm{G}_3 - 5$). The third and final stage of modifications will be designed to produce a single anomalously long b-cycle which passes somewhat close to every vertex in the graph. We wish to perform these modifications without dramatically increasing the distance between nearby vertexes. Therefore in each set of modifications we will cut and move certain edges, and then construct 'bypasses' around the cut edges so as to limit the increase in distance resulting from the modification.

We now describe the first set of modifications in order to construct a new action τ_1 of $\mathbb{F}$ on $\mathrm{Sym}(d)$. Let $P_1, \dots, P_n$ be the $\tau_0(a)$-cycles in the graph Θ_0. We have that $|P_m| \geq R$ for all $m \in [n]$. For each $m \in [n]$ let $P_{m,\circ}$ be a set of vertexes in P_m which is maximal among subsets of P_m which are R-separated in P_m. By an R-separated set in P_m we mean a subset S of P_m such that for two vertexes v, w in S the minimal value of n such that $\tau_0^n(a)v = w$ is at least R. The maximality condition guarantees the distance between two consecutive vertexes in $P_{m,\circ}$ is no more than $2R$. Therefore our hypothesis that $|P_m| \geq 4R$ implies that $P_{m,\circ} \geq 2$.

For each $v \in P_{m,\circ}$ cut the $\tau_0(a)$-edge coming out of v. Then place a new $\tau_1(a)$-labelled edge from v to the vertex following the previous element of $P_{m,\circ}$. Call the graph so obtained Θ_1. Since we cut the $\tau_0(a)$-edge coming into the vertex following the previous element of $P_{m,\circ}$, each vertex in the graph Θ_1 has four edges incident to it. These edges are appropriately labelled a, b, a^{-1}, b^{-1} so this graph defines an action of $\mathbb{F}_{\triangleleft}$ on $\mathrm{Sym}(d)$, which we refer to as τ_1.

We now construct bypasses around the edges cut in the previous step. Let D be a set disjoint from $\mathrm{Sym}(d)$ with

$$D = \left| \bigcup_{m \in J} P_{m,\circ} \right|$$

Let D' be a disjoint copy of D. Let

$$f : \bigcup_{m \in J} P_{m,\circ} \to D$$

and let

$$f' : \bigcup_{m \in J} P_{m,\circ} \to D'$$

be bijections. For each $v \in P_{m,\circ}$ cut the $\tau_1(b)$-edge coming out of v in the middle. Place $f(v)$ in the middle of this edge, so that there is a b-labelled edge from v to $f(v)$ and from $f(v)$ to $\tau_1(b)v$. Also, for each $v \in P_{m,\circ}$ cut the $\tau_1(b)$-edge coming out of $\tau_0(a)^2 v$ in the middle. Place $f'(v)$ in the middle of this edge, so there is a b-labelled edge from $\tau_0(a)^2 v$ to $f'(v)$ and from $f'(v)$ to $\tau_1(b)\tau_0^2(a)v$. Place a $\tau_2(a)$-labelled edge from $f(v)$ to $f'(v)$. Our hypothesis that the distance in the $\tau_0(a)$-cycle between consecutive elements of $P_{m,\circ}$ is at least R implies that there are no conflicts arising in this modification. At this stage of the modification the vertexes $f(v)$ are missing an incoming $\tau_2(a)$ edge and the vertexes $f'(v)$ are missing an outgoing $\tau_2(a)$ edge.

Since $|P_{m,\circ}| \geq 2$ we can divide the set $P_{m,\circ}$ into subsets consisting of pairs and at most one triple of vertexes which are consecutive in the $\tau_0(a)$-ordering. If v_1, v_2 is such a pair, place a $\tau_2(a)$ edge from $f'(v_1)$ to $f(v_2)$ and from $f'(v_2)$ to $f(v_1)$. If v_1, v_2, v_3 is such a triple, place a $\tau_2(a)$ edge from $f'(v_1)$ to $f(v_2)$, from $f'(v_2)$ to $f(v_3)$, from $f'(v_3)$ to $f(v_1)$.

Denote the resulting graph by Θ_2. Again, for each element v of $\mathrm{Sym}(d) \cup D$ there exist exactly four edges in Θ_2 incident to v, labelled by a, b, a^{-1}, b^{-1}. Therefore Θ_2 defines an action of $\mathbb{F}_\triangleleft$ on $\mathrm{Sym}(d) \cup D \cup D'$, which we denote τ_2. We make the following claims about Θ_2 and τ_2.

($\mathrm{G}_2 - 1$) In passing from Θ_0 to Θ_2 we have modified the edges incident to at most $5R^{-1}d!$ vertexes.

($\mathrm{G}_2 - 2$) The a-cycles in Θ_2 have length at most $2R$.

($\mathrm{G}_2 - 3$) If $v, w \in \mathrm{Sym}(d)$ then the undirected distance from v to w in Θ_2 is at most 4 times the undirected distance from v to w in Θ_0.

($\mathrm{G}_2 - 4$) If $v, w \in \mathrm{Sym}(d)$ then the undirected distance from v to w in Θ_0 is at most R times the undirected distance from v to w in Θ_2.

($\mathrm{G}_2 - 5$) All $\tau_2(a)$ cycles in Θ_2 have length at least 4. All $\tau_2(b)$ cycles in Θ_2 have length at least $4R$.

We first verify the condition ($\mathrm{G}_2 - 1$). The separation hypothesis on each $P_{m,\circ}$ guarantees that the size of the union of the $P_{m,\circ}$ is at most $R^{-1}d!$. For each vertex $v \in P_{m,\circ}$ we modified the a-edge coming out of v. This gives $2R^{-1}d!$ vertexes whose incident edges were modified. When inserting the elements of D and D' we modified the b edges incident to v and $\tau_0(a)^2 v$. Since we have already accounted for v this gives $3R^{-1}d!$ additional vertexes incident to modified edges. Thus the condition ($\mathrm{G}_2 - 1$) holds.

The condition ($\mathrm{G}_2 - 2$) holds by construction. We now verify the condition ($\mathrm{G}_2 - 3$). Let $v, w \in \mathrm{Sym}(d)$ and let S be an undirected path from v to w in $\mathrm{Sym}(d)$. Suppose the path S passes through a $\tau_0(a)$-edge coming out of some $u \in P_{m,\circ}$. This edge was cut in passing from Θ_0 to Θ_1. However, we can bypass this cut in Θ_2 by using the route

$$u \to \tau_2(b)u = f(u) \to \tau_2(a)\tau_1(b)u = f'(u) \to \tau_1(b)^{-1}f'(u) = \tau_0(a)^2 u$$

Since $\tau_0(a)^2 u = \tau_2(a)\tau_0(a)u$, we see from condition $(\mathrm{G}_2 - 2)$ that bypassing a cut $\tau_0(a)$-edge can be done in 4 undirected steps. Thus the condition $(\mathrm{G}_2 - 3)$ is verified.

The condition $(\mathrm{G}_2 - 4)$ is clear since in passing from Θ_0 to Θ_1 we placed an edge between the beginning and end of a $\tau_0(a)$-path of length R. The first item in the condition $(\mathrm{G}_2 - 5)$ holds by the way we placed $\tau_2(a)$ edges between the vertexes of D and D'. The second item in the condition $(\mathrm{G}_2 - 5)$ holds since in passing from Θ_0 to Θ_2 we only lengthened the b cycles.

We now perform a second set of modifications to obtain a new action τ_3 of $\mathbb{F}$ on $\mathrm{Sym}(d) \cup D \cup D'$. This second set of modifications will be analogous to the first set, with b replacing a. Let $Q_1, \ldots, Q_\ell$ be the $\tau_2(b)$-cycles in the graph Θ_2. Let $I \subseteq [\ell]$ be the set of indexes m such that $|Q_m| \geq R$. If $m \in [\ell]$ let $Q_{m,\circ}$ be a set of vertexes in Q_m which maximal among R-separated subsets of Q_m. The second item in the condition $(\mathrm{G}_2 - 5)$ guarantees that $|Q_{m_\circ}| \geq 2$. For each $v \in Q_{m,\circ}$ cut the $\tau_2(b)$-edge coming out of v. Then place a new $\tau_2(b)$-labelled edge from v to the vertex following the previous element $Q_{m,\circ}$. Note that we cut the $\tau_2(b)$-edge coming into the vertex following the previous element of $Q_{m,\circ}$, so this graph again has the appropriate quartet of labelled edges incident to each vertex.

Now, let E be a set disjoint from $\mathrm{Sym}(d) \cup D \cup D'$ with

$$|E| = \left| \bigcup_{m \in I} Q_{m,\circ} \right|$$

Let E' be a disjoint copy of E and let

$$h : \bigcup_{m \in I} Q_{m,\circ} \to E$$

and

$$h' : \bigcup_{m \in I} Q_{m,\circ} \to E'$$

be bijections. For each $v \in Q_{m,\circ}$ cut the $\tau_2(a)$-edge coming out of v in the middle. Place $h(v)$ in the middle of this cut edge, so that there is an a-labelled edge from v to $h(v)$ and from $h(v)$ to $\tau_2(a)v$. Also cut the $\tau_2(a)$-edge coming out of $\tau_2(b)^2 v$ in the middle. Place $h'(v)$ in the middle of this cut edge, so that there is an a-labelled edge from $\tau_2(b)^2 v$ to $h'(v)$ and from $h'(v)$ to $\tau_2(a)\tau_2(b)^2 v$. The first item in the condition $(\mathrm{G}_2 - 5)$ guarantees that there are no idempotent edges in θ_2, so this construction does not cause any conflicts.

Since $|Q_{m,\circ}| \geq 2$ we can divide the set $Q_{m,\circ}$ into subsets consisting of pairs and at most one triple of vertexes which are consecutive in the $\tau_2(b)$-ordering. If v_1, v_2 is such a pair, place a $\tau_2(b)$ edge from $h'(v_1)$ to $h(v_2)$ and from $h'(v_2)$ to $h(v_1)$. If v_1, v_2, v_3 is such a triple, place a $\tau_2(b)$ edge from $h'(v_1)$ to $h(v_2)$, from $h'(v_2)$ to $h(v_3)$, from $h'(v_3)$ to $h(v_1)$. Call the graph so obtained Θ_3. Again, we have four edges with the appropriate labels attached to every vertex, so Θ_3 defines an action τ_3 of $\mathbb{F}_\triangleleft$ on $\mathrm{Sym}(d) \cup D \cup D' \cup E \cup E'$. We make the following claims about Θ_3 and τ_3.

$(\mathrm{G}_3 - 1)$ In passing from Θ_0 to Θ_3 we have modified the edges incident to at most $10R^{-1}d!$ vertexes.

$(\mathrm{G}_3 - 2)$ Each a-cycles and each b-cycle in Θ_2 has length at most $4R$.

$(\mathrm{G}_3 - 3)$ If $v, w \in \mathrm{Sym}(d)$ then the undirected distance v to w in Θ_3 is at most 16 times the undirected distance from v to w in Θ_0.

($G_3 - 4$) If $v, w \in \mathrm{Sym}(d)$ then the undirected distance from v to w in Θ_0 is at most R^2 times the undirected distance from v to w in Θ_3.

($G_3 - 5$) The $\tau_3(a)$ and $\tau_3(b)$ cycles in Θ_3 have length at least 4.

We first verify the condition ($G_3 - 1$). The separation hypothesis on each $Q_{m,\circ}$ guarantees that the size of the union of the $Q_{m,\circ}$ is at most $R^{-1}d$. For each vertex $v \in Q_{m,\circ}$ we modified the b-edge coming out of v. This gives $2R^{-1}d!$ vertexes whose incident edges were modified. When inserting the elements of E and E' we modified the a edges incident to v to and $\tau_0(b)v$. Since we have already accounted for v this gives $3R^{-1}d!$ additional vertexes incident to modified edges. Combining this with (G_2-1) we obtain the condition (G_3-1).

By construction, in the graph Θ_3 each b-cycle has length at most $2R$. In passing from Θ_2 to Θ_3 we modified a-edges only by sometimes inserting a single vertex in the middle. Thus from the condition ($G_2 - 2$) we obtain the condition ($G_3 - 2$).

Let $v, w \in \mathrm{Sym}(d)$ and let S be an undirected path from v to w in Θ_2. Suppose the path S passes through a $\tau_2(b)$-edge coming out of some $u \in Q_{m,\circ}$. This edge was cut in passing from Θ_2 to Θ_3. However, we can bypass this cut in Θ_3 by using the route

$$u \to \tau_3(a)u = h(u) \to \tau_3(b)\tau_3(a)u = h'(u) \to \tau_3(a)^{-1}h'(u) = \tau_2^2(b)u$$

Since $\tau_2^2(b)u = \tau_3(b)\tau_2(b)u$, we see using condition ($G_3 - 2$) that we can bypass a cut $\tau_2(b)$-edge in 4 undirected steps. Thus condition ($G_3 - 3$) follows from condition ($G_2 - 3$). The condition ($G_3 - 4$) follows from the condition ($G_2 - 4$) since in passing from Θ_2 to Θ_3 we placed edges between the endpoints of $\tau_2(b)$-paths of length at most R.

We now perform a third set of modifications. Choose a subset A of $\mathrm{Sym}(d) \cup D \cup D' \cup E \cup E'$ which is maximal among subsets which are $10R$-separated according to the undirected version of Θ_3, in the sense that any pair of distinct elements of A are at distance at least $10R$ in Θ_3. The hypothesis that the distance between any pair of distinct vertexes v and w in A is at least $10R$ implies that no vertex can be at distance less than $4R$ from both v and w. Thus the balls of radius $4R$ in Θ_3 around the elements of A are pairwise disjoint. Recalling that the balls of radius $4R$ in Θ_0 are isomorphic with balls of radius $4R$ in $\mathbb{F}$, it is clear that in passing from Θ_0 to Θ_3 we maintained the existence of a path of length R in the ball of radius $4R$ around any given point. Thus the balls of radius $4R$ around the elements of A have size at least R and so we have

$$|A| \leq R^{-1}(d! + 2|D| + 2|E|) \leq 5R^{-1}d!$$

Furthermore, the hypothesis that A is maximal among $10R$-separated sets implies that every element of $\mathrm{Sym}(d) \cup D \cup D' \cup E \cup E'$ has distance at most $10R$ from an element of A.

Let B be a set disjoint from $\mathrm{Sym}(d) \cup D \cup D' \cup E \cup E'$ with $|B| = |A|$. Let $\ell : A \to B$ be a bijection. For each element v of A, cut the edge between $\tau_3(a)^{-1}v$ and v in the middle. Insert $\ell(v)$ in between the two pieces of the cut edge, so that there is an a-labelled edge from $\tau_3(a)^{-1}v$ to $\ell(v)$ and from $\ell(v)$ to v. Also let $\{v_1, \ldots, v_s\}$ be an enumeration of B and place a b-labelled edge from v_m to v_{m+1} for $m \in [s]$, where the indexes m are taken modulo s. Write $V = \mathrm{Sym}(d) \cup D \cup D' \cup E \cup E' \cup B$. Call the graph on V so obtained Θ_4. Let τ_4 be the action of $\mathbb{F}_\triangleleft$ on V associated to Θ_4.

Now, for each vertex $v \in A$ let C_v be the $\tau_4(b)$-cycle containing v and let C'_v be the $\tau_4(b)$-cycle containing $\tau_3^{-1}(a)v = \tau_4^{-2}(a)v$. The condition $(\mathrm{G}_3 - 5)$ guarantees that there exist vertexes in these cycles which are not elements of A. Choose a vertex $w(v) \in C_v$ and $w'(v) \in C'_v$ which are not elements of A. Cut the $\tau_4(b)$-edge going out of $w(v)$ and going out of $w'(v)$. Place an b edge from $w(v)$ to $\tau_4(b)w'(v)$ and from $w'(v)$ to $\tau_4(b)w(v)$. Thus C_v and C'_v become a single b-cycle. Note that since A was $10R$-separated in the undirected version of Θ_4, the cycles C_v and C'_v contain no elements of B. Let Θ be the resulting graph and let τ be the corresponding action. We now verify that the conditions $(\mathrm{G}-1)$ through $(\mathrm{G}-6)$ holds for Θ.

In passing from Θ_3 to Θ_4 we modified the edges incident to at most $2|A| \leq 10R^{-1}d!$ vertexes. In combination with the condition $(\mathrm{G}_3 - 2)$ we see that in passing from Θ_0 to Θ_4 we modified the edges incident to at most $20R^{-1}d!$ vertexes. We can take W to be the complement in $\mathrm{Sym}(d)$ of the balls of radius r around vertexes incident to modified edges. Thus the condition $(\mathrm{G} - 1)$ is verified.

By construction the a and b cycles in Θ_3 have length at most $4R$. In passing from Θ_3 to Θ_4 we added at most 1 edge to an a-cycle and no edges to a b-cycle. In passing from Θ_4 to Θ we at most doubled the length of a cycle. Thus condition $(\mathrm{G} - 2)$ is verified.

We have that $\Theta_4 \restriction (V \setminus B)$ is a subgraph of Θ_3. Therefore if $v, w \in \mathrm{Sym}(d)$ we see from the condition $(\mathrm{G}_3 - 4)$ that the undirected distance from v to w in Θ_0 is at most R^2 times the undirected distance from v to w in $\Theta_4 \restriction (V \setminus B)$. In passing from Θ_4 to Θ we connected pairs of $\tau_3(b)$-cycles which were joined by a $\tau_3(a)$-edge. Using the condition $(\mathrm{G}_3 - 2)$ we see that the undirected distance from v to w in $\Theta_4 \restriction (V \setminus B)$ is at most $4R+1$ times the undirected distance from v to w in $\Theta \restriction (V \setminus B)$. This implies the condition $(\mathrm{G}-3)$.

The maximality of A guarantees that for every $v \in V$ there exists $u \in A$ and an undirected path from u to v in Θ_3 with length at most $10R$. Therefore if we set $w = \tau(a)^{-1}(u) \in B$ then there exists an undirected path from w to v in Θ_3 with length at most $10R+1$. In passing from Θ_3 to Θ_4 we may have placed vertexes in the middle of $\tau_3(a)$-edges, thereby at most doubling the length of this path. Thus we obtain an undirected path from w to v of length at most $2(10R+1)$ in Θ_4. In passing from Θ_4 to Θ we combined certain pairs of cycles, and therefore we multiplied the length of any path by at most the length of a cycle in Θ_4. Condition $(\mathrm{G} - 2)$ guarantees the length of such a cycle is at most $4R + 1$, so we obtain an undirected path from w to v in Θ of length at most $2(4R + 1)(10R + 1)$. Using condition $(\mathrm{G} - 2)$ we obtain condition $(\mathrm{G} - 4)$.

Let $g \in \mathbb{B}_{r,\lhd}$ and let $v \in \mathrm{Sym}(d)$. The condition $(\mathrm{G}_3 - 3)$ guarantees there exists an undirected path from v to $\tau_0(g)v$ in Θ_3 with length at most $16r$ which might pass through an element of B. In passing from Θ_3 to Θ_4 we may have placed a vertex in the middle of a $\tau_3(a)$-edge on the path. Thus there is an undirected path from v to $\tau_0(g)v$ in Θ_4 of length at most $32r$ which might pass through an element of B. Using condition $(\mathrm{G} - 2)$ we see that there exists a directed path from v to $\tau_0(g)v$ in Θ_4 of length at most $64r(4R + 1)$ which might pass through an element of B.

Suppose u is an element of B on this directed path. The modification made in passing from Θ_4 to Θ ensures that by following the $\tau_4(b)$-cycles we can travel from $\tau_4(a)^{-1}u$ to $\tau_4(a)u$ without meeting an element of B. Using condition $(\mathrm{G}-2)$ we see that this bypass has length at most $4(4R+1)$. Thus condition $(\mathrm{G}-5)$ is verified.

In the three steps of the modification, we inserted vertexes into a-edges twice and into b-edges once. This observation implies the condition $(\mathrm{G}-6)$. The condition $(\mathrm{G}-7)$ is immediate from (G_3-5). This completes the proof of Proposition 3.5. $\square$

3.3 Verifying the properties of the graph

3.3.1 Estimating transport operators in the modified graph

Let $\imath$ be the identity permutation in $\mathrm{Sym}(d)$. Let V and τ be as in Proposition 3.5. For $v \in V \setminus \mathrm{Sym}(d)$ let $\widehat{v}$ be the element of $\mathrm{Sym}(d)$ guaranteed by condition (G − 6) in Proposition 3.5, so that $\widehat{v}$ is adjacent to v. Define positive definite functions $(\mathsf{E}_v)_{v\in V}$ by setting

$$\mathsf{E}_v = \begin{cases} \mathsf{D}_v & \text{if } v \in \mathrm{Sym}(d) \\ \mathsf{D}_\imath & \text{if } v \in B \\ \mathsf{D}_{\widehat{v}} & \text{if } v \in V \setminus (\mathrm{Sym}(d) \cup B) \end{cases} \tag{3.38}$$

Using the hypothesis that $\mathrm{encost}_{r,\epsilon}(\Theta) \leq M$ we can find positive definite functions $(\widehat{\mathsf{E}}_v)_{v\in V}$ defined on all of $\mathbb{F}_\triangleright$ such that

$$||\mathsf{E}_v - (\widehat{\mathsf{E}}_v \restriction \mathbb{B}_r)||_1 \leq \epsilon \tag{3.39}$$

for all $v \in V$ and such that

$$\mathfrak{e}\big(\widehat{\mathsf{E}}_v, \widehat{\mathsf{E}}_{\tau(a)v}\big) - 1 \leq M\big(\mathfrak{e}(\mathsf{E}_v, \mathsf{E}_{\tau(a)v}) - 1\big) \tag{3.40}$$

and

$$\mathfrak{e}\big(\widehat{\mathsf{E}}_v, \widehat{\mathsf{E}}_{\tau(b)v}\big) - 1 \leq M\big(\mathfrak{e}(\mathsf{E}_v, \mathsf{E}_{\tau(b)v}) - 1\big) \tag{3.41}$$

for all $v \in V$. We regard each $\widehat{\mathsf{E}}_v$ as a function from $\mathbb{F}_\triangleright$ to $\mathrm{Mat}_{d\times d}(\mathbb{C})$. From Proposition 3.4 and (3.40) and (3.41) we obtain

$$\mathfrak{e}\big(\widehat{\mathsf{E}}_v, \widehat{\mathsf{E}}_{\tau(a)v}\big) \leq 1 + 2MK_r\delta \tag{3.42}$$

and

$$\mathfrak{e}\big(\widehat{\mathsf{E}}_v, \widehat{\mathsf{E}}_{\tau(b)v}\big) \leq 1 + 2MK_r\delta \tag{3.43}$$

for all $v \in V$.

Suppose $g \in \mathbb{B}_{r,\triangleleft}$ and $v, \tau(g)v \in \mathrm{Sym}(d)$. Let

$$P = (v_1 = v, v_2, \ldots, v_{n-1}, v_n = \tau(g)v)$$

be the directed path guaranteed by the condition (G − 5) in Proposition 3.5. Using the condition (G − 3) in Proposition 3.5 we see that if $v_m, v_{m+1} \in \mathrm{Sym}(d)$ then v_m was at undirected distance at most $(4R + 1)R^2$ from v_{m+1} in Θ_0. Moreover if $v_m, v_{m+1} \in V \setminus (\mathrm{Sym}(d) \cup B)$ we see that the undirected distance from $\widehat{v_m}$ to $\widehat{v_{m+1}}$ in $\Theta \restriction (V \setminus B)$ is at most 3, so again using the condition (G − 3) in Proposition 3.5 we see that $\widehat{v_m}$ was at undirected distance at most $3(4R+1)R^2$ from $\widehat{v_{m+1}}$ in Θ_0. A similar argument applies when $v_m \in \mathrm{Sym}(d)$ and $v_{m+1} \in V \setminus (\mathrm{Sym}(d) \cup B)$, and when $v \in V \setminus (\mathrm{Sym}(d) \cup B)$ and $v_m \in \mathrm{Sym}(d)$. Since the edges in Θ_0 correspond to multiplication by $\sigma(a)^{\pm 1}$ and $\sigma(b)^{\pm 1}$, Proposition 3.4 implies that

$$\mathfrak{e}(\mathsf{D}_{v_m}, \mathsf{D}_{v_{m+1}}) \leq (1 + 2MK_rR\delta)^{3(4R+1)R^2} \leq \exp(6M(4R + 1)K_rR^3\delta) \tag{3.44}$$

Thus for $g \in \mathbb{B}_{r,\triangleleft}$ we have

$$\mathfrak{e}(\widehat{\mathsf{E}}_v, \widehat{\mathsf{E}}_{\tau(g)v}) \le \prod_{m=1}^{n-1} \mathfrak{e}(\widehat{\mathsf{E}}_{v_m}, \widehat{\mathsf{E}}_{v_{m+1}}) \tag{3.45}$$

$$= \prod_{m=1}^{n-1} \mathfrak{e}(\mathsf{D}_{v_m}, \mathsf{D}_{v_{m+1}}) \tag{3.46}$$

$$\le \prod_{m=1}^{n-1} \exp(6(4R+1)K_r R^3 \delta) \tag{3.47}$$

$$\le \exp(1536 r M K_r (4R+1)^3 R^3 \delta) \tag{3.48}$$

$$= e^{s\delta} \tag{3.49}$$

Here,

- (3.46) follows from (3.45) by (3.38) since P does not pass through an element of B,
- (3.47) follows from (3.46) by (3.44),
- and (3.48) follows from (3.47) since the condition $(\mathrm{G}-5)$ in Proposition 3.5 guarantees $m \le 256r(4R+1)^2$.
- and (3.49) follows from (3.48) by (3.3)

Furthermore, from the conditions $(\mathrm{G}-3)$ and $(\mathrm{G}-4)$ in Proposition 3.5 we see that for every $v \in V$ there exists $w \in B$ such that there is a path P in Θ from w to v of length at most $8R^2(4R+1)^3(10R+1)$ such that each edge in P corresponds to an undirected edge in Θ_0. Thus by simplifying the upper bound in Proposition 3.4 to 2 we obtain

$$\mathfrak{e}(\widehat{\mathsf{E}}_w, \widehat{\mathsf{E}}_v) \le (2M)^{8R^2(4R+1)^3(10R+1)} = \exp\big(8R^2(4R+1)^3(10R+1)\log(2M)\big) \tag{3.50}$$

Define

$$\widehat{\mathsf{F}}_v = \begin{cases} \widehat{\mathsf{E}}_v & \text{if } v \in \mathrm{Sym}(d) \\ \widehat{\mathsf{E}}_{\widehat{v}} & \text{if } v \in V \setminus \mathrm{Sym}(d) \end{cases} \tag{3.51}$$

From (3.49) and (3.42) we see that for all $v \in V$ and all $g \in \mathbb{B}_{r,\triangleleft}$ we have

$$\mathfrak{e}(\widehat{\mathsf{F}}_v, \widehat{\mathsf{F}}_{\tau(g)v}) \le e^{s\delta} \tag{3.52}$$

Fix $w_\circ \in B$ and write $\widehat{\mathsf{E}}_\circ$ for $\widehat{\mathsf{E}}_{w_\circ}$. We have $\mathsf{E}_w = \mathsf{E}_\circ$ for all $w \in B$ and so from (3.42) we obtain

$$\mathfrak{e}(\widehat{\mathsf{E}}_\circ, \widehat{\mathsf{E}}_w) \le \exp\big(8R^2(4R+1)^3(10R+1)\log(2M)\big) \tag{3.53}$$

for all $w \in B$.

3.3.2 Constructing a representation through permutations

Again let V be as in Proposition 3.5. For $v \in V$ we now define $\mathsf{K}_v = R^{-1}\widehat{\mathsf{E}}_\circ + (1-R^{-1})\widehat{\mathsf{F}}_v$. Write $\mathcal{Y}_v$ for $\mathcal{X}(\mathsf{K}_v)$ and let $\zeta_{\triangleright,v} : \mathbb{F}_\triangleright \to \mathrm{U}(\mathcal{Y}_v)$ be the associated representation of K_v. Define $\mathcal{Y} = \bigoplus_{v\in V} \mathcal{Y}_v$ and $\zeta_\triangleright = \bigoplus_{v\in V} \zeta_{\triangleright,v}$. Define a representation $\theta : \mathbb{F}_\triangleleft \to \mathrm{GL}(\mathcal{Y})$ by setting

$$\theta(g) = \bigoplus_{v\in V} t[\mathsf{K}_v, \mathsf{K}_{\tau(g)v}]$$

Note that θ factors through the finite group $\Gamma = \tau(\mathbb{F}_\triangleleft)$. Moreover, we have

$$t[\mathsf{K}_v, \mathsf{K}_{\tau(g)v}]\zeta_{\triangleright.v} = \zeta_{\triangleright,\tau(g)v} t[\mathsf{K}_v, \mathsf{K}_{\tau(g)v}]$$

for all $v \in V$ and $g \in \mathbb{F}_\triangleleft$. Therefore θ commutes with $\zeta_\triangleright$ so that $\theta \times \zeta_\triangleright$ is a half-finite linear representation of G.

From (3.52) we see that

$$||\theta(g)||_{\mathrm{op}} \leq e^{s_\delta} \tag{3.54}$$

for all $g \in \mathbb{B}_{r,\triangleleft}$. Now, let $g \in \mathbb{F}_\triangleleft$ be arbitrary. Let $\beta : \mathbb{F}_\triangleright \to \mathbb{C}^d$ be such that

$$\sum_{g'\in\mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\mathsf{K}_v}(g')_j$$

is a unit vector in $\mathcal{Y}_v$. Thus we have

$$\begin{aligned} 1 &= \left(1 - \frac{1}{R}\right) \left|\left| \sum_{g'\in\mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\widehat{\mathsf{F}}_v}(g')_j \right|\right|^2 + \frac{1}{R} \left|\left| \sum_{g'\in\mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\widehat{\mathsf{E}}_\circ}(g')_j \right|\right|^2 \\ &\geq \frac{1}{R} \left|\left| \sum_{g'\in\mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\widehat{\mathsf{E}}_\circ}(g')_j \right|\right|^2 \end{aligned} \tag{3.55}$$

We have

$$\begin{aligned} &\left|\left| t[\mathsf{K}_v, \mathsf{K}_{\tau(g)v}] \sum_{g'\in\mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\mathsf{K}_v}(g')_j \right|\right|^2 \\ &\qquad = \left|\left| \sum_{g'\in\mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\mathsf{K}_{\tau(g)v}}(g')_j \right|\right|^2 \\ &\qquad = \left(1 - \frac{1}{R}\right) \left|\left| \sum_{g'\in\mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\widehat{\mathsf{F}}_{\tau(g)v}}(g')_j \right|\right|^2 + \frac{1}{R} \left|\left| \sum_{g'\in\mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\widehat{\mathsf{E}}_\circ}(g')_j \right|\right|^2 \\ &\qquad = \left(1 - \frac{1}{R}\right) \left|\left| \sum_{g'\in\mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\widehat{\mathsf{E}}_\circ}(g')_j \right|\right|^2 + \frac{1}{R} \left|\left| t[\widehat{\mathsf{E}}_\circ, \widehat{\mathsf{F}}_{\tau(g)v}] \sum_{g'\in\mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\widehat{\mathsf{E}}_\circ}(g')_j \right|\right|^2 \end{aligned}$$

$$\leq \left(1 - \frac{1}{R} + \frac{\mathfrak{e}(\widehat{\mathsf{E}}_\circ, \widehat{\mathsf{F}}_{\tau(g)v})}{R}\right) \left|\left| \sum_{g' \in \mathbb{F}_\triangleright} \sum_{j=1}^{d} \beta(g')_j \Phi_{\widehat{\mathsf{E}}_\circ}(g')_j \right|\right|^2 \tag{3.56}$$

$$\leq (1 + \mathfrak{e}(\widehat{\mathsf{E}}_\circ, \widehat{\mathsf{F}}_{\tau(g)v}))R \tag{3.57}$$

$$\leq \big(1 + \exp\big(8R^2(4R+1)^3(10R+1)\log(2M)\big)\big)R \tag{3.58}$$

Here, (3.57) follows from (3.56) by (3.55) and (3.58) follows from (3.57) by (3.53). From (3.2) and (3.58) we obtain

$$||\theta(g)||_{\mathrm{op}} \leq L_{r,R} \tag{3.59}$$

for all $g \in \mathbb{F}_\triangleleft$.

3.4 Repairing the representation to be unitary

3.4.1 Conjugation by an average

In Segments 3.4.1 and 3.4.2 we regard θ as a representation of the finite group Γ. Define a positive operator $q \in \mathrm{B}(\mathfrak{Y})$ by

$$q = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} \theta(\gamma)^* \theta(\gamma).$$

By applying (3.59) to g^{-1} we see that each $\theta(\gamma)$ is invertible. Hence each operator $\theta(\gamma)^*\theta(\gamma)$ is strictly positive and so q is invertible. Define a representation $\zeta_\triangleleft$ of Γ on $\mathfrak{Y}$ by setting $\zeta_\triangleleft(\gamma) = q^{\frac{1}{2}}\theta(\gamma)q^{-\frac{1}{2}}$. For all $\gamma \in \Gamma$ and all $g' \in \mathbb{F}_\triangleright$ we have that $\zeta_\triangleright(g')$ commutes with each $\theta(\gamma)$. Since $\zeta_\triangleright(g')$ is unitary, this implies that $\zeta_\triangleright(g')$ commutes with $\theta(\gamma)^*$ and hence $\zeta_\triangleright(g')$ commutes with q. Therefore $\zeta_\triangleright$ commutes with $\zeta_\triangleleft$ and so if we set $\zeta = \zeta_\triangleleft \times \zeta_\triangleright$ then ζ is a half finite linear representation of G.

We claim that ζ is in fact unitary. Write I for the identity operator on $\mathfrak{Y}$. For $\gamma \in \Gamma$ we have

$$\begin{aligned}
\zeta_\triangleleft(\gamma)^*\zeta_\triangleleft(\gamma) &= \big(q^{\frac{1}{2}}\theta(\gamma)q^{-\frac{1}{2}}\big)^*\big(q^{\frac{1}{2}}\theta(\gamma)q^{-\frac{1}{2}}\big) \\
&= q^{-\frac{1}{2}}\theta(\gamma)^* q \theta(\gamma) q^{-\frac{1}{2}} \\
&= q^{-\frac{1}{2}}\theta(\gamma)^* \left(\frac{1}{|\Gamma|}\sum_{\nu \in \Gamma}\theta(\nu)^*\theta(\nu)\right)\theta(\gamma)q^{-\frac{1}{2}} \\
&= q^{-\frac{1}{2}} \left(\frac{1}{|\Gamma|}\sum_{\nu \in \Gamma}\theta(\gamma)^*\theta(\nu)^*\theta(\nu)\theta(\gamma)\right) q^{-\frac{1}{2}} \\
&= q^{-\frac{1}{2}} \left(\frac{1}{|\Gamma|}\sum_{\nu \in \Gamma}\theta(\nu\gamma)^*\theta(\nu\gamma)\right) q^{-\frac{1}{2}} \\
&= q^{-\frac{1}{2}} \left(\frac{1}{|\Gamma|}\sum_{\nu \in \Gamma}\theta(\nu)^*\theta(\nu)\right) q^{-\frac{1}{2}} \\
&= I
\end{aligned} \tag{3.60}$$

so that $\zeta_\triangleleft(\gamma)$ is unitary and therefore ζ is a unitary representation.

3.4.2 Bounding the spectrum of the average

Proposition 3.6. *We have* $\mathrm{spec}(q) \subseteq [L_{r,R}^{-2}, L_{r,R}^2]$.

Proof of Proposition 3.6. Using (3.59) we see that for any unit vector $y \in \mathcal{Y}$ we have

$$\begin{aligned} \langle qy, y\rangle &= \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\big\langle \theta(\gamma)^*\theta(\gamma)y, y\big\rangle \\ &= \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}||\theta(\gamma)y||^2 \\ &\le L_{r,R}^2 \end{aligned} \tag{3.61}$$

By applying (3.59) to g^{-1} we see that

$$\inf\big\{||\theta(\gamma)y||^2 : y \in \mathcal{Y} \text{ is a unit vector}\big\} \ge \frac{1}{L_{r,R}^2}$$

and so

$$\inf\big\{\langle qy, y\rangle : y \in \mathcal{Y} \text{ is a unit vector}\big\} \ge \frac{1}{L_{r,R}^2}. \tag{3.62}$$

Now suppose $\lambda \in \mathrm{spec}(q)$. Since q is self-adjoint, there exists a sequence $(y_n)_{n=1}^\infty$ of unit vectors in $\mathcal{Y}$ such that $\lim_{n\to\infty} ||(q-\lambda I)y_n|| = 0$. This implies that $\lim_{n\to\infty}\langle (q-\lambda I)y_n, y_n\rangle = 0$ and so $\lim_{n\to\infty}\langle qy_n, y_n\rangle = \lambda$. Thus from (3.61) and (3.62) we have $L_{r,R}^{-2} \le \lambda \le L_{r,R}^2$. □

3.4.3 Estimating the distance to the repaired representation

Proposition 3.7. *Suppose* $g \in \mathbb{B}_{r,\triangleleft}$. *Then* $||\zeta_\triangleleft(g) - \theta(g)||_{\mathrm{op}} \le R^{-1}$.

Proof of Proposition 3.7. Fix $g \in \mathbb{B}_{r,\triangleleft}$. By applying (3.54) to g and g^{-1} we see

$$e^{-s_\delta} I \le \theta(g)^*\theta(g) \le e^{s_\delta} I$$

Since $\theta(g)^*\theta(g)$ is unitarily conjugate to $\theta(g)\theta(g)^*$ we obtain

$$e^{-s_\delta} I \le \theta(g)\theta(g)^* \le e^{s_\delta} I$$

so that

$$||\theta(g)\theta(g)^* - I||_{\mathrm{op}} \le e^{s_\delta} - 1 \tag{3.63}$$

Since $q^{-\frac{1}{2}}\theta(g)^* q\theta(g) q^{-\frac{1}{2}} = I$ we have $\theta(g)^* q\theta(g) = q$. Therefore

$$\begin{aligned} ||q\theta(g) - \theta(g)q||_{\mathrm{op}} &= ||q\theta(g) - \theta(g)\theta(g)^* q\theta(g)||_{\mathrm{op}} \\ &\le ||I - \theta(g)\theta(g)^*||_{\mathrm{op}}||q||_{\mathrm{op}}||\theta(g)||_{\mathrm{op}} && (3.64) \\ &\le 2||I - \theta(g)\theta(g)^*||_{\mathrm{op}}||q||_{\mathrm{op}} && (3.65) \\ &\le 2L_{r,R}^2||I - \theta(g)\theta(g)^*||_{\mathrm{op}} && (3.66) \\ &\le 2L_{r,R}^2(e^{s_\delta} - 1) && (3.67) \end{aligned}$$

Here,

- (3.65) follows from (3.64) by (3.54) since $e^{s_\delta} \leq 2$,
- (3.66) follows from (3.65) by Proposition 3.6 since q is self-adjoint,
- and (3.67) follows from (3.66) by (3.63).

Let $z \in \mathbb{C} \setminus \mathrm{spec}(q)$. We compute

$$\begin{aligned}
&||(q - zI)^{-1}\theta(g) - \theta(g)(q - zI)^{-1}||_{\mathrm{op}} \\
&= ||(q - zI)^{-1}\theta(g) - (q - zI)^{-1}(q - zI)\theta(g)(q - zI)^{-1}||_{\mathrm{op}} \\
&\leq ||(q - zI)^{-1}||_{\mathrm{op}}||\theta(g) - (q - zI)\theta(g)(q - zI)^{-1}||_{\mathrm{op}} \\
&= \frac{1}{\mathrm{dist}(z, \mathrm{spec}(q))}||\theta(g) - (q - zI)\theta(g)(q - zI)^{-1}||_{\mathrm{op}} \\
&= \frac{1}{\mathrm{dist}(z, \mathrm{spec}(q))}\Big|\Big|\theta(g) - q\theta(g)(q - zI)^{-1} + z\theta(g)(q - zI)^{-1}\Big|\Big|_{\mathrm{op}} \\
&= \frac{1}{\mathrm{dist}(z, \mathrm{spec}(q))}\Big|\Big|\theta(g) - \theta(g)q(q - zI)^{-1} + z\theta(g)(q - zI)^{-1} \\
&\qquad\qquad + \theta(g)q(q - zI)^{-1} - q\theta(g)(q - zI)^{-1}\Big|\Big|_{\mathrm{op}} \\
&= \frac{1}{\mathrm{dist}(z, \mathrm{spec}(q))}\Big|\Big|\theta(g) - \theta(g)(q - zI)(q - zI)^{-1} + \theta(g)q(q - zI)^{-1} - q\theta(g)(q - zI)^{-1}\Big|\Big|_{\mathrm{op}} \\
&= \frac{1}{\mathrm{dist}(z, \mathrm{spec}(q))}||\theta(g)q(q - zI)^{-1} - q\theta(g)(q - zI)^{-1}||_{\mathrm{op}} \\
&\leq \frac{1}{\mathrm{dist}(z, \mathrm{spec}(q))}||\theta(g)q - q\theta(g)||_{\mathrm{op}}||(q - zI)^{-1}||_{\mathrm{op}} \\
&= \frac{1}{\mathrm{dist}(z, \mathrm{spec}(q))^2}||\theta(g)q - q\theta(g)||_{\mathrm{op}} && (3.68) \\
&\leq \frac{2L_{r,R}^2(e^{s_\delta} - 1)}{\mathrm{dist}(z, \mathrm{spec}(q))^2} && (3.69)
\end{aligned}$$

Here, (3.69) follows from (3.68) by (3.67). Now, let $c : [0, 1] \to \mathbb{C}$ be a simple closed contour with the following properties.

(i) We have $\mathrm{Re}(c(x)) > 0$ for all $x \in [0, 1]$.

(ii) The interval $[L_{r,R}^{-2}, L_{r,R}^2]$ is enclosed by c.

(iii) We have $\mathrm{dist}(c(x), [L_{r,R}^{-2}, L_{r,R}^2]) \geq \frac{1}{2}L_{r,R}^{-2}$ for all $x \in [0, 1]$.

(iv) We have $\sup\{|c(x)| : x \in [0, 1]\} \leq 2L_{r,R}^2$

(v) We have $\ell(c) \leq 10L_{r,R}^2$ where $\ell(c)$ denotes the length of c.

By Clause (i) we can consistently define a square root function on the image of c. Proposition 3.6 together with Clause (ii) in the definition of c implies that c encloses $\mathrm{spec}(q)$. Therefore we can use the holomorphic functional calculus to make the following computation.

$$
\begin{aligned}
&||\theta(g)q^{\frac{1}{2}} - q^{\frac{1}{2}}\theta(g)||_{\mathrm{op}} \\
&\quad = \frac{1}{2\pi}\left|\left|\theta(g)\left(\int_0^1 c(x)^{\frac{1}{2}}(c(x)I-q)^{-1}\,\mathrm{d}x\right) - \left(\int_0^1 c(x)^{\frac{1}{2}}(c(x)I-q)^{-1}\,\mathrm{d}x\right)\theta(g)\right|\right|_{\mathrm{op}} \\
&\quad = \frac{1}{2\pi}\left|\left|\int_0^1 c(x)^{\frac{1}{2}}\Big(\theta(g)(c(x)I-q)^{-1} - (c(x)I-q)^{-1}\theta(g)\Big)\,\mathrm{d}x\right|\right|_{\mathrm{op}} \\
&\quad \le \frac{\ell(c)}{2\pi}\sup_{0\le x\le 1}\left(|c(x)|^{\frac{1}{2}}\left|\left|\theta(g)(c(x)I-q)^{-1} - (c(x)I-q)^{-1}\theta(g)\right|\right|_{\mathrm{op}}\right) && (3.70)\\
&\quad \le 10L_{r,R}^2\sup_{0\le x\le 1}\left(|c(x)|^{\frac{1}{2}}\left|\left|\theta(g)(c(x)I-q)^{-1} - (c(x)I-q)^{-1}\theta(g)\right|\right|_{\mathrm{op}}\right) && (3.71)\\
&\quad \le 20L_{r,R}^3\sup_{0\le x\le 1}\left|\left|\theta(g)(c(x)I-q)^{-1} - (c(x)I-q)^{-1}\theta(g)\right|\right|_{\mathrm{op}} && (3.72)\\
&\quad \le \frac{40L_{r,R}^5(e^{s_\delta}-1)}{\mathrm{dist}(z,\mathrm{spec}(q))^2} && (3.73)\\
&\quad \le 160L_{r,R}^9(e^{s_\delta}-1) && (3.74)
\end{aligned}
$$

Here,

- (3.71) follows from (3.70) by Clause (v) in the definition of c,
- (3.72) follows from (3.71) by Clause (iv) in the definition of c,
- (3.73) follows from (3.72) by (3.69),
- and (3.74) follows from (3.73) by Clause (iii) in the definition of c.

Now, since $\mathrm{spec}(q) \subseteq [L_{r,R}^{-2}, L_{r,R}^2]$, the spectral mapping theorem implies that $\mathrm{spec}(q^{-\frac{1}{2}}) \subseteq [L_{r,R}^{-1}, L_{r,R}]$. Since $q^{-\frac{1}{2}}$ is self-adjoint, this implies $||q^{-\frac{1}{2}}||_{\mathrm{op}} \le L_{r,R}$. Therefore

$$
\begin{aligned}
||\zeta_{\triangleleft}(g) - \theta(g)||_{\mathrm{op}} &= ||q^{\frac{1}{2}}\theta(g)q^{-\frac{1}{2}} - \theta(g)||_{\mathrm{op}} \\
&= ||q^{\frac{1}{2}}\theta(g)q^{-\frac{1}{2}} - \theta(g)q^{\frac{1}{2}}q^{-\frac{1}{2}}||_{\mathrm{op}} \\
&\le ||q^{\frac{1}{2}}\theta(g) - \theta(g)q^{\frac{1}{2}}||_{\mathrm{op}}||q^{-\frac{1}{2}}||_{\mathrm{op}} \\
&\le 160L_{r,R}^{10}(e^{s_\delta}-1)
\end{aligned}
$$

Therefore Proposition 3.7 follows from (3.4) □

3.5 Finding a witness vector

Define a vector $y \in \mathcal{Y}$ by setting

$$
y = \frac{1}{|W|}\bigoplus_{\varsigma\in W}\sum_{j=1}^{d}\alpha_{\varsigma j}\Phi_{\mathsf{K}_\varsigma}(e)_j
$$

where $W \subseteq \mathrm{Sym}(d)$ is the set from the condition $(\mathrm{G}-1)$. Since each K_ς is normalized we have from (3.7) that y is a unit vector. Let $g \in \mathbb{B}_{r,\triangleleft}$ and let $g' \in \mathbb{B}_{r,\triangleright}$. From Proposition 3.7 we have

$$\langle \zeta(g,g')y,y\rangle = \langle \zeta_\triangleleft(g)\zeta_\triangleright(g')y,y\rangle \approx[R^{-1}] \langle \theta(g)\zeta_\triangleright(g')y,y\rangle \tag{3.75}$$

We have

$$\begin{aligned}\langle \theta(g)\zeta_\triangleright(g')y,y\rangle &= \frac{1}{R|W|}\left\langle \theta(g)\zeta_\triangleright(g')\bigoplus_{\varsigma\in W}\sum_{j=1}^{d}\alpha_{\varsigma j}\Phi_{\widehat{\mathsf{E}}_\circ}(e)_j, \bigoplus_{\varsigma\in W}\sum_{k=1}^{d}\alpha_{\varsigma k}\Phi_{\widehat{\mathsf{E}}_\circ}(e)_k\right\rangle \\ &\quad + \frac{1-R^{-1}}{|W|}\left\langle \theta(g)\zeta_\triangleright(g')\bigoplus_{\varsigma\in W}\sum_{j=1}^{d}\alpha_{\varsigma j}\Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_j, \bigoplus_{\varsigma\in W}\sum_{k=1}^{d}\alpha_{\varsigma k}\Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_k\right\rangle\end{aligned} \tag{3.76}$$

We have

$$\frac{1}{R|W|}\left|\left\langle \theta(g)\zeta_\triangleright(g')\bigoplus_{\varsigma\in W}\sum_{j=1}^{d}\alpha_{\varsigma j}\Phi_{\widehat{\mathsf{E}}_\circ}(e)_j, \bigoplus_{\varsigma\in W}\sum_{k=1}^{d}\alpha_{\varsigma k}\Phi_{\widehat{\mathsf{E}}_\circ}(e)_k\right\rangle\right| \leq \frac{1}{R} \tag{3.77}$$

From (3.75), (3.76) and (3.77) we have

$$\langle \zeta(g,g')y,y\rangle \approx[2R^{-1}]\frac{1}{|W|}\left\langle \theta(g)\zeta_\triangleright(g')\bigoplus_{\varsigma\in W}\sum_{j=1}^{d}\alpha_{\varsigma j}\Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_j, \bigoplus_{\varsigma\in W}\sum_{k=1}^{d}\alpha_{\varsigma k}\Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_k\right\rangle \tag{3.78}$$

By construction we have

$$\zeta_\triangleright(g')\Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_j = \Phi_{\widehat{\mathsf{F}}_\varsigma}(g')_j \tag{3.79}$$

We have

$$\begin{aligned}&\frac{1}{d!}\left\langle \theta(g)\bigoplus_{\varsigma\in W}\sum_{j=1}^{d}\alpha_{\varsigma j}\Phi_{\widehat{\mathsf{F}}_\varsigma}(g')_j, \bigoplus_{\varsigma\in W}\sum_{k=1}^{d}\alpha_{\varsigma k}\Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_k\right\rangle \\ &\qquad = \frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\left\langle \sum_{j=1}^{d}\alpha_{\tau(g)^{-1}\varsigma j}\Phi_{\widehat{\mathsf{F}}_\varsigma}(g')_j, \sum_{k=1}^{d}\alpha_{\varsigma k}\Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_k\right\rangle\end{aligned} \tag{3.80}$$

From (3.78), (3.79) and (3.80) we obtain

$$\langle \zeta(g,g')y,y\rangle \approx[2R^{-1}]\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\tau(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\left\langle \Phi_{\widehat{\mathsf{F}}_\varsigma}(g')_j, \Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_k\right\rangle \tag{3.81}$$

Using identification given by the condition $(\mathrm{G}-1)$ we have

$$\begin{aligned}&\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\tau(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\left\langle \Phi_{\widehat{\mathsf{F}}_\varsigma}(g')_j, \Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_k\right\rangle \\ &\qquad = \frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\left\langle \Phi_{\widehat{\mathsf{F}}_\varsigma}(g')_j, \Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_k\right\rangle\end{aligned} \tag{3.82}$$

From our choice of $\widehat{\mathsf{F}}_\varsigma$ in (3.51) we have

$$\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\left\langle\Phi_{\widehat{\mathsf{F}}_\varsigma}(g')_j,\Phi_{\widehat{\mathsf{F}}_\varsigma}(e)_k\right\rangle$$
$$=\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\left\langle\Phi_{\widehat{\mathsf{E}}_\varsigma}(g')_j,\Phi_{\widehat{\mathsf{E}}_\varsigma}(e)_k\right\rangle \tag{3.83}$$

From (3.81), (3.82) and (3.83) we have

$$\langle\zeta(g,g')y,y\rangle\approx[2R^{-1}]\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\left\langle\Phi_{\widehat{\mathsf{E}}_\varsigma}(g')_j,\Phi_{\widehat{\mathsf{E}}_\varsigma}(e)_k\right\rangle \tag{3.84}$$

Since $g'\in\mathbb{B}_r$ from (3.39) we have

$$\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\widehat{\mathsf{E}}_\varsigma(g')_{j,k}\approx[R^{-1}]\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\mathsf{E}_\varsigma(g')_{j,k} \tag{3.85}$$

From our choice of E in (3.38) we have

$$\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\mathsf{E}_\varsigma(g')_{j,k}=\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\mathsf{D}_\varsigma(g')_{j,k} \tag{3.86}$$

From (3.84), (3.85) and (3.86) we have

$$\langle\zeta(g,g')y,y\rangle\approx[2R^{-1}]\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\mathsf{D}_\varsigma(g')_{j,k} \tag{3.87}$$

Since the condition $(\mathrm{G}-1)$ guarantees $|W|\geq(1-20K_rR^{-1})d!$ we have

$$\frac{1}{|W|}\sum_{\varsigma\in(W\cap\tau(g)^{-1}W)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\left\langle\Phi_{\mathsf{D}_\varsigma}(g')_j,\Phi_{\mathsf{D}_\varsigma}(e)_k\right\rangle$$
$$\approx[40K_rR^{-1}]\frac{1}{d!}\sum_{\varsigma\in\mathrm{Sym}(d)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\left\langle\Phi_{\mathsf{D}_\varsigma}(g')_j,\Phi_{\mathsf{D}_\varsigma}(e)_k\right\rangle \tag{3.88}$$

From the construction of D we have

$$\frac{1}{d!}\sum_{\varsigma\in\mathrm{Sym}(d)\setminus B}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\mathsf{D}_\varsigma(g')_{j,k}\approx[R^{-1}]\frac{1}{d!}\sum_{\varsigma\in\mathrm{Sym}(d)\setminus B}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\mathsf{C}_\varsigma(g')_{j,k} \tag{3.89}$$

From (3.87), (3.88) and (3.89) we have

$$\langle \zeta(g,g')y,y\rangle \approx[43K_rR^{-1}]\frac{1}{d!}\sum_{\varsigma\in\mathrm{Sym}(d)\setminus B}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\mathsf{C}_\varsigma(g')_{j,k} \tag{3.90}$$

We have

$$\frac{1}{d!}\sum_{\varsigma\in\mathrm{Sym}(d)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\mathsf{C}_\varsigma(g')_{j,k} = \frac{1}{d!}\sum_{\varsigma\in\mathrm{Sym}(d)}\sum_{j,k=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}\overline{\alpha_{\varsigma k}}\left\langle \rho_\triangleright(g')x_{\varsigma j},x_{\varsigma k}\right\rangle \tag{3.91}$$

$$= \frac{1}{d!}\sum_{\varsigma\in\mathrm{Sym}(d)}\left\langle \rho_\triangleright(g')\sum_{j=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}x_{\varsigma j},\sum_{k=1}^{d}\alpha_{\varsigma k}x_{\varsigma k}\right\rangle \tag{3.92}$$

where the equality in (3.91) holds by (3.22). From (3.90) and (3.92) we have

$$\langle \zeta(g,g')y,y\rangle \approx[43K_rR^{-1}]\frac{1}{d!}\sum_{\varsigma\in\mathrm{Sym}(d)}\left\langle \rho_\triangleright(g')\sum_{j=1}^{d}\alpha_{\sigma(g)^{-1}\varsigma j}x_{\varsigma j},\sum_{k=1}^{d}\alpha_{\varsigma k}x_{\varsigma k}\right\rangle \tag{3.93}$$

By making the changes of variables $j\mapsto\varsigma^{-1}\sigma(g)j$ in the left sum and $k\mapsto\varsigma^{-1}k$ in the right sum of (3.93) we obtain

$$(3.93) = \frac{1}{d!}\sum_{\varsigma\in\mathrm{Sym}(d)}\left\langle \rho_\triangleright(g')\sum_{j=1}^{d}\alpha_j x_{\sigma(g)j},\sum_{k=1}^{d}\alpha_k x_k\right\rangle$$

or equivalently

$$(3.93) = \left\langle \rho_\triangleright(g')\sum_{j=1}^{d}\alpha_j x_{\sigma(g)j},\sum_{k=1}^{d}\alpha_k x_k\right\rangle \tag{3.94}$$

From (3.93) and (3.94) we obtain

$$\langle \zeta(g,g')y,y\rangle \approx[43K_rR^{-1}]\left\langle \rho_\triangleright(g')\sum_{j=1}^{d}\alpha_j x_{\sigma(g)j},\sum_{k=1}^{d}\alpha_k x_k\right\rangle \tag{3.95}$$

From (3.10) we have

$$\left\langle \rho_\triangleright(g')\sum_{j=1}^{d}\alpha_j x_{\sigma(g)j},\sum_{k=1}^{d}\alpha_k x_k\right\rangle = \left\langle \rho_\triangleright(g')\kappa(g)\sum_{j=1}^{d}\alpha_j x_j,\sum_{k=1}^{d}\alpha_k x_k\right\rangle \tag{3.96}$$

From (3.9) we have

$$\left\langle \rho_\triangleright(g')\kappa(g)\sum_{j=1}^{d}\alpha_j x_j,\sum_{k=1}^{d}\alpha_k x_k\right\rangle \approx[r\delta]\left\langle \rho_\triangleright(g')\rho_\triangleleft(g)\sum_{j=1}^{d}\alpha_j x_j,\sum_{k=1}^{d}\alpha_k x_k\right\rangle \tag{3.97}$$

From (3.95), (3.96) and (3.97) we have

$$\langle \zeta(g,g')y,y\rangle \approx[44K_rR^{-1}] \left\langle \rho_{\triangleright}(g')\rho_{\triangleleft}(g)\sum_{j=1}^{d}\alpha_j x_j, \sum_{k=1}^{d}\alpha_k x_k \right\rangle \tag{3.98}$$

From (3.8) and (3.98) we obtain

$$\langle \zeta(g,g')y,y\rangle \approx[45K_rR^{-1}]\langle \rho(g,g')x,x\rangle$$

Using (3.1) we see that this completes the proof of Theorem 1.2.

References

[1] M. Bakonyi and D. Timotin. Extensions of positive definite functions on free groups. *J. Funct. Anal.*, 246(1):31–49, 2007.

[2] Mihály Bakonyi and Hugo J. Woerdeman. *Matrix completions, moments, and sums of Hermitian squares.* Princeton University Press, Princeton, NJ, 2011.

[3] Bachir Bekka, Pierre de la Harpe, and Alain Valette. *Kazhdan's property (T)*, volume 11 of *New Mathematical Monographs.* Cambridge University Press, Cambridge, 2008.

[4] Nathanial P. Brown and Narutaka Ozawa. *C^*-algebras and finite-dimensional approximations*, volume 88 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2008.

[5] Kenneth R. Davidson. *C^*-algebras by example*, volume 6 of *Fields Institute Monographs.* American Mathematical Society, Providence, RI, 1996.

[6] Felix R. Gantmacher. *Matrizentheorie.* Springer-Verlag, Berlin, 1986. With an appendix by V. B. Lidskij, With a preface by D. P. Želobenko, Translated from the second Russian edition by Helmut Boseck, Dietmar Soyka and Klaus Stengert.

[7] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. MIP*=RE. *arXiv e-prints*, page arXiv:2001.04383, Jan 2020.

[8] Richard V. Kadison and John R. Ringrose. *Fundamentals of the theory of operator algebras. Vol. II*, volume 16 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1997. Advanced theory, Corrected reprint of the 1986 original.

[9] Tosio Kato. *Perturbation theory for linear operators.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[10] Alexander S. Kechris. *Global aspects of ergodic group actions*, volume 160 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2010.

[11] Alexander S. Kechris. Weak containment in the space of actions of a free group. *Israel J. Math.*, 189:461–507, 2012.

[12] Eberhard Kirchberg. On nonsemisplit extensions, tensor products and exactness of group C^*-algebras. *Invent. Math.*, 112(3):449–489, 1993.

[13] Narutaka Ozawa. About the Connes embedding conjecture: algebraic approaches. *Jpn. J. Math.*, 8(1):147–183, 2013.

Department of Mathematics
The University of Texas at Austin
Austin, TX 78712

pjburton@math.utexas.edu
kate.juschenko@gmail.com