ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal of Functional Analysis 262 (2012) 1011–1042

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Absolutely continuous spectrum for the Anderson model on a product of a tree with a finite graph

Richard Froese [a,*], Florina Halasan [b], David Hasler [c]

[a] Department of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada
[b] Fakultät für Mathematik und Informatik, Friedrich-Schiller-Universität Jena, Jena, Germany
[c] Department of Mathematics, Ludwig Maximilians University, Munich, Germany

Received 2 September 2010; accepted 25 October 2011

Available online 5 November 2011

Communicated by L. Gross

Abstract

We prove the almost sure existence of absolutely continuous spectrum at low disorder for the Anderson model on the simplest example of a product of a regular tree with a finite graph. This graph contains loops of unbounded size.

Keywords: Anderson model; Extended states

1. Introduction

Since Klein's theorem on the existence of absolutely continuous spectrum for the Anderson model on a regular tree [9] was given new proofs, in [1,3], there have been several generalizations of this result to the Anderson model on other trees. For example, decorated trees was considered in [7] while substitution trees were treated in [8]. In this paper we show the almost sure existence of purely absolutely continuous spectrum at weak disorder for the Anderson model on the simplest example of a product of a regular tree with a finite graph. To our knowledge this is the first proof of extended states for the Anderson model on a graph with loops of unbounded size. Graphs with unbounded loops were considered in [4] for other types of randomness.

* Corresponding author.
E-mail address: rfroese@math.ubc.ca (R. Froese).

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2011.10.009

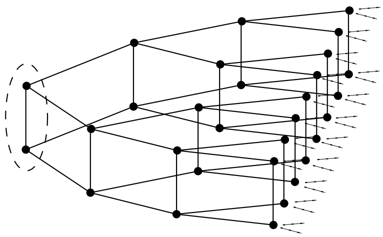

Fig. 1. The graph $T \times G$.

The Laplace operator on the product of a regular tree with a finite graph is unitarily equivalent to a direct sum of shifted Laplace operators on the base tree, where the shifts are determined by the spectrum of the Laplacian on the finite factor graph. This implies that the spectrum of the Laplace operator is the union of shifted copies of the spectrum of the base tree Laplacian. What happens when a random potential of Anderson type is added? In our example, we are able to prove the existence of absolutely continuous spectrum on the intersection of the shifted copies, namely, the interval $[-2\sqrt{2}+1, 2\sqrt{2}-1]$. We conjecture that the analogous theorem is true for general products of trees with finite graphs. (Added in proof: This has now been proved by Klein and Sadel [10].) Notice, however, that if the norm of the finite factor graph Laplacian is too large, this intersection will be empty. It is an interesting open problem to determine the nature of the spectrum for energies where only some of the shifted copies of the Laplace operator in the decomposition of the free Laplacian have spectrum. In our example these would be the energies contained in the intervals $[-2\sqrt{2}, -2\sqrt{2}+1]$ and $[2\sqrt{2}-1, 2\sqrt{2}]$. The analogous problem for slowly decaying random potentials on the strip was considered in [5], but the methods used there do not apply to the Anderson model.

In this paper the base tree T is a binary rooted tree and the finite factor graph G is the graph with two vertices connected with a single edge. This graph $T \times G$ is depicted in Fig. 1.

The Laplacian for the product graph is $\Delta = \Delta_T \otimes 1 + 1 \otimes \Delta_G$, acting on the Hilbert space $\ell^2(T \times G) = \ell^2(T) \otimes \ell^2(G) = \ell^2(T) \otimes \mathbb{C}^2$. In what follows we will think of elements of $\ell^2(T) \otimes \mathbb{C}^2$ as $\mathbb{C}^2$ valued functions on T. From this point of view, the analysis of this model can be considered to be a 2×2 matrix valued version of the model on the original tree. Roughly speaking, the hyperbolic plane $\mathbb{H}$ is replaced by the Siegel upper half space $\mathbb{SH}_2$. So, although the outline of the proof is the same as for the tree, we are confronted with non-commuting variables and the much more complicated geometry at infinity of $\mathbb{SH}_2$.

For convenience we will actually work with the adjacency matrix, which amounts to setting the diagonal matrix elements of the Laplacian to zero. Then $\Delta_G = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, and the Laplacian acts on $\varphi \in \ell^2(T) \otimes \mathbb{C}^2$ as

$$\Delta\varphi(x) = \sum_{y:y\sim x} \varphi(y) + \Delta_G \varphi(x).$$

Here $y \sim x$ means that y is connected to x by a single edge.

Let Q denote an i.i.d. random potential on T taking values in the set of 2×2 real symmetric matrices $\mathrm{Sym}(2, \mathbb{R})$. Assume that the single site distribution is given by the measure ν satisfying

$$\mathbb{E}\big[\|Q\|^{2(1+p)}\big] = \int\limits_{\mathrm{Sym}(2,\mathbb{R})} \|Q\|^{2(1+p)}\, d\nu(Q) < \infty \tag{1.1}$$

for some $p > 0$.

We study the spectral properties of the Anderson Hamiltonian

$$H_k = \Delta + kQ \tag{1.2}$$

for small coupling constant k, which we take to be positive. The goal of this paper is to prove the following theorem.

Theorem 1.1. *Let H_k be the random Anderson Hamiltonian defined by* (1.2)*, where the potential Q satisfies* (1.1)*. Let I be any closed subinterval of* $(-2\sqrt{2}+1, 2\sqrt{2}-1)$*. Then, for sufficiently small k, H has purely absolutely continuous spectrum in I almost surely.*

Here are some of the new ingredients in this paper. After a preliminary symplectic change of variables to move the fixed point of our recursion relation to iI, we define a weight function in (1.9) with some extra convexity compared to the functions we used previously (the analogue on the original tree is described in the conference proceedings review [6]). This allows a simple geometric characterization ((2.1) and (2.2)) of the places where our key inequality degenerates. This characterization involves an unusual co-ordinate system for $\mathbb{SH}_2$ given by (1.11).

1.1. The forward Green function and the recursion relation

Let P denote the rank two projection onto the space of functions supported on the vertices above the root (inside the oval in Fig. 1). Then, for λ in the resolvent set of H_k, we define the Green function at the root to be

$$G(\lambda) = P(H-\lambda)^{-1}P. \tag{1.3}$$

This Green function is a λ dependent random variable taking values in the Siegel upper half space $\mathbb{SH}_2$.

By definition, $\mathbb{SH}_2$ is the set of symmetric 2×2 matrices with complex entries whose imaginary parts are positive definite. The symplectic group $\mathrm{Sp}(4,\mathbb{R})$ acts on $\mathbb{SH}_2$ via generalized linear fractional transformations. For $\Gamma = \left[\begin{smallmatrix} A & B \\ C & D\end{smallmatrix}\right] \in \mathrm{Sp}(4,\mathbb{R})$ and $Z \in \mathbb{SH}_2$ we write the action as $\Gamma \cdot Z = (AZ+B)(CZ+D)^{-1}$. Properties of $\mathbb{SH}_2$, its compactification, and the $\mathrm{Sp}(4,\mathbb{R})$ action that we need can be found in the thesis of Freitas [2].

The forward Green functions are defined by disconnecting the tree as indicated, and restricting the resolvent for the Hamiltonians of the two disconnected subtrees to the range of the projections corresponding to the root nodes of the subtrees (see Fig. 2).

The analogue of (1.3) gives rise to two forward Green functions $G_1(\lambda)$ and $G_2(\lambda)$ that, for a given realization of the potential, are related to $G(\lambda)$ by

$$G(\lambda) = \Phi_\lambda\left(\frac{G_1(\lambda)+G_2(\lambda)}{2} - \frac{k}{2}Q\right) \tag{1.4}$$

where

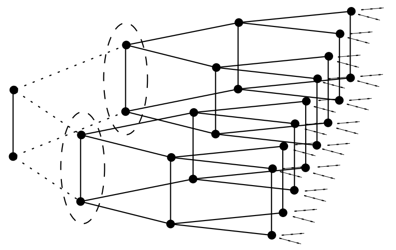

Fig. 2. Definition of the forward Green functions.

$$\Phi_\lambda(G) = (2G + \lambda - \Delta_G)^{-1}$$

and Q is the value of the potential at the root. Note that Φ_λ is the generalized linear fractional transformation which we can identify with the matrix

$$\Phi_\lambda = \begin{bmatrix} 0 & I/\sqrt{2} \\ \sqrt{2}I & (\lambda I - \Delta_G)/\sqrt{2} \end{bmatrix}.$$

If λ is real, then $\Phi_\lambda \in \mathrm{Sp}(4, \mathbb{R})$. Otherwise, Φ_λ is a composition of a complex shift with a transformation in $\mathrm{Sp}(4, \mathbb{R})$.

We define G_λ to be the fixed point of Φ_λ. Solving the fixed point equation $G_\lambda = \Phi_\lambda(G_\lambda)$ yields

$$G_\lambda = -\left(\frac{\lambda - \Delta_G}{4}\right) + i\sqrt{\frac{1}{2} - \left(\frac{\lambda - \Delta_G}{4}\right)^2}.$$

Since the eigenvalues of Δ_G are ± 1, both eigenvalues of G_λ lie on a circle of radius $1/\sqrt{2}$ in the upper half plane when $\lambda \in (-2\sqrt{2} + 1, 2\sqrt{2} - 1)$. For these values of λ, $G_\lambda \in \mathbb{SH}_2$, while for real λ outside this range, G_λ lies on the boundary at infinity. This explains the range of λ for which we can prove absolutely continuous spectrum.

We now choose a closed interval $J \subset (-2\sqrt{2} + 1, 2\sqrt{2} - 1)$ that will remain fixed for the rest of the paper. Define

$$R_\epsilon = \{\lambda \in \mathbb{C}\colon \operatorname{Re}\lambda \in J,\ 0 < \operatorname{Im}\lambda \leqslant \epsilon\} \tag{1.5}$$

with ϵ sufficiently small so that $G_\lambda \in \mathbb{SH}_2$ for $\lambda \in R_\epsilon$.

We want the fixed point G_λ to serve as an origin for $\mathbb{SH}_2$. To avoid difficulties that result from the fact that G_λ does not commute with all of $\mathbb{SH}_2$, we perform a λ dependent symplectic change of variables to move the origin to iI. For $\lambda \in R_\epsilon$, write $G_\lambda = X_\lambda + iY_\lambda$ and let Γ_λ be the symplectic transformation given by the matrix

$$\Gamma_\lambda = \begin{bmatrix} Y_\lambda^{-1/2} & -Y_\lambda^{-1/2}X_\lambda \\ 0 & Y_\lambda^{1/2} \end{bmatrix}.$$

Then $\Gamma_\lambda \cdot G_\lambda = iI$. We will work with the new variables Z in $\mathbb{SH}_2$ related to G by

$$Z = \Gamma_\lambda \cdot G = Y_\lambda^{-1/2} G Y_\lambda^{-1/2} - Y_\lambda^{-1/2} X_\lambda Y_\lambda^{-1/2}.$$

With these variables, Eq. (1.4) becomes

$$Z(\lambda) = \Psi_\lambda\left(\frac{Z_1(\lambda) + Z_2(\lambda)}{2} - \frac{k}{2}\widehat{Q}\right) \tag{1.6}$$

where $\Psi_\lambda = \Gamma_\lambda \circ \Phi_\lambda \circ \Gamma_\lambda^{-1}$ and $\widehat{Q} = Y_\lambda^{-1/2} Q Y_\lambda^{-1/2}$. For future reference we compute the matrix for Ψ_λ explicitly. This yields

$$\Psi_\lambda = \begin{bmatrix} (\lambda - \Delta_G)/(2\sqrt{2}) & -\sqrt{1 - (\lambda - \Delta_G)^2/8} \\ \sqrt{1 - (\lambda - \Delta_G)^2/8} & (\lambda - \Delta_G)/(2\sqrt{2}) \end{bmatrix} = \begin{bmatrix} \cos(\Theta_\lambda) & -\sin(\Theta_\lambda) \\ \sin(\Theta_\lambda) & \cos(\Theta_\lambda) \end{bmatrix} \tag{1.7}$$

where $\Theta_\lambda = \cos^{-1}((\lambda - \Delta_G)/(2\sqrt{2}))$ with the branch of $\cos^{-1}$ chosen to make $\sin(\Theta_\lambda)$ positive definite when $\lambda \in J$. Notice that Ψ_λ is an orthogonal symplectic matrix for $\lambda \in J$.

Eq. (1.6), and the self-similarity of the tree imply that for any positive measurable function w on $\mathbb{SH}_2$,

$$\mathbb{E}\big[w\big(Z(\lambda)\big)\big] = \mathbb{E}\left[w\left(\Psi_\lambda\left(\frac{Z_1(\lambda) + Z_2(\lambda)}{2} - \frac{k}{2}\widehat{Q}\right)\right)\right], \tag{1.8}$$

where $Z_1(\lambda)$, $Z_2(\lambda)$ are independent copies of $Z(\lambda)$ and Q is independently distributed according to ν.

1.2. The functions $w_p(Z_1, Z_2)$ and $\mu^*_{2,p}(Z_1, Z_2)$

The following symplectically invariant function will play an important role in our analysis. For $Z_j = X_j + iY_j$, $j = 1, 2$ and $p > 0$, let

$$w_p(Z_1, Z_2) = \big\| Y_2^{-1/2}(Z_1 - Z_2)^* Y_1^{-1}(Z_1 - Z_2) Y_2^{-1/2} \big\|_{1+p}^{1+p} \tag{1.9}$$

where $\|\cdot\|_{1+p}$ denotes the Schatten $(1 + p)$ norm. When $p = 0$ the norm gives the trace, and the resulting definition is a function of the Riemannian distance in the Siegel space. As we will see below, w_p is still invariant under the symplectic action when $p > 0$, and the extra convexity that results for positive p will be important.

The weight function that we use to measure growth in $\mathbb{SH}_2$ is defined to be

$$w_p(Z) = w_p(Z, iI). \tag{1.10}$$

The following lemma collects some properties of $w_p(Z_1, Z_2)$ and $w_p(Z)$.

Lemma 1.2.

(i) *Let Γ be an element of* $\mathrm{Sp}(4, \mathbb{R})$ *acting on* $\mathbb{SH}_2$. *Then*

$$w_p(\Gamma \cdot Z_1, \Gamma \cdot Z_2) = w_p(Z_1, Z_2).$$

(ii) *Let T be a complex translation given by the action $T \cdot Z = Z + it$ with $t > 0$. Then*

$$w_p(T \cdot Z_1, T \cdot Z_2) < w_p(Z_1, Z_2).$$

(iii) *There are constants C_1 and C_2 such that for every $Z \in \mathbb{SH}_2$,*

$$\|\mathrm{Im}(Z)\|^{1+p} \leqslant C_1 w_p(Z) + C_2.$$

(iv) *For any $\epsilon > 0$ there exists C_ϵ such that for any $Q \in \mathrm{Sym}(2, \mathbb{R})$*

$$w_p(Z+Q) \leqslant \bigl(1+\epsilon + C_\epsilon \|Q\|^{2(1+p)}\bigr) w_p(Z) + C_\epsilon \|Q\|^{2(1+p)}.$$

This lemma is proved in Appendix A.

The ratio

$$\mu_{2,p,\lambda}(Z_1, Z_2) = w_p\left(\Psi_\lambda\left(\frac{Z_1+Z_2}{2}\right)\right) \Big/ \left(\frac{1}{2} w_p(Z_1) + \frac{1}{2} w_p(Z_2)\right)$$

plays a central role in our analysis. To understand this function we introduce an unusual co-ordinate system for $\mathbb{SH}_2$. For $Z = X + iY \in \mathbb{SH}_2$, define

$$U(Z) = Y^{-1/2}(Z - iI). \tag{1.11}$$

We will study this co-ordinate system in detail below. Clearly $w_p(Z) = \|U(Z)^* U(Z)\|_{1+p}^{1+p} = \|U(Z)\|_{2(1+p)}^{2(1+p)}$. The quantity $U(Z)$ appears in the following crucial formula.

Proposition 1.3. *For* $\mathrm{Im}\,\lambda \geqslant 0$,

$$w_p\left(\Psi_\lambda\left(\frac{Z_1+Z_2}{2}\right)\right) \leqslant \left\| \frac{1}{2}\bigl[U(Z_1)^*, U(Z_2)^*\bigr] P(Y_1, Y_2) \begin{bmatrix} U(Z_1) \\ U(Z_2) \end{bmatrix} \right\|_{1+p}^{1+p} \tag{1.12}$$

where

$$P(Y_1, Y_2) = \begin{bmatrix} Y_1^{1/2} \\ Y_2^{1/2} \end{bmatrix} (Y_1 + Y_2)^{-1} \bigl[Y_1^{1/2}, Y_2^{1/2}\bigr]$$

is the orthogonal projection onto the range of $\begin{bmatrix} Y_1^{1/2} \\ Y_2^{1/2} \end{bmatrix}$. *The inequality is an equality if* $\lambda \in \mathbb{R}$.

Notice that the left side of (1.12) does not depend on λ, so we can define the λ independent upper bound for $\mu_{2,p,\lambda}$

$$\mu_{2,p}^*(Z_1, Z_2) = \left\| \frac{1}{2}\bigl[U(Z_1)^*, U(Z_2)^*\bigr] P(Y_1, Y_2) \begin{bmatrix} U(Z_1) \\ U(Z_2) \end{bmatrix} \right\|_{1+p}^{1+p} \Big/ \left(\frac{1}{2} w_p(Z_1) + \frac{1}{2} w_p(Z_2)\right).$$

It follows from Proposition 1.3 that

$$\mu_{2,p,\lambda} \leqslant \mu_{2,p}^*. \tag{1.13}$$

Proposition 1.4. *The ratio $\mu^*_{2,p}(Z_1, Z_2) \leqslant 1$, or equivalently*

$$\left\| \frac{1}{2}\big[U(Z_1)^*, U(Z_2)^*\big] P(Y_1, Y_2) \begin{bmatrix} U(Z_1) \\ U(Z_2) \end{bmatrix} \right\|_{1+p}^{1+p} \leqslant \frac{1}{2} w_p(Z_1) + \frac{1}{2} w_p(Z_2). \tag{1.14}$$

Equality holds if and only if $Z_1 = Z_2$.

These propositions are proved below, where we also determine in what form they survive on the compactifications considered below.

1.3. Reduction to estimates on $\mu^*_{2,p}$

If $Z = \Gamma_\lambda \cdot G$ then $\operatorname{Im} G = Y_\lambda^{1/2} \operatorname{Im} Z \, Y_\lambda^{1/2}$. Thus, $\|\operatorname{Im} G\| \leqslant C \|\operatorname{Im} Z\|$ uniformly for $\lambda \in R_\epsilon$ with ϵ small. So, given Lemma 1.2(iii), Theorem 1.1 follows from the following theorem (see, e.g., Lemma 1 of [4]).

Theorem 1.5. *Let $G(\lambda)$ be the Green function for the random Hamiltonian $H_k = \Delta + kQ$ defined by* (1.3)*, and let $Z(\lambda) = \Gamma_\lambda \cdot G(\lambda)$. Then, for sufficiently small coupling constant k, and small ϵ, there exists a constant C such that*

$$\sup_{\lambda \in R_\epsilon} \mathbb{E}\big[w_p\big(Z(\lambda)\big)\big] \leqslant C.$$

In this section we will indicate how this theorem follows from estimates of $\mu^*_{2,p}$ at infinity. This part of the proof follows the same lines as [3]. Using (1.8) twice we find that

$$\mathbb{E}\big[w_p\big(Z(\lambda)\big)\big] = \mathbb{E}\left[w_p\left(\Psi_\lambda\left(\frac{1}{2} Z_1 + \frac{1}{2} \Psi_\lambda\left(\frac{1}{2} Z_2 + \frac{1}{2} Z_3 - \frac{1}{2} k \widehat{Q}_2\right) - \frac{1}{2} k \widehat{Q}_1\right)\right)\right]$$

where Z_1, Z_2 and Z_3 are independent copies of $Z(\lambda)$ and Q_1 and Q_2 are independent copies of the single site (matrix) potential. Since we may permute Z_1, Z_2 and Z_3 without changing the expectation, we find

$$\mathbb{E}\big[w_p\big(Z(\lambda)\big)\big] = \frac{1}{3} \mathbb{E}\big[\Sigma(Z_1, Z_2, Z_3, k\widehat{Q}_1, k\widehat{Q}_2, \lambda)\big]$$

where Σ is the symmetrization of the expression above given by

$$\Sigma(Z_1, Z_2, Z_3, Q_1, Q_2, \lambda) = \sum_\sigma w_p\left(\Psi_\lambda\left(\frac{1}{2} Z_{\sigma_1} + \frac{1}{2} \Psi_\lambda\left(\frac{1}{2} Z_{\sigma_2} + \frac{1}{2} Z_{\sigma_3} - \frac{1}{2} Q_2\right) - \frac{1}{2} Q_1\right)\right).$$

In the sum, σ ranges over the three cyclic permutations of $(1, 2, 3)$.

Introduce the ratio

$$\mu_3(Z_1, Z_2, Z_3, Q_1, Q_2, \lambda) = \frac{\Sigma(Z_1, Z_2, Z_3, Q_1, Q_2, \lambda)}{w_p(Z_1) + w_p(Z_2) + w_p(Z_3)}.$$

To prove our main theorem, we will prove that

Proposition 1.6. *There exists a compact set $K \subseteq \mathbb{SH}_2 \times \mathbb{SH}_2 \times \mathbb{SH}_2$, $\epsilon > 0$, $\epsilon_1 > 0$ and $\delta > 0$ so that*

$$\sup_{(Z_1,Z_2,Z_3)\notin K,\ \|Q_1\|\leqslant\epsilon_1,\ \|Q_2\|\leqslant\epsilon_1,\ \lambda\in R_\epsilon} \mu_3(Z_1, Z_2, Z_3, Q_1, Q_2, \lambda) \leqslant (1-\delta).$$

Given Proposition 1.6 we can prove Theorem 1.5 as follows.

Proof of Theorem 1.5. Choose ϵ, ϵ_1, K and δ so that the estimate in Proposition 1.6 holds. Let $\chi(\cdot)$ denote the characteristic function of the indicated set. We can then estimate $\mathbb{E}[w_p(Z(\lambda))]$ by introducting cutoffs as follows.

$$\begin{aligned}\mathbb{E}\big[w_p\big(Z(\lambda)\big)\big] \leqslant{}& \frac{1}{3}\mathbb{E}\big[\chi\big(\|k\widehat{Q}_1\| \leqslant \epsilon_1, \|k\widehat{Q}_2\| \leqslant \epsilon_1, (Z_1, Z_2, Z_3) \notin K\big)\Sigma\big] \\ &+ \frac{1}{3}\mathbb{E}\big[\chi\big(\|k\widehat{Q}_1\| \leqslant \epsilon_1, \|k\widehat{Q}_2\| \leqslant \epsilon_1, (Z_1, Z_2, Z_3) \in K\big)\Sigma\big] \\ &+ \frac{1}{3}\mathbb{E}\big[\chi\big(\|k\widehat{Q}_1\| > \epsilon_1\big)\Sigma\big] + \frac{1}{3}\mathbb{E}\big[\chi\big(\|k\widehat{Q}_2\| > \epsilon_1\big)\Sigma\big]. \end{aligned} \tag{1.15}$$

Here Σ stands for $\Sigma(Z_1, Z_2, Z_3, k\widehat{Q}_1, k\widehat{Q}_2, \lambda)$. In the first term on the right of (1.15) we may replace Σ with $(1-\delta)(w_p(Z_1) + w_p(Z_2) + w_p(Z_3))$ for any $\lambda \in R_\epsilon$, thanks to *Proposition* 1.6. This results in the following estimate for the first term on the right of (1.15), valid for all $\lambda \in R_\epsilon$

$$\frac{1}{3}\mathbb{E}\big[\chi\big(\|k\widehat{Q}_1\| \leqslant \epsilon_1, \|k\widehat{Q}_2\| \leqslant \epsilon_1, (Z_1, Z_2, Z_3) \notin K\big)\Sigma\big] \leqslant (1-\delta)\mathbb{E}\big[w_p\big(Z(\lambda)\big)\big].$$

The second term in the right of (1.15) is estimated by noting that $\Sigma(Z_1, Z_2, Z_3, Q_1, Q_2, \lambda)$ is continuous on $\mathbb{SH}_2 \times \mathbb{SH}_2 \times \mathbb{SH}_2 \times \mathrm{Sym}(2,\mathbb{R}) \times \mathrm{Sym}(2,\mathbb{R}) \times \overline{R}_\epsilon$ and therefore bounded on a compact subset. This yields the following estimate for the second term on the right of (1.15), again valid for all $\lambda \in R_\epsilon$

$$\frac{1}{3}\mathbb{E}\big[\chi\big(\|k\widehat{Q}_1\| \leqslant \epsilon_1, \|k\widehat{Q}_2\| \leqslant \epsilon_1, (Z_1, Z_2, Z_3) \in K\big)\Sigma\big] \leqslant C(\epsilon, \epsilon_1, K).$$

The last two terms on the right of (1.15) are handled identically, so we will focus on the third term. This is where the assumption of low disorder, i.e., that k is sufficiently small, enters. We wish to exploit the fact that $\chi(\|k\widehat{Q}_1\| > \epsilon_1) \to 0$ as $k \to 0$, pointwise in Q_1. To do this we will need the following upper bound for Σ

$$\begin{aligned}&\Sigma(Z_1, Z_2, Z_3, Q_1, Q_2, \lambda) \\ &\quad\leqslant C\big(1 + \|Q_1\|^{2(1+p)} + \|Q_2\|^{2(1+p)}\big)\big(w_p(Z_1) + w_p(Z_2) + w_p(Z_3) + 1\big).\end{aligned} \tag{1.16}$$

Before proving this inequality, let us see how it can be used to complete the proof. Recall that $\widehat{Q}$ denotes $Y_\lambda^{-1/2} Q Y_\lambda^{-1/2}$ so that $\|\widehat{Q}\| \leqslant C\|Q\|$ with C uniform for $\lambda \in R_\epsilon$. Thus, using (1.16) and the independence of the random variables Q_1, Q_2, Z_1, Z_2, Z_3 we find that for bounded k there exists a constant C such that

$$\begin{aligned}
&\mathbb{E}\big[\chi\big(\|k\widehat{Q}_1\| > \epsilon_1\big)\Sigma\big] \\
&\quad \leqslant C\mathbb{E}_{Q_1,Q_2}\big[\chi\big(\|k\widehat{Q}_1\| > \epsilon_1\big)\big(1 + \|Q_1\|^{2(1+p)} + \|Q_2\|^{2(1+p)}\big)\big]\big(3\mathbb{E}\big[w_p\big(Z(\lambda)\big)\big] + 1\big) \\
&\quad = \delta(k, \epsilon_1)\big(\mathbb{E}\big[w_p\big(Z(\lambda)\big)\big] + 1\big)
\end{aligned}$$

where $\delta(k, \epsilon_1) \to 0$ as $k \to 0$. Given (1.1), this follows from the Lebesgue dominated convergence theorem applied to $\mathbb{E}_{Q_1,Q_2}[\chi(\|k\widehat{Q}_1\| > \epsilon_1)(1 + \|Q_1\|^{2(1+p)} + \|Q_2\|^{2(1+p)})]$. Combining this estimate with the previous estimates for the first and second terms on the right of (1.15) we obtain

$$\begin{aligned}
\mathbb{E}\big[w_p\big(Z(\lambda)\big)\big] &\leqslant \big(1 - \delta + \delta(k, \epsilon_1)\big)\mathbb{E}\big[w_p\big(Z(\lambda)\big)\big] + C \\
&\leqslant (1 - \delta/2)\mathbb{E}\big[w_p\big(Z(\lambda)\big)\big] + C
\end{aligned}$$

for k sufficiently small, valid for all $\lambda \in R_\epsilon$. Since the constants here are independent of $\lambda \in R_\epsilon$, this implies the bound of Theorem 1.5 and completes the proof, provided we can rule out $\mathbb{E}[w_p(Z(\lambda))] = \infty$.

It remains to establish (1.16) and to prove an a priori estimate for $\mathbb{E}[w_p(Z(\lambda))]$.

We begin by proving (1.16). Proposition 1.3 and Proposition 1.4 imply that $w_p(\Psi_\lambda(\frac{Z_1+Z_2}{2})) \leqslant \frac{1}{2}w_p(Z_1) + \frac{1}{2}w_p(Z_2)$. Repeated applications of this inequality, together with Lemma 1.2(iv) with any choice of ϵ, which we write in the less precise form $w_p(Z - Q) \leqslant C(1 + \|Q\|^{2(1+p)})(w_p(Z) + 1)$, yield

$$\begin{aligned}
&\Sigma(Z_1, Z_2, Z_3, Q_1, Q_2, \lambda) \\
&\quad \leqslant \sum_\sigma \left[\frac{1}{2}w_p(Z_{\sigma_1} - Q_1) + \frac{1}{4}w_p(Z_{\sigma_2} - Q_2) + \frac{1}{4}w_p(Z_{\sigma_3})\right] \\
&\quad \leqslant \sum_\sigma \Big[C\big(1 + \|Q_1\|^{2(1+p)}\big)\big(w_p(Z_{\sigma_1}) + 1\big) + C\big(1 + \|Q_2\|^{2(1+p)}\big) \\
&\qquad \times \big(w_p(Z_{\sigma_2}) + 1\big) + \frac{1}{4}w_p(Z_{\sigma_3})\Big].
\end{aligned}$$

This implies (1.16).

Finally we turn to the a priori bound. We need to prove $\mathbb{E}[w_p(Z(\lambda))] \leqslant C(\lambda)$, where the constant $C(\lambda)$ may blow up as $\operatorname{Im}\lambda$ becomes small. We will show that for any realization of the potential,

$$w_p\big(Z(\lambda)\big) \leqslant C(\lambda)\big(1 + \|Q\|^{2(1+p)}\big), \tag{1.17}$$

where $C(\lambda)$ does not depend on the potential. Then the bound follows by taking the expectation.

For this bound it is more convenient to work with the original forward Green function $G(\lambda)$. By Lemma 1.2(i) we have $w_p(Z(\lambda)) = w_p(Z(\lambda), iI) = w_p(G(\lambda), G_\lambda)$. For any realization of the potential, the recursion relation can be written $G(\lambda) = -(G_1 + G_2 + \lambda - \Delta_G - Q)^{-1}$, where we are writing G_i for $G_i(\lambda)$. Thus

$$\left\|G(\lambda)\right\| \leqslant \sup_{Z\in S\mathbb{H}_2} \left\|(Z+i\operatorname{Im}\lambda)^{-1}\right\| \leqslant C/\operatorname{Im}\lambda$$

(see Lemma A.1 in Appendix A for the second inequality). The same estimate holds for $\|G_1\|$ and $\|G_2\|$. Now let $Y_1 = \operatorname{Im}(G_1 + G_2 + \lambda - \Delta_G - Q)$ and $Y_2 = \operatorname{Im}(2G_\lambda + \lambda - \Delta_G)$. Notice that for $i = 1, 2$, $Y_i \geqslant \operatorname{Im}\lambda$ and so, since Y_i is real symmetric, $Y_i^{-1} \leqslant 1/\operatorname{Im}\lambda$. Now we use the invariance of w_p in Lemma 1.2(i), and the fixed point property of G_λ to write

$$\begin{aligned}
w_p\bigl(G(\lambda), G_\lambda\bigr) &= w_p\bigl(-(G_1+G_2+\lambda-\Delta_G-Q)^{-1}, -(2G_\lambda+\lambda-\Delta_G)^{-1}\bigr)\\
&= w_p(G_1+G_2+\lambda-\Delta_G-Q, 2G_\lambda+\lambda-\Delta_G)\\
&= \bigl\|Y_2^{-1/2}\bigl(G_1^*+G_2^*-Q-2G_\lambda^*\bigr)Y_1^{-1}(G_1+G_2-Q-2G_\lambda)Y_2^{-1/2}\bigr\|_{1+p}^{1+p}\\
&\leqslant \bigl(\operatorname{Im}\lambda^{-1}\bigl(\|G_1\|_{1+p}+\|G_2\|_{1+p}+\|Q\|_{1+p}+2\|G_\lambda\|_{1+p}\bigr)\bigr)^{2(1+p)}\\
&\leqslant \bigl(\operatorname{Im}\lambda^{-1}\bigl(C/\operatorname{Im}\lambda+\|Q\|_{1+p}+2\|G_\lambda\|_{1+p}\bigr)\bigr)^{2(1+p)}.
\end{aligned}$$

Since $\|G_\lambda\|_{1+p}$ is a λ dependent constant, independent of the potential, and all norms are equivalent for 2×2 matrices, this inequality implies (1.17). $\square$

Our next task is to reduce Proposition 1.6 to a statement about $\mu_{2,p}^*$.

The standard compactification $\overline{\mathbb{SH}}_2$ of $\mathbb{SH}_2$ is obtained by using the ball model. This is the set of all symmetric 2×2 complex matrices W with $\|W\| < 1$. Here the norm is the operator norm, W being regarded as an operator on a two-dimensional ℓ^2 space. The upper half space model and the ball model are related by the map $Z \mapsto (Z-iI)(Z+iI)^{-1} = (Z+iI)^{-1}(Z-iI)$ and its inverse. The ball model can be compactified in a natural way, by taking its closure in the Euclidean topology. The boundary of this closure, which we identify with the boundary at infinity, $\partial_\infty\mathbb{SH}_2$, of $\mathbb{SH}_2$, contains all symmetric 2×2 complex matrices W with $\|W\| = 1$ Thus, $\overline{\mathbb{SH}}_2 = \mathbb{SH}_2 \cup \partial_\infty\mathbb{SH}_2$. For more information, see [2]. We now extend μ_3 to the compactification $\overline{\mathbb{SH}}_2 \times \overline{\mathbb{SH}}_2 \times \overline{\mathbb{SH}}_2 \times \operatorname{Sym}(2,\mathbb{R}) \times \operatorname{Sym}(2,\mathbb{R}) \times \overline{R}_\epsilon$ by defining its value at a boundary point as the supremum of all values along all sequences converging to the boundary point in the topology of the compactification. Since the resulting function is upper semicontinuous, Proposition 1.6 follows if we show that the value of μ_3 on any point of the boundary $\overline{\mathbb{SH}}_2 \times \overline{\mathbb{SH}}_2 \times \overline{\mathbb{SH}}_2 \times \{0\} \times \{0\} \times J$ is < 1. Recall that J is the real interval at the base of R_ϵ.

First, let us show that $\mu_3 \leqslant 1$ on the boundary. Let $(Z_1, Z_2, Z_3) \in \partial_\infty(\mathbb{SH}_2 \times \mathbb{SH}_2 \times \mathbb{SH}_2)$ (this means that at least one Z_i is in $\partial_\infty\mathbb{SH}_2$) and $\lambda \in J \subset \mathbb{R}$. To estimate the value of μ_3 at the boundary point $(Z_1, Z_2, Z_3, 0, 0, \lambda)$ let $(Z_{1,n}, Z_{2,n}, Z_{3,n}, Q_{1,n}, Q_{2,n}, \lambda_n)$ converge to this point in the topology of the compactification. We must bound μ_3 along this sequence.

A calculation together with the inequality (1.13) shows that

$$\begin{aligned}
&\mu_3(Z_1, Z_2, Z_3, Q_1, Q_2, \lambda)\\
&\quad= \sum_\sigma \mu_{2,p}^*\left(Z_{\sigma_1} - 2Q_1, \Psi_\lambda\left(\frac{1}{2}Z_{\sigma_2} + \frac{1}{2}Z_{\sigma_3} - Q_2\right)\right)\\
&\quad\times\left(\frac{\frac{1}{2}w_p(Z_{\sigma_1}-2Q_1) + \frac{1}{4}\mu_{2,p}^*(Z_{\sigma_2}-Q_2, Z_{\sigma_3}-Q_2)(w_p(Z_{\sigma_2}-Q_2)+w_p(Z_{\sigma_3}-Q_2))}{w_p(Z_1)+w_p(Z_2)+w_p(Z_3)}\right).
\end{aligned} \tag{1.18}$$

By Proposition 1.4 we know $\mu^*_{2,p} \leqslant 1$ so when evaluated at $(Z_{1,n}, Z_{2,n}, Z_{3,n}, Q_{1,n}, Q_{2,n}, \lambda_n)$,

$$\mu_3 \leqslant \sum_\sigma \left(\frac{\frac{1}{2} w_p(Z_{\sigma_1,n} - 2Q_{1,n}) + \frac{1}{4}(w_p(Z_{\sigma_2,n} - Q_{2,n}) + w_p(Z_{\sigma_3,n} - Q_{2,n}))}{w_p(Z_{1,n}) + w_p(Z_{2,n}) + w_p(Z_{3,n})} \right).$$

Thus by Lemma 1.2(iv), since $Q_{1,n}, Q_{2,n}$ tend to zero, the limit is $\leqslant 1$.

Since the symplectic action $Z \mapsto Z + Q$ for $Q \in \mathrm{Sym}(2, \mathbb{R})$ extends continuously to the boundary at infinity, the sequence $Z_n + Q_n$ will converge to Z in the compactification if $Z_n \to Z$ and $Q_n \to 0$. Thus (1.18) implies that if $\mu_3 \to 1$ along a sequence converging to $(Z_1, Z_2, Z_3, 0, 0, \lambda)$ in the compactification, then there are sequences $Z_{1,n} \to Z_1$, $Z_{2,n} \to Z_2$, $Z_{3,n} \to Z_3$ and $\lambda_n \to \lambda$ such that

$$\mu^*_{2,p}(Z_{1,n}, Z_{2,n}) \to 1, \qquad \mu^*_{2,p}(Z_{1,n}, Z_{3,n}) \to 1, \qquad \mu^*_{2,p}(Z_{2,n}, Z_{3,n}) \to 1, \tag{1.19}$$

and

$$\begin{aligned} &\mu^*_{2,p}\left(Z_{1,n}, \Psi_{\lambda_n}\left(\frac{Z_{2,n} + Z_{3,n}}{2}\right)\right) \to 1, \\ &\mu^*_{2,p}\left(Z_{2,n}, \Psi_{\lambda_n}\left(\frac{Z_{3,n} + Z_{1,n}}{2}\right)\right) \to 1, \\ &\mu^*_{2,p}\left(Z_{3,n}, \Psi_{\lambda_n}\left(\frac{Z_{1,n} + Z_{2,n}}{2}\right)\right) \to 1. \end{aligned} \tag{1.20}$$

The sequences in each limit may be different.

The way one might hope to use these equations is to show that if $\mu^*_{2,p}(Z_{1,n}, Z_{2,n}) \to 1$ then the limits Z_1 and Z_2 are equal, that is $Z_1 = Z_2 = Z$, and that $(Z_{1,n} + Z_{2,n})/2 \to Z$ too. The second statement is not automatic because addition does not extend continuously to the compactification. This would be a plausible extension of Proposition 1.4, and can be shown to hold for a tree. Then (1.19) and (1.20) would imply that there is a Z on the boundary at infinity with $\Psi_\lambda(Z) = Z$. This contradiction would prove the desired inequality and hence Proposition 1.6.

This approach fails for the product graph we are considering. However the following two propositions can be used in an analogous way. The next proposition says that even though it is possible that $\mu^*_{2,p}(Z_{1,n}, Z_{2,n}) \to 1$ without $Z_1 = Z_2$, the limit condition does imply that both Z_1 and Z_2 belong to the same set, an image of $\overline{\mathbb{H}}$ imbedded in $\mathbb{SH}_2$ described by (ii) or (iii) below.

Proposition 1.7. *Let $(Z_{1,n}, Z_{2,n})$ be a sequence converging to a point in $\overline{\mathbb{SH}}_2 \times \overline{\mathbb{SH}}_2$ with $\mu^*_{2,p}(Z_{1,n}, Z_{2,n}) \to 1$. Then either:*

(i) *$Z_{1,n}$, $Z_{2,n}$ and the average $Z_{a,n} = (Z_{1,n} + Z_{2,n})/2$ (possibly for a subsequence) all converge to the same point in $\overline{\mathbb{SH}}_2$. In other words, the corresponding points $W_{1,n}$, $W_{2,n}$ and $W_{a,n}$ in the ball model converge to the same point in the Euclidean topology.*

(ii) *There exists a real orthogonal matrix V such that $W_{i,n} \to V\left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha_i \end{smallmatrix}\right]V^t$ for $i = 1, 2, a$ (possibly for a subsequence). Here $|\alpha_i| \leqslant 1$ and the limit W_i lies on the boundary of the ball model.*

(iii) *There exists a real orthogonal matrix V and $r, p \in \mathbb{R}$ such that $Z_{i,n} \to V\left[\begin{smallmatrix} z_i & r \\ r & p \end{smallmatrix}\right]V^t$ for $i = 1, 2, a$ (possibly for a subsequence). Here $z_i \in \overline{\mathbb{H}}$ and the limit Z_i lies on the boundary of the upper half space model.*

The next proposition says that the sets described above do not intersect their images under Ψ_λ.

Proposition 1.8.

(i) *If Z lies on the boundary of $\mathbb{SH}_2$ in the upper half space model, then $\Psi_\lambda(Z) \neq Z$ for every $\lambda \in J$.*
(ii) *Suppose V is a real orthogonal matrix and $V\left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right]V^t$ and $V\left[\begin{smallmatrix} 1 & 0 \\ 0 & \beta \end{smallmatrix}\right]V^t$ with $|\alpha| \leqslant 1$, $|\beta| \leqslant 1$ are two points in the boundary of the ball model. Then for every $\lambda \in J$*

$$\tilde{\Psi}_\lambda\left(V\begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix}V^t\right) \neq V\begin{bmatrix} 1 & 0 \\ 0 & \beta \end{bmatrix}V^t.$$

Here $\tilde{\Psi}_\lambda$ denotes the action of Ψ_λ conjugated to act on the ball model.
(iii) *Suppose V is a real orthogonal matrix and $V\left[\begin{smallmatrix} z & r \\ r & p \end{smallmatrix}\right]V^t$ and $V\left[\begin{smallmatrix} z' & r \\ r & p \end{smallmatrix}\right]V^t$ with $z, z' \in \overline{\mathbb{H}}$ are two points on the boundary in the upper half space model. Then for every $\lambda \in J$*

$$\Psi_\lambda\left(V\begin{bmatrix} z & r \\ r & p \end{bmatrix}V^t\right) \neq V\begin{bmatrix} z' & r \\ r & p \end{bmatrix}V^t.$$

We now show how Proposition 1.7 and Proposition 1.8 imply Proposition 1.6 and thus our main result.

Proof of Proposition 1.6. Suppose $\mu_3 \to 1$ along a sequence $(Z_{1,n}, Z_{2,n}, Z_{3,n}, Q_{1,n}, Q_{2,n}, \lambda_n)$ converging to $(Z_1, Z_2, Z_3, 0, 0, \lambda)$ in the compactification. Then, there are sequences so that (1.19) and (1.20) hold. Then, by Proposition 1.7 there are three possibilities. (i): $Z_{i,n}$ and $(Z_{i,n} + Z_{j,n})/2$ (possibly for a subsequence) all converge to the same point Z and thus $\Psi_\lambda(Z) = Z$. This is not possible since the only fixed point of Ψ_λ in $\overline{\mathbb{SH}_2}$ is iI. (We leave the proof of the required continuity in λ to the reader.) The second possibility is (ii): $Z_{i,n}$ and $(Z_{i,n} + Z_{j,n})/2$ (possibly for a subsequence) when viewed in the ball model all converge to matrices of the form $V\left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right]V^t$ for the same real rotation matrix V but possibly different values of α with $|\alpha| \leqslant 1$. Then (1.20) implies that there exist α and β with $|\alpha| \leqslant 1$, $|\beta| \leqslant 1$ such that $\tilde{\Psi}_\lambda \cdot (V\left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right]V^t) = V\left[\begin{smallmatrix} 1 & 0 \\ 0 & \beta \end{smallmatrix}\right]V^t$. This is impossible by Proposition 1.8(ii). Otherwise (iii): $Z_{i,n}$ and $(Z_{i,n} + Z_{j,n})/2$ (possibly for a subsequence) all converge to matrices of the form $V\left[\begin{smallmatrix} z & r \\ r & p \end{smallmatrix}\right]V^t$ for the same real rotation matrix V and the same values of $r, p \in \mathbb{R}$ but possibly different values of $z \in \overline{\mathbb{H}}$. Then (1.20) implies that there exist $z, z' \in \overline{\mathbb{H}}$ such that $\Psi_\lambda \cdot (V\left[\begin{smallmatrix} z & r \\ r & p \end{smallmatrix}\right]V^t) = V\left[\begin{smallmatrix} z' & r \\ r & p \end{smallmatrix}\right]V^t$. This is impossible by Proposition 1.8(iii). Since all cases lead to a contradiction, we conclude that $\mu_3 \to 1$ is not possible. Therefore Proposition 1.6 holds. □

2. Proofs

In this section we will show how the geometric formula for $\mu^*_{2,p}$ given after Proposition 1.3 allows us to prove Proposition 1.4 and its extension Proposition 1.7. What emerges is that there are two separate relevant quantities – the projections of $U(Z_1)$ and $U(Z_2)$ onto their unit spheres, and the range of $P(Y_1, Y_2)$ – whose limits are constrained when $\mu_{2,p,\lambda}$ tends to 1. Understanding these constraints leads to a proof of our results. We conclude the section with a proof of Proposition 1.8.

The proof of Proposition 1.3 is a simple calculation.

Proof of Proposition 1.3. Since $\Psi_\lambda \cdot (iI) = iI$ we have

$$\begin{aligned} w_p\left(\Psi_\lambda\left(\frac{Z_1+Z_2}{2}\right)\right) &= w_p\left(\Psi_\lambda\left(\frac{Z_1+Z_2}{2}\right), iI\right) \\ &= w_p\left(\Psi_\lambda\left(\frac{Z_1+Z_2}{2}\right), \Psi_\lambda(iI)\right) \\ &\leqslant w_p\left(\frac{Z_1+Z_2}{2}, iI\right) \\ &= \left\| \frac{1}{2}\big(\big(Z_1^*+Z_2^*+2iI\big)(Y_1+Y_2)^{-1}(Z_1+Z_2-2iI)\big)\right\|_{1+p}^{1+p} \\ &= \left\| \frac{1}{2}\big[U(Z_1)^*, U(Z_2)^*\big]P(Y_1,Y_2)\begin{bmatrix} U(Z_1) \\ U(Z_2)\end{bmatrix}\right\|_{1+p}^{1+p}. \end{aligned}$$

The inequality in the third line follows from Lemma 1.2(i) and (ii) and the fact that Ψ_λ is a composition of a transformation in $\mathrm{Sp}(4,\mathbb{R})$ and a complex translation by $\operatorname{Im}\lambda$. If $\operatorname{Im}\lambda = 0$ the complex translation is missing and the inequality becomes an equality. □

Proof of Proposition 1.4. We need to estimate a quantity of the form $\|\frac{1}{2}[U_1^*, U_2^*]P\big[\begin{smallmatrix} U_1 \\ U_2\end{smallmatrix}\big]\|_{1+p}^{1+p}$ where U_1 and U_2 are 2×2 matrices and P is a self-adjoint rank 2 projection. The first inequality is

$$\left\| \frac{1}{2}\big[U_1^*, U_2^*\big]P\begin{bmatrix} U_1 \\ U_2\end{bmatrix}\right\|_{1+p}^{1+p} \leqslant \left\| \frac{1}{2}\big[U_1^*, U_2^*\big]\begin{bmatrix} U_1 \\ U_2\end{bmatrix}\right\|_{1+p}^{1+p} = \left\| \frac{1}{2}\big(U_1^*U_1 + U_2^*U_2\big)\right\|_{1+p}^{1+p}.$$

Since the $(1+p)$ norm takes account of all the singular values, this inequality is strict unless

$$\operatorname{Ran}\begin{bmatrix} U_1 \\ U_2\end{bmatrix} \subseteq \operatorname{Ran} P. \tag{2.1}$$

Next we use the triangle inequality for the norm $\|\cdot\|_{1+p}$ to conclude

$$\left\| \frac{1}{2}\big(U_1^*U_1 + U_2^*U_2\big)\right\|_{1+p}^{1+p} \leqslant \left(\frac{1}{2}\big\|U_1^*U_1\big\|_{1+p} + \frac{1}{2}\big\|U_2^*U_2\big\|_{1+p}\right)^{1+p}.$$

Since $p > 0$, the unit ball in the norm $\|\cdot\|_{1+p}$ is convex. This implies that the inequality is strict unless $U_1^*U_1$ is a multiple of $U_2^*U_2$. Since both $U_1^*U_1$ and $U_2^*U_2$ are positive definite matrices, this multiple must be a positive number. Finally, by convexity,

$$\left(\frac{1}{2}\|U_1^*U_1\|_{1+p}+\frac{1}{2}\|U_2^*U_2\|_{1+p}\right)^{1+p}\leqslant\frac{1}{2}\|U_1^*U_1\|_{1+p}^{1+p}+\frac{1}{2}\|U_2^*U_2\|_{1+p}^{1+p}$$

with a strict inequality unless $\|U_1^*U_1\|_{1+p}=\|U_2^*U_2\|_{1+p}$. Thus equality implies that the multiple above equals 1 and

$$U_1^*U_1=U_2^*U_2. \tag{2.2}$$

In the case of the present proposition we have that $U_i=U(Z_i)=Y_1^{-1/2}(Z_i-iI)$, $i=1,2$ and that P projects onto

$$\operatorname{Ran}\begin{bmatrix}Y_1^{1/2}\\ Y_2^{1/2}\end{bmatrix}=\operatorname{Ran}\begin{bmatrix}I\\ Y_2^{1/2}Y_1^{-1/2}\end{bmatrix}.$$

The equality holds since $Y_1^{1/2}$ is invertible for $Z_1\in\mathbb{SH}_2$. Now the range condition

$$\operatorname{Ran}\begin{bmatrix}U(Z_1)\\ U(Z_2)\end{bmatrix}\subseteq\operatorname{Ran}\begin{bmatrix}I\\ Y_2^{1/2}Y_1^{-1/2}\end{bmatrix}$$

is equivalent to $U(Z_2)=Y_2^{1/2}Y_1^{-1/2}U(Z_1)$ or $X_2+i(I-Y_2^{-1})=X_1+i(I-Y_1^{-1})$. Equating real and imaginary parts, this implies $Z_1=Z_2$. □

Notice that we did not use (2.2) in the proof, but it will be important later.

The following function will be used below

$$R(t,\epsilon)=t/2+\sqrt{t^2/4+\epsilon^2}.$$

Its asymptotics when $\epsilon\to 0$ and $t\to t_0$ depend on the sign of t_0:

$$R(t,\epsilon)\begin{cases}=t+O(\epsilon^2) & \text{if } t_0>0,\\ \to 0 & \text{if } t_0=0,\\ =\epsilon^2/|t|+O(\epsilon^4) & \text{if } t_0<0.\end{cases} \tag{2.3}$$

We will also need the fact that if $\epsilon\to 0$ and $t_1\to 0$ and $t_2\to t_0<0$ then

$$R(t_2,\epsilon)/R(t_1,\epsilon)\to 0. \tag{2.4}$$

This follows from $1/R(t,\epsilon)=\epsilon^{-2}(R(t,\epsilon)-t)$.

Let $Z=X+iY\in\mathbb{SH}_2$ and $U(Z)=Y^{-1/2}(Z-iI)$. Here are some facts that we need. Write

$$U(Z)=\epsilon^{-1}(S+iT)$$

where $\epsilon=1/\|U(Z)\|_{2(1+p)}$ and $\|S+iT\|_{2(1+p)}=1$. Then

$$Y^{1/2} = \epsilon^{-1} R(T, \epsilon),$$
$$X = \epsilon^{-1} Y^{1/2} S = \epsilon^{-2} R(T, \epsilon) S,$$
$$T = \epsilon\big(Y^{1/2} - Y^{-1/2}\big) = R(T, \epsilon) - \epsilon^2 R(T, \epsilon)^{-1},$$
$$S = \epsilon Y^{-1/2} X.$$

Notice that T is a real symmetric matrix, but not necessarily positive definite. The matrix S need not be symmetric, but $R(T, \epsilon)S$ is.

Proof of Proposition 1.7. We are given sequences $(Z_{1,n}, Z_{2,n}) \to (Z_1, Z_2) \in \overline{\mathbb{SH}_2} \times \overline{\mathbb{SH}_2}$ with $\mu^*_{2,p}(Z_{1,n}, Z_{2,n}) \to 1$. Let $U_{k,n} = U(Z_{k,n}) = \epsilon^{-1}_{k,n}(S_{k,n} + iT_{k,n})$, $k = 1, 2$ and define P_n to be the rank 2 projection onto the range of

$$\operatorname{Ran}\begin{bmatrix} Y^{1/2}_{1,n} \\ Y^{1/2}_{2,n} \end{bmatrix} = \operatorname{Ran}\begin{bmatrix} \epsilon^{-1}_{1,n} R(T_{1,n}, \epsilon_{1,n}) \\ \epsilon^{-1}_{2,n} R(T_{2,n}, \epsilon_{2,n}) \end{bmatrix}^t.$$

Then

$$\mu^*_{2,p}(Z_{1,n}, Z_{2,n}) = \left\| \frac{1}{2}\big[r_{1,n}\big(S^t_{1,n} - iT_{1,n}\big), r_{2,n}\big(S^t_{2,n} - iT_{2,n}\big)\big] P_n \begin{bmatrix} r_{1,n}(S_{1,n} + iT_{1,n}) \\ r_{2,n}(S_{2,n} + iT_{2,n}) \end{bmatrix} \right\|^{1+p}_{1+p}$$

with

$$r^{2(1+p)}_{1,n} = \frac{2\epsilon^{-2(1+p)}_{2,n}}{\epsilon^{-2(1+p)}_{1,n} + \epsilon^{-2(1+p)}_{2,n}}, \qquad r^{2(1+p)}_{2,n} = \frac{2\epsilon^{-2(1+p)}_{1,n}}{\epsilon^{-2(1+p)}_{1,n} + \epsilon^{-2(1+p)}_{2,n}},$$

so that $r^{2(1+p)}_{1,n} + r^{2(1+p)}_{2,n} = 2$. By going to a subsequence we may assume that

$$S_{k,n} + iT_{k,n} \to S_k + iT_k, \quad k = 1, 2,$$
$$r_{k,n} \to r_k, \quad k = 1, 2,$$
$$P_n \to P$$

since these quantities vary in compact sets. Now every term in the expression for $\mu^*_{2,p}$ converges, so that

$$\left\| \frac{1}{2}\big[r_1\big(S^t_1 - iT_1\big), r_2\big(S^t_2 - iT_2\big)\big] P \begin{bmatrix} r_1(S_1 + iT_1) \\ r_2(S_2 + iT_2) \end{bmatrix} \right\|^{1+p}_{1+p} = 1.$$

Given this equality we can follow the reasoning in the proof of Proposition 1.4 to conclude that (2.1) and (2.2) hold when U_1 and U_2 in those equations are replaced by $r_1(S^t_1 - iT_1)$ and $r_2(S^t_2 - iT_2)$. After this replacement (2.2) implies $r_1 = r_2 = 1$. Thus by (2.1) we find that

$$\operatorname{Ran}\begin{bmatrix} S_1 + iT_1 \\ S_2 + iT_2 \end{bmatrix} \subseteq \operatorname{Ran} P \quad \text{or} \quad P\begin{bmatrix} S_1 + iT_1 \\ S_2 + iT_2 \end{bmatrix} = \begin{bmatrix} S_1 + iT_1 \\ S_2 + iT_2 \end{bmatrix}. \tag{2.5}$$

The equality $r_1 = r_2 = 1$ also implies that $\epsilon_{1,n}/\epsilon_{2,n} \to 1$ and

$$\big(S_1^t - iT_1\big)(S_1 + iT_1) = \big(S_2^t - iT_2\big)(S_2 + iT_2). \tag{2.6}$$

If the common limit for $\epsilon_{1,n}$ and $\epsilon_{2,n}$ is non-zero, then $Z_{k,n}$, $k = 1, 2$ converge to points in the interior of $\mathbb{SH}_2$. In this case the conclusion of the proposition follows from Proposition 1.4. Thus we may assume that $\epsilon_{k,n} \to 0$, $k = 1, 2$.

Let $Z_{a,n} = (Z_{1,n} + Z_{2,n})/2$ and define $X_{a,n}$, $Y_{a,n}$, $U_{a,n}$, $\epsilon_{a,n}$, $S_{a,n}$ and $T_{a,n}$ and their limiting values as above. Then a calculation shows that

$$U_{a,n}^* U_{a,n} = \frac{1}{2}\big[U_{1,n}^* U_{2,n}^*\big] P_n \begin{bmatrix} U_{1,n} \\ U_{2,n} \end{bmatrix}.$$

Taking norms, this implies that

$$\epsilon_{a,n}^{-2(1+p)} = \mu_{2,p}^*(Z_{1,n}, Z_{2,n}) \frac{1}{2}\big(\epsilon_{1,n}^{-2(1+p)} + \epsilon_{2,n}^{-2(1+p)}\big).$$

Since we are assuming that $\mu_{2,p}^*(Z_{1,n}, Z_{2,n}) \to 1$, this implies that $\epsilon_{a,n}/\epsilon_{k,n} \to 1, k = 1, 2$. In particular, $\epsilon_{a,n} \to 0$. This means that the average point $Z_{a,n}$ is moving to infinity, that is, possible cancellations in the sum $Z_{1,n} + Z_{2,n}$ that would keep $Z_{a,n}$ finite do not occur.

We will use that

$$\begin{aligned} T_{a,n} &= \epsilon_{a,n}\left(\frac{Y_{1,n} + Y_{2,n}}{2}\right)^{1/2} - \epsilon_{a,n}\left(\frac{Y_{1,n} + Y_{2,n}}{2}\right)^{-1/2} \\ &= \frac{1}{\sqrt{2}}\left(\left(\frac{\epsilon_{a,n}}{\epsilon_{1,n}}\right)^2 R(T_{1,n}, \epsilon_{1,n})^2 + \left(\frac{\epsilon_{a,n}}{\epsilon_{2,n}}\right)^2 R(T_{2,n}, \epsilon_{2,n})^2\right)^{1/2} \\ &\quad - \epsilon_{a,n}^2 \sqrt{2}\left(\left(\frac{\epsilon_{a,n}}{\epsilon_{1,n}}\right)^2 R(T_{1,n}, \epsilon_{1,n})^2 + \left(\frac{\epsilon_{a,n}}{\epsilon_{2,n}}\right)^2 R(T_{2,n}, \epsilon_{2,n})^2\right)^{-1/2}. \end{aligned} \tag{2.7}$$

Beginning with $T_a = \epsilon_a Y_a^{-1/2}(Y_a - I)$ we also compute that

$$T_{a,n}^2 = \frac{1}{2}\big[T_{1,n}^* T_{2,n}^*\big] P_n \begin{bmatrix} T_{1,n} \\ T_{2,n} \end{bmatrix}.$$

Then taking account of the imaginary part of (2.5) we find that in the limit

$$T_a^2 = \frac{1}{2}\big(T_1^2 + T_2^2\big), \tag{2.8}$$

which is not immediately apparent from (2.7). Similarly

$$S_a^t S_a = \frac{1}{2}\big(S_1^t S_1 + S_2^t S_2\big) \tag{2.9}$$

and

$$(S_a + iT_a)^*(S_a + iT_a) = \frac{1}{2}\big((S_1 + iT_1)^*(S_1 + iT_1) + (S_2 + iT_2)^*(S_2 + iT_2)\big). \tag{2.10}$$

The points corresponding to $Z_{k,n}$, $k = 1, 2, a$ in the disk model are given by

$$W_{k,n} = (Z_{k,n} + iI)^{-1}(Z_{k,n} - iI) = \Big(S_{k,n} + i\sqrt{T_{k,n}^2 + 4\epsilon_{k,n}^2}\Big)^{-1}(S_{k,n} + iT_{k,n}). \tag{2.11}$$

Our task is to show that the limiting values satisfy either (i) $W_1 = W_2 = W_a$ or the relations described in one of part (ii) or (iii) of the proposition.

We will break our analysis into cases depending on the eigenvalues of the real symmetric 2×2 matrices T_1 and T_2. Let t_1 and t_2 be the eigenvalues of T_1 and τ_1, τ_2 be the eigenvalues of T_2. For T_1 we have 6 cases which we will label ++, +0, +−, 00, 0−, −− depending on whether t_1 and t_2 are positive, zero or negative. Pairing the possibilities for T_1 and T_2 and taking account of symmetry leaves 21 cases to consider.

Case ++ ++

In this case T_1 and T_2 and, by (2.7), also T_a are positive definite. So (2.11) implies that $W_{1,n}$, $W_{2,n}$ and $W_{a,n}$ all converge to I. So (i) holds.

Cases ++ +0, ++ +−, ++ 00, ++ 0− *and* ++ −−

In these cases, using (2.3), we have $\lim_{n\to\infty} R(T_{1,n}, \epsilon_{1,n}) = T_1$ and we see that the limit of $\left[\begin{smallmatrix} R(T_{1,n},\epsilon_{1,n}) \\ R(T_{2,n},\epsilon_{2,n}) \end{smallmatrix}\right]$ has the form $\left[\begin{smallmatrix} T_1 \\ B \end{smallmatrix}\right]$ where $B = \lim_{n\to\infty} R(T_{2,n}, \epsilon_{2,n})$. By assumption T_1 is invertible, hence $\operatorname{Ran}\left[\begin{smallmatrix} T_1 \\ B \end{smallmatrix}\right]$ is two-dimensional, and hence equal to $\operatorname{Ran} P$. From (2.5) we may deduce that $\operatorname{Ran}(S_2 + iT_2) \subseteq \operatorname{Ran} B$. Referring again to (2.3) we see that $\operatorname{Ran} B$ is less than two-dimensional, so that $S_2 + iT_2$ has rank less than two. On the other hand $S_1 + iT_1$ is invertible. This contradicts (2.6). Therefore these cases do not occur.

Case +0 +0

By (2.6) $S_1 + iT_1$ and $S_2 + iT_2$ are either both invertible or both not invertible. If they are both invertible, then, since $\lim_{n\to\infty} S_{k,n} + i\sqrt{T_{k,n}^2 + 4\epsilon_{k,n}^2} = S_k + iT_k$ for $k = 1, 2$ we see from (2.11) that $W_1 = W_2 = I$. From (2.10) we see that $(S_a + iT_a)$ is invertible. Also, from (2.7) we can conclude that $T_a \geqslant 0$. Then (2.11) implies that $W_a = I$, too.

Now we must consider the case where $S_1 + iT_1$ and $S_2 + iT_2$ are both not invertible. First we show that T_1 and T_2 have the same eigenvectors. We argue by contradiction. Suppose the eigenvector of T_1 corresponding to its positive eigenvalue is different from that of T_2. Then the limit $\lim_{n\to\infty}\left[\begin{smallmatrix} R(T_{1,n},\epsilon_{1,n}) \\ R(T_{2,n},\epsilon_{2,n}) \end{smallmatrix}\right] = \left[\begin{smallmatrix} T_1 \\ T_2 \end{smallmatrix}\right]$ has rank 2, which implies that P is the projection onto its range. Thus (2.1) implies that $\operatorname{Ran}\left[\begin{smallmatrix} S_1+iT_1 \\ S_2+iT_2 \end{smallmatrix}\right] \subseteq \operatorname{Ran}\left[\begin{smallmatrix} T_1 \\ T_2 \end{smallmatrix}\right]$.

For the moment, let us focus on S_1 and T_1. Denote the projections onto the positive and zero eigenvectors for T_1 by P_+ and P_0. The range condition above implies that $\operatorname{Ran} S_1 \subseteq \operatorname{Ran} T_1$ which implies that $\operatorname{Ran} P_0 S_1 \subseteq \operatorname{Ran} P_0 T_1 = 0$. So $P_0 S_1 = 0$. In addition, we know that $R(T_{1,n}, \epsilon_{1,n}) S_{1,n}$ is symmetric, so taking limits, we find that $T_1 S_1 = S_1^t T_1$. This implies that $P_+ S P_0 = 0$. Taken together, these equalities show that $S_1 = P_+ S_1 P_+$. Now we can deduce that $\operatorname{Ran} P_0 \subseteq$

$\mathrm{Ker}(S_1^T - iT_1)(S_1 + iT_1)$. In fact, we must have equality: $\mathrm{Ker}(S_1^T - iT_1)(S_1 + iT_1)$ cannot be more than one-dimensional because, lying on the unit sphere, $(S_1 + iT_1) \neq 0$. So $\mathrm{Ker}(S_1^T - iT_1)(S_1 + iT_1) = \mathrm{Ker}\, T_1$.

Now an analogous argument shows that $\mathrm{Ker}(S_2^T - iT_2)(S_2 + iT_2) = \mathrm{Ker}\, T_2$. We are assuming that $\mathrm{Ker}\, T_1 \neq \mathrm{Ker}\, T_2$. However, (2.6) implies $\mathrm{Ker}(S_1^T - iT_1)(S_1 + iT_1) = \mathrm{Ker}(S_2^T - iT_2)(S_2 + iT_2)$. This contradiction proves our claim that the eigenvectors of T_1 and T_2 are the same.

Now we focus again on $S_{1,n} + iT_{1,n}$ and compute the limiting value of $W_{1,n}$. To simplify notation slightly, we drop the subscript 1. Let $t_{1,n}, t_{2,n}$ be the eigenvalues of T_n, and let V_n be the real orthogonal matrix whose columns are the eigenvectors of T_n. For the case we are considering $t_{1,n} \to t_1 > 0$ and $t_{2,n} \to 0$. Clearly

$$T_n = V_n \begin{bmatrix} t_{1,n} & 0 \\ 0 & t_{2,n} \end{bmatrix} V_n^t. \tag{2.12}$$

The symmetry of $R(T_n, \epsilon_n)S_n$ implies that

$$S_n = V_n \begin{bmatrix} s_{1,1,n} & R(t_{2,n}, \epsilon_{1,n})s_{1,2,n}/R(t_{1,n}, \epsilon_{1,n}) \\ s_{1,2,n} & s_{2,2,n} \end{bmatrix} V_n^t. \tag{2.13}$$

Since the limit $S + iT$ is not invertible we have $s_{2,2,n} \to 0$. With this notation, the expression for W_n is

$$\begin{aligned} W_n = V_n &\begin{bmatrix} s_{1,1,n} + i\sqrt{t_{1,n}^2 + 4\epsilon_n^2} & R(t_{2,n}, \epsilon_n)s_{1,2,n}/R(t_{1,n}, \epsilon_n) \\ s_{1,2,n} & s_{2,2,n} + i\sqrt{t_{2,n}^2 + 4\epsilon_n^2} \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} s_{1,1,n} + it_{1,n} & R(t_{2,n}, \epsilon_n)s_{1,2,n}/R(t_{1,n}, \epsilon_n) \\ s_{1,2,n} & s_{2,2,n} + it_{2,n} \end{bmatrix} V_n^t. \end{aligned}$$

Now we can compute the $(1, 1)$ entry of $V_n^t W_n V_n$ explicitly, yielding

$$\frac{(s_{2,2,n} + i\sqrt{t_{2,n}^2 + 4\epsilon_n^2})(s_{1,1,n} + it_{1,n}) - R(t_{2,n}, \epsilon_n)s_{1,2,n}^2/R(t_{1,n}, \epsilon_n)}{(s_{2,2,n} + i\sqrt{t_{2,n}^2 + 4\epsilon_n^2})(s_{1,1,n} + i\sqrt{t_{1,n}^2 + 4\epsilon_n^2}) - R(t_{2,n}, \epsilon_n)s_{1,2,n}^2/R(t_{1,n}, \epsilon_n)}.$$

Write $(s_{2,2,n}, t_{2,n}, \epsilon_n) = r_n(\omega_{1,n}, \omega_{2,n}, \omega_{3,n})$ with $\omega_{1,n}^2 + \omega_{2,n}^2 + \omega_{3,n}^2 = 1$. Then $r_n \to 0$ and, by going to a subsequence if needed, we may assume that the $\omega_{k,n} \to \omega_k$, $k = 1, 2, 3$. The numerator and denominator of the expression above converge to the same value, namely,

$$\left(\omega_1 + i\sqrt{\omega_2^2 + 4\omega_3^2}\right)(s_{1,1} + it_1) - R(\omega_2, \omega_3)s_{1,2}^2/t_1.$$

We claim that this value cannot be zero. If it is, then calculating the real and imaginary parts yields

$$\begin{aligned} \omega_1 s_{1,1} - t_1\sqrt{\omega_2^2 + 4\omega_3^2} - R(\omega_2, \omega_3)s_{1,2}^2/t_1 = 0, \\ s_{1,1}\sqrt{\omega_2^2 + 4\omega_3^2} + \omega_1 t_1 = 0. \end{aligned}$$

Recall that $t_1 > 0$ and $R(\omega_2, \omega_3) \geqslant 0$. The second equation implies that each term in the first equation is non-positive, and thus must be zero separately. This yields $\omega_2 = \omega_3 = 0$ so $\omega_1 = \pm 1$ and thus $s_{1,1} = 0$. Returning to the expression for the common value of the numerator and denominator, this is now $it_{1,1}$ which is non-zero, contradicting our assumption. We conclude that this common value of the numerator and denominator above is non-zero, and thus the $(1,1)$ entry of the limit $V^t W V$ is 1.

Thus we have shown that

$$W = V \begin{bmatrix} 1 & \beta \\ \beta & \alpha \end{bmatrix} V^t,$$

where we have taken into account that since W is a matrix in the ball model for $\mathbb{SH}_2$, it is symmetric. In addition, we know that $\|W\| \leqslant 1$ so we can conclude that $\beta = 0$. To see this we compute the eigenvalues of $\left[\begin{smallmatrix} 1,\beta \\ \beta,\alpha \end{smallmatrix}\right]^* \left[\begin{smallmatrix} 1,\beta \\ \beta,\alpha \end{smallmatrix}\right]$ explicitly. This yields a value for the larger eigenvalue of

$$\frac{1+|\alpha|^2+2|\beta|^2}{2} + \sqrt{\frac{(1-|\alpha|^2)^2}{4} + |1+\alpha|^2|\beta|^2} \geqslant 1+|\beta|^2.$$

This must be $\leqslant 1$ so $\beta = 0$. Then we must also have $|\alpha| \leqslant 1$ to keep $\|W\| \leqslant 1$.

Re-introducing the subscript 1, this shows that W_1 has the form prescribed in conclusion (ii) of the proposition. The argument for W_2 is the same, and the matrix V, containing eigenvectors for T_1 or T_2 is the same matrix in both cases. Using (2.7) we can see that the matrix $T_a = ((T_1^2 + T_2^2)/2)^{1/2}$ has the same eigenvectors as T_1 and T_2, and also has one positive and one zero eigenvector. So a similar argument shows that W_a also has the form prescribed in (ii) (possibly $W_a = I$ which is a special case of (ii)), again with the same matrix V. This concludes the proof of this case.

Case $+0 \quad +-$

We begin by showing that T_1 and T_2 have the same eigenvectors. To begin, we consider $S_2 + iT_2$ and note that by (2.12) and (2.13) this matrix has the form

$$S_2 + iT_2 = V \begin{bmatrix} \sigma_{1,1} + i\tau_1 & 0 \\ \sigma_{2,1} & \sigma_{2,2} + i\tau_2 \end{bmatrix} V^T,$$

where $\tau_1 > 0$ and $\tau_2 < 0$ are the eigenvalues of T_2. Thus

$$\det\bigl(S_2^t - iT_2\bigr)(S_2 + iT_2) = |\sigma_{1,1} + i\tau_1|^2 |\sigma_{2,2} + i\tau_2|^2 \neq 0$$

so $S_2 + iT_2$ is invertible. By (2.6), $S_1 + iT_1$ is invertible too.

If the eigenvectors of T_1 and T_2 are different, then by (2.3) the limit $\lim_{n\to\infty} \left[\begin{smallmatrix} R(T_{1,n}, \epsilon_{1,n}) \\ R(T_{2,n}, \epsilon_{2,n}) \end{smallmatrix}\right] = \left[\begin{smallmatrix} T_1 \\ T_{2,+} \end{smallmatrix}\right]$, where $T_{2,+}$ is the matrix T_2 projected onto its positive eigenspace. The matrix $\left[\begin{smallmatrix} T_1 \\ T_{2,+} \end{smallmatrix}\right]$ has rank 2 so its range must coincide with the range of P. Then (2.5) implies that $\operatorname{Ran} S_1 + iT_1 \subseteq \operatorname{Ran} T_1$ which is impossible since $S_1 + iT_1$ is invertible and $\dim \operatorname{Ran} T_1 = 1$. Therefore the eigenvectors of T_1 and T_2 are the same. Let V be the orthogonal matrix containing the common eigenvectors.

Since $S_1 + iT_1$ is invertible, we obtain from (2.11) that $W_1 = (S_1 + iT_1)^{-1}(S_1 + iT_1) = I$. Similarly $W_2 = (S_2 + i|T_2|)^{-1}(S_2 + iT_2)$. An explicit computation shows that this has the form $V\left[\begin{smallmatrix}1 & 0\\ 0 & \alpha\end{smallmatrix}\right]V^t$ with $\alpha = (\sigma_{2,2} - i|\tau_2|)/(\sigma_{2,2} + i|\tau_2|)$.

It remains to consider W_a. Using the formula (2.7) and the asymptotics (2.3) we find that $T_a = ((T_1^2 + T_{2,+}^2)/2)^{1/2}$. Thus T_a has one positive and one zero eigenvalue with the same eigenvectors as T_1 and T_2. The arguments from the previous case show that W_a has the form $V\left[\begin{smallmatrix}1 & 0\\ 0 & \alpha\end{smallmatrix}\right]V^t$ with $|\alpha| \leqslant 1$.

Case +0 00

In this case, $T_2 = 0$ so by (2.6) $S_1 + iT_1$ and S_2 are either both invertible or both not invertible. If they are both invertible, then by (2.11) $W_1 = W_2 = I$. By (2.7) $T_a = T_1/\sqrt{2}$ and therefore has one positive and one zero eigenvalue. Then the argument from case +0 +0 shows that $W_a = V\left[\begin{smallmatrix}1 & 0\\ 0 & \alpha\end{smallmatrix}\right]V^t$ with $|\alpha| \leqslant 1$, where V contains the eigenvectors of T_1.

Now we consider the case where $S_1 + iT_1$ and S_2 are both not invertible.

First we show that $\operatorname{Ker} S_2 = \operatorname{Ker} T_1$. Notice that since $R(T_{1,n}, \epsilon_{1,n})S_{1,n} = S_{1,n}^t R(T_{1,n}, \epsilon_{1,n})$ and $R(T_{1,n}, \epsilon_{1,n}) \to T_1$, upon taking limits we find that $T_1 S_1 = S_1^t T_1$. Thus

$$\big(S_1^t - iT_1\big)(S_1 + iT_1) = S_1^t S_1 + i\big(S_1^t T_1 - T_1 S_1\big) + T_1^2 = S_1^t S_1 + T_1^2.$$

So, by (2.6), if $S_2 v = 0$ then $\|S_1 v\|^2 + \|T_1 v\|^2 = 0$ which implies that $T_1 v = 0$. Thus $\operatorname{Ker} S_2 \subseteq \operatorname{Ker} T_1$. By assumption $\operatorname{Ker} T_1$ has dimension 1, so we must have equality.

The arguments in case +0 +0 now imply that W_1 has the form $V\left[\begin{smallmatrix}1 & 0\\ 0 & \alpha\end{smallmatrix}\right]V^t$ with $|\alpha| \leqslant 1$, where V contains the eigenvectors of T_1. Since $T_a = T_1/\sqrt{2}$, W_a has the same form.

It remains to consider W_2. Let $\tau_{1,n}$ and $\tau_{2,n}$ be the eigenvalues of $T_{2,n}$ which, by assumption, both converge to zero. We will use the notation

$$a_{j,n} = \epsilon_{j,n}^{-1} R(\tau_{j,n}, \epsilon_{j,n}), \quad j = 1, 2.$$

These are the eigenvalues of $Y_2^{1/2}$. Then, since $\left[\begin{smallmatrix}a_{1,n} & 0\\ 0 & a_{2,n}\end{smallmatrix}\right]S_{2,n}$ is a real symmetric matrix, it has real eigenvalues $\tilde{\lambda}_n$ and $\tilde{\delta}_n$ and eigenvectors $\left[\begin{smallmatrix}c_n\\ s_n\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}-s_n\\ c_n\end{smallmatrix}\right]$ where $c_n = \cos(\theta_n)$ and $s_n = \sin(\theta_n)$ for some θ_n. To declutter the notation, we will now drop the subscript n with the understanding that variables are evaluated along a subsequence. We find that

$$S_2 = V_2 \begin{bmatrix} \frac{\tilde{\lambda}c^2 + \tilde{\delta}s^2}{a_1} & \frac{\tilde{\lambda}-\tilde{\delta}}{a_1}cs \\ \frac{\tilde{\lambda}-\tilde{\delta}}{a_2}cs & \frac{\tilde{\lambda}s^2 + \tilde{\delta}c^2}{a_2} \end{bmatrix} V_2^t$$

where V_2 diagonalizes T_2. Then we obtain

$$\begin{aligned} W_2 = V_2 &\begin{bmatrix} \frac{\tilde{\lambda}c^2 + \tilde{\delta}s^2}{a_1} + i\epsilon(a_1 + 1/a_1) & \frac{\tilde{\lambda}-\tilde{\delta}}{a_1}cs \\ \frac{\tilde{\lambda}-\tilde{\delta}}{a_2}cs & \frac{\tilde{\lambda}s^2 + \tilde{\delta}c^2}{a_2} + i\epsilon(a_2 + 1/a_2) \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} \frac{\tilde{\lambda}c^2 + \tilde{\delta}s^2}{a_1} + i\epsilon(a_1 - 1/a_1) & \frac{\tilde{\lambda}-\tilde{\delta}}{a_1}cs \\ \frac{\tilde{\lambda}-\tilde{\delta}}{a_2}cs & \frac{\tilde{\lambda}s^2 + \tilde{\delta}c^2}{a_2} + i\epsilon(a_2 - 1/a_2) \end{bmatrix} V_2^t \end{aligned}$$

$$= V_2 \begin{bmatrix} (\lambda c^2+\delta s^2)+i\epsilon'(a_1^2+1) & (\lambda-\delta)cs \\ (\lambda-\delta)cs & (\lambda s^2+\delta c^2)+i\epsilon'(a_2^2+1) \end{bmatrix}^{-1}$$
$$\times \begin{bmatrix} (\lambda c^2+\delta s^2)+i\epsilon'(a_1^2-1) & (\lambda-\delta)cs \\ (\lambda-\delta)cs & (\lambda s^2+\delta c^2)+i\epsilon'(a_2^2-1) \end{bmatrix} V_2^t$$

where $\lambda = \tilde{\lambda}/a_1$, $\delta = \tilde{\delta}/a_1$, $\epsilon' = \epsilon/a_1$, and we have cancelled a common factor of a_2/a_1 from the bottom row of each matrix. Since we are assuming that S_2 is converging to a rank 1 matrix, we may assume that λ converges to a non-zero finite number and δ converges to zero. Moreover, since not only ϵ but also $\tau_1 = \epsilon(a_1 - 1/a_1)$ converges to zero, we find that $\epsilon' = \epsilon/a_1$ converges to zero too.

Now write $(\delta, \epsilon') = r(\omega_1, \omega_2)$ where $r \to 0$ and $\omega_1^2 + \omega_2^2 = 1$. Going to a subsequence if needed, we may assume that ω_1 and ω_2 converge. Then a lengthy calculation shows that in the limit (the limiting values of a_1 and a_2 could be infinite here) we have

$$W_2 - I = \frac{-2i\omega_2}{\omega_1 + i\omega_2(a_1^2 s^2 + a_2^2 c^2 + 1)} V_2 \begin{bmatrix} s^2 & -cs \\ -cs & c^2 \end{bmatrix} V_2^t.$$

The limiting vector $V_2\left[\begin{smallmatrix} c \\ s \end{smallmatrix}\right]$ is orthogonal to the kernel of S_2. Since $\operatorname{Ker} S_2 = \operatorname{Ker} T_1$, this vector must be the eigenvector of T_1 with positive eigenvalue. Thus $V_2\left[\begin{smallmatrix} s^2 & -cs \\ -cs & c^2 \end{smallmatrix}\right]V_2^t = V\left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]V^t$, where V contains the eigenvectors for T_1. Therefore we may conclude that $W_2 = V\left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right]V^t$ with $|\alpha| \leqslant 1$.

Case `+0 0-`

We will show that this case is not possible.

First, suppose that $(S_1 + iT_1)$ is invertible. Then, by (2.6) $S_2 + iT_2$ is invertible too. Let $V_{1,n}$ be an orthogonal matrix diagonalizing $T_{1,n}$ so that $V_{1,n}^t T_{1,n} V_{1,n} = \left[\begin{smallmatrix} t_{1,n} & 0 \\ 0 & t_{2,n} \end{smallmatrix}\right]$. We will work in the basis where $T_{1,n}$ is diagonal, so let $\tilde{S}_{k,n} + i\tilde{T}_{k,n} = V_{1,n}^t(S_{k,n} + iT_{k,n})V_{1,n}$. To apply (2.5) we need to compute the limit of

$$\operatorname{Ran}\begin{bmatrix} \begin{bmatrix} R(t_{1,n},\epsilon_{1,n}) & 0 \\ 0 & R(t_{2,n},\epsilon_{1,n}) \end{bmatrix} \\ B_n \end{bmatrix} \tag{2.14}$$

where $B = V_n\left[\begin{smallmatrix} R(\tau_{1,n},\epsilon_{2,n}) & 0 \\ 0 & R(\tau_{2,n},\epsilon_{2,n}) \end{smallmatrix}\right]V_n^t$ for some orthogonal V_n. Here $t_{1,n}$ and $t_{2,n}$ are the eigenvalues of $T_{1,n}$ and $\tau_{1,n}$ and $\tau_{2,n}$ are the eigenvalues of $T_{2,n}$. Using (2.3) we find that

$$\begin{bmatrix} \begin{bmatrix} R(t_{1,n},\epsilon_{1,n}) & 0 \\ 0 & R(t_{2,n},\epsilon_{1,n}) \end{bmatrix} \\ B_n \end{bmatrix} \to \begin{bmatrix} t_1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since this matrix has rank 1, the limiting range in (2.14) must be larger. To determine what it can be, we multiply the matrix in (2.14) on the left by $\left[\begin{smallmatrix} 1 & 0 \\ 0 & r_n \end{smallmatrix}\right]$ where r_n is chosen to scale the second column of the matrix in (2.14) to produce a non-zero limit, possibly after going to a subsequence. Multiplying on the right side with an invertible matrix does not change the range. So, using (2.5) we find that

$$\operatorname{Ran}\begin{bmatrix} \tilde{S}_1 + i\tilde{T}_1 \\ \tilde{S}_2 + i\tilde{T}_2 \end{bmatrix} \subseteq \operatorname{Ran}\begin{bmatrix} t_1 & 0 \\ 0 & \omega_1 \\ 0 & \omega_2 \\ 0 & \omega_3 \end{bmatrix}$$

for some ω_1, ω_2 and ω_3. This implies that $\tilde{S}_2 + i\tilde{T}_2$ is not invertible, which contradicts our assumption.

Now we consider the case when $S_1 + iT_1$ and $S_2 + iT_2$ are both not invertible. By (2.6) their kernels are equal. Let V_1 be an orthogonal matrix diagonalizing $T_{1,n}$ so that $V_{1,n}^t T_{1,n} V_{1,n} = \left[\begin{smallmatrix} t_{1,n} & 0 \\ 0 & t_{2,n} \end{smallmatrix}\right]$. As we have seen above, the fact that $R(T_{1,n}, \epsilon_{1_n}) S_{1,n}$ is symmetric together with the fact that $R(t_{2,n}, \epsilon_{1_n})/R(t_{1,n}, \epsilon_{1_n}) \to 0$ imply

$$S_1 + iT_1 = V_1 \begin{bmatrix} s_{1,1} + it_1 & 0 \\ s_{2,1} & s_{2,2} \end{bmatrix} V_1^t = V_1 \begin{bmatrix} s_{1,1} + it_1 & 0 \\ s_{2,1} & 0 \end{bmatrix} V_1^t.$$

We used that since $t_1 > 0$ and $S_1 + iT_1$ is not invertible, we must have $s_{2,2} = 0$. Similarly, the fact that $\tau_2 < 0$ and $\tau_1 = 0$ implies that $R(\tau_{2,n}, \epsilon_{1_n})/R(\tau_{1,n}, \epsilon_{1_n}) \to 0$ so we can conclude that

$$S_2 + iT_2 = V_2 \begin{bmatrix} \sigma_{1,1} & 0 \\ \sigma_{2,1} & \sigma_{2,2} + i\tau_2 \end{bmatrix} V_2^t = V_2 \begin{bmatrix} 0 & 0 \\ \sigma_{2,1} & \sigma_{2,2} + i\tau_2 \end{bmatrix} V_2^t,$$

since $S_2 + iT_2$ is not invertible either. Now we invoke the fact that $S_1 + iT_1$ and $S_2 + iT_2$ have the same kernel. This implies that

$$V_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = V_2 \frac{1}{\sqrt{\sigma_{2,1}^2 + \sigma_{2,2}^2 + \tau_2^2}} \begin{bmatrix} \sigma_{2,2} + i\tau_2 \\ -\sigma_{2,1} \end{bmatrix}.$$

Write $V_1^{-1} V_2 = \left[\begin{smallmatrix} c & s \\ -s & c \end{smallmatrix}\right]$ where $c = \cos\theta$ and $s = \sin\theta$ for some θ. Then, the first line of the previous matrix equation reads

$$c(\sigma_{2,2} + i\tau_2) + s\sigma_{2,1} = 0.$$

Since $\tau_2 < 0$ the imaginary part of this equation implies $c = 0$. Since $c^2 + s^2 = 1$, this implies $s = \pm 1$ and thus $\sigma_{2,1} = 0$. Therefore

$$S_2 + iT_2 = V_1 \begin{bmatrix} 0 & \pm 1 \\ \mp 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \sigma_{2,2} + i\tau_2 \end{bmatrix} \begin{bmatrix} 0 & \mp 1 \\ \pm 1 & 0 \end{bmatrix} V_1^t = V_1 \begin{bmatrix} \sigma_{2,2} + i\tau_2 & 0 \\ 0 & 0 \end{bmatrix} V_1^t.$$

Now we turn to (2.5). We conjugate all the matrices with $V_{1,n}$, that is, we work in the basis where $T_{1,n}$ is diagonal. Then we find

$$\operatorname{Ran}\begin{bmatrix} s_1 1 + it_1 & 0 \\ s_{2,1} & 0 \\ \sigma_{2,2} + i\tau_2 & 0 \\ 0 & 0 \end{bmatrix} \subseteq \lim \operatorname{Ran}\begin{bmatrix} \begin{bmatrix} R(t_{1,n}, \epsilon_{1,n}) & 0 \\ 0 & R(t_{2,n}, \epsilon_{1,n}) \end{bmatrix} \\ V_n \begin{bmatrix} R(\tau_{1,n}, \epsilon_{2,n}) & 0 \\ 0 & R(\tau_{2,n}, \epsilon_{2,n}) \end{bmatrix} V_n^t \end{bmatrix}$$

where $V_n = \left[\begin{smallmatrix} c_n & -s_n \\ s_n & c_n \end{smallmatrix}\right]$ with $c_n \to 0$, $s_n \to \pm 1$. Write $(R(t_{2,n}, \epsilon_{1,n}), R(\tau_{1,n}, \epsilon_{2,n})) = \delta_n(\omega_{1,n}, \omega_{2,n})$ with $\delta_n \to 0$ and $(\omega_{1,n}, \omega_{2,n}) \to (\omega_1, \omega_2)$ and $\omega_{1,n}^2 + \omega_{2,n}^2 = 1$. Now multiply the matrix on the right side of the previous equation with $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1/\delta_n \end{smallmatrix}\right]$. This leaves the range unchanged, so the limit on the right is the limiting range of

$$\begin{bmatrix} R(t_{1,n}, \epsilon_{1,n}) & 0 \\ 0 & \omega_{1,n} \\ R(\tau_{1,n}, \epsilon_{2,n})c_n^2 + R(\tau_{2,n}, \epsilon_{2,n})s_n^2 & \omega_{2,n}s_nc_n - R(\tau_{2,n}, \epsilon_{2,n})s_nc_n/\delta_n \\ R(\tau_{1,n}, \epsilon_{2,n})s_nc_n - R(\tau_{2,n}, \epsilon_{2,n})s_nc_n & \omega_{2,n}s_n^2 + R(\tau_{2,n}, \epsilon_{2,n})c_n^2/\delta_n \end{bmatrix}.$$

This limiting range will be the span of the limiting values of the columns, provided these are linearly independent. Using $R(\tau_{2,n}, \epsilon_{2,n})/\delta_n \to 0$, we see that this is true, and therefore

$$\operatorname{Ran}\begin{bmatrix} s_1 1 + it_1 & 0 \\ s_{2,1} & 0 \\ \sigma_{2,2} + i\tau_2 & 0 \\ 0 & 0 \end{bmatrix} \subseteq \operatorname{Ran}\begin{bmatrix} t_1 & 0 \\ 0 & \omega_1 \\ 0 & 0 \\ 0 & \omega_2 \end{bmatrix}.$$

But this is impossible because $\tau_2 < 0$.

Case `+0 --`

In this case the limiting range of $\left[\begin{smallmatrix} R(T_{1,n}, \epsilon_{1,n}) \\ R(T_{2,n}, \epsilon_{2,n}) \end{smallmatrix}\right]$ is the range of a matrix of the form $\left[\begin{smallmatrix} A \\ 0 \end{smallmatrix}\right]$ for some invertible 2×2 matrix A. This follows from the asymptotics (2.3) which imply that the eigenvalues of $R(T_{2,n}, \epsilon_{2,n})$ tend to zero much more quickly than those of $R(T_{1,n}, \epsilon_{1,n})$. Thus (2.5) implies $S_2 + iT_2 = 0$ which is not possible. So this case does not occur.

Case `00 00`

If S_1 is invertible, then, since $T_1 = T_2 = 0$, (2.6) and (2.9) imply that S_2 and S_a are invertible too. Then formula (2.11) shows that $W_1 = W_2 = W_a = I$.

If S_1 is not invertible, then (2.6) and (2.9) show that S_1, S_2 and S_a have the same kernel. Following the computation of W_2 in the case `0+ 00`, we see that for the present case, W_1, W_2 and W_a each have the form $V\left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right]V^t$ with $|\alpha| \leqslant 1$, where in each case V contains the common eigenvectors of $S_1^tS_1$, $S_2^tS_2$ and $S_a^tS_a$.

Case `00 0-`

If S_1 and $S_2 + iT_2$ are both invertible then, starting with (2.5) and possibly rescaling the limit on the right, we will end up with

$$\operatorname{Ran}\begin{bmatrix} S_1 \\ S_2 + iT_2 \end{bmatrix} \subseteq \operatorname{Ran}\begin{bmatrix} A \\ B \end{bmatrix}$$

where A and B are invertible matrices with real entries. Since the ranges are unchanged under multiplication on the right by invertible matrices, this is equivalent to

$$\operatorname{Ran}\begin{bmatrix} I \\ (S_2 + iT_2)S_1^{-1} \end{bmatrix} \subseteq \operatorname{Ran}\begin{bmatrix} I \\ BA^{-1} \end{bmatrix}$$

which implies that $S_2S_1^{-1} + iT_2S_1^{-1} = BA^{-1}$. Taking the imaginary part of this equation yields $T_2S_1^{-1} = 0$ which implies $T_2 = 0$, since S_1 is invertible. But $T_2 \neq 0$ so this is impossible.

Now suppose that S_1 and $S_2 + iT_2$ are both not invertible. From (2.6) they have a common kernel, which must be one-dimensional. If this kernel is spanned by v then, since S_1 is a real matrix and $S_1v = 0$, we may assume that v has real entries too. Then $S_2v + iT_2v = 0$ implies, by taking real and imaginary parts, that $S_2v = 0$ and $T_2v = 0$. If V_2 is an orthogonal matrix diagonalizing T_2, we have $T_2 = V_2\left[\begin{smallmatrix}0 & 0\\ 0 & t_2\end{smallmatrix}\right]V_2^t$. Thus, $v = T_2\left[\begin{smallmatrix}1\\ 0\end{smallmatrix}\right]$. Now, it follows that $S_1 = V_2\left[\begin{smallmatrix}0 & s_{1,2}\\ 0 & s_{2,2}\end{smallmatrix}\right]V_2^t$ and $S_2 + iT_2 = V_2\left[\begin{smallmatrix}0 & \sigma_{1,2}\\ 0 & \sigma_{2,2}+i\tau_2\end{smallmatrix}\right]V_2^t$. So, starting with (2.5) and conjugating with V_2 we obtain

$$\begin{bmatrix} s_{1,2} \\ s_{2,2} \\ \sigma_{1,2} \\ \sigma_{2,2} + i\tau_2 \end{bmatrix} \in \lim \operatorname{Ran} \begin{bmatrix} V_n \left[\begin{smallmatrix} R(t_{1,n},\epsilon_{1,n}) & 0 \\ 0 & R(t_{2,n},\epsilon_{1,n}) \end{smallmatrix}\right] V_n^t \\ \left[\begin{smallmatrix} R(\tau_{1,n},\epsilon_{2,n}) & 0 \\ 0 & R(\tau_{2,n},\epsilon_{2,n}) \end{smallmatrix}\right] \end{bmatrix} \tag{2.15}$$

where $V_n = V_{2,n}^{-1}V_{1,n} = \left[\begin{smallmatrix} c_n & -s_n \\ s_n & c_n \end{smallmatrix}\right]$ for some $c_n = \cos(\theta_n)$ and $s_n = \sin(\theta_n)$. Going to a subsequence if needed, we assume that c_n and s_n converge. To simplify notation, drop the n subscript and let $R_1 = R(t_{1,n}, \epsilon_{1,n})$, $R_2 = R(t_{2,n}, \epsilon_{1,n})$, $R_3 = R(\tau_{1,n}, \epsilon_{2,n})$, and $R_4 = R(\tau_{2,n}, \epsilon_{2,n})$. With this notation we need to find the limiting range of

$$B = \begin{bmatrix} R_1c^2 + R_2s^2 & (R_1 - R_2)sc \\ (R_1 - R_2)sc & R_1s^2 + R_2c^2 \\ R_3 & 0 \\ 0 & R_4 \end{bmatrix}.$$

Let $\delta_1 = \sqrt{R_1^2c^2 + R_2^2s^2 + R_3^2}$ and $\delta_2 = \sqrt{R_1^2s^2 + R_2^2c^2 + R_4^2}$ be the Euclidean norms of the columns of B. If $\lim R_3/\delta_1 > 0$, then $B\left[\begin{smallmatrix} 1/\delta_1 & 0 \\ 0 & 1/\delta_2 \end{smallmatrix}\right]$ converges to a matrix of the form

$$\begin{bmatrix} * & * \\ * & * \\ + & 0 \\ 0 & 0 \end{bmatrix}$$

where $+$ denotes a positive entry and $*$ is an arbitrary entry and each column has Euclidean norm equal to 1. Here we used that $R_4/\delta_2 \to 0$, which follows from the estimate $R_4^2/\delta_2^2 \leqslant 2R_4^2/R_k^2$ for k either 1 or 2 and the fact that $R_4/R_k \to 0$. The matrix above has rank 2, and thus its range must be the same as the limiting range on the right side of (2.15). Now, given (2.15), the fact that both entries in the last row are zero contradicts $\tau_2 < 0$.

Thus we must have $\lim R_3/\delta_1 = 0$ which implies that either $R_3/(R_1c) \to 0$ or $R_3/(R_2s) \to 0$. (It could be that one or the other of these sequences is undefined, if c or s is identically zero along the sequence.) If $R_3/(R_1c) \to 0$ we compute the limiting value of $BV\left[\begin{smallmatrix} 1/R_1 & 0 \\ 0 & 1/\sqrt{R_2^2+s^2R_3^2} \end{smallmatrix}\right]$ and find that this has the form

$$\begin{bmatrix} 1 & 0 \\ 0 & R_2/\sqrt{R_2^2+s^2R_3^2} \\ cR_3/R_1 & sR_3/\sqrt{R_2^2+s^2R_3^2} \\ -sR_4/R_1 & cR_4/\sqrt{R_2^2+s^2R_3^2} \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & * \\ 0 & * \\ 0 & 0 \end{bmatrix}$$

where the second column has Euclidean norm equal to 1. As above, this contradicts (2.15). Finally, if $R_3/(R_2 s) \to 0$ we compute the limiting value of $BV\left[\begin{smallmatrix} 1/\sqrt{R_1^2+c^2R_3^2} & 0 \\ 0 & 1/R_2 \end{smallmatrix}\right]$ and find that this has the form

$$\begin{bmatrix} R_1/\sqrt{R_1^2+c^2R_3^2} & 0 \\ 0 & 1 \\ cR_3/\sqrt{R_1^2+c^2R_3^2} & sR_3/R_2 \\ -sR_4/\sqrt{R_1^2+c^2R_3^2} & cR_4/R_2 \end{bmatrix} \to \begin{bmatrix} * & 0 \\ 0 & 1 \\ * & 0 \\ 0 & 0 \end{bmatrix}$$

where the first column has Euclidean norm equal to 1. Again this contradicts (2.15).

In conclusion, we see that this case is not possible.

Case `00 --`

This case is analogous to `++ 00` and is not possible.

Case `0- 0-`

Let V_1 and V_2 be orthogonal matrices diagonalizing T_1 and T_2 respectively. By switching the sign of a column, if needed, we may assume that V_1 and V_2 are rotation matrices. We will show that they are equal. Using (2.5) we write

$$\begin{bmatrix} S_1+iT_1 \\ S_2+iT_2 \end{bmatrix} \in \lim \operatorname{Ran} \begin{bmatrix} V_1 \left[\begin{smallmatrix} R_1 & 0 \\ 0 & R_2 \end{smallmatrix}\right] V_1^t \\ V_2 \left[\begin{smallmatrix} R_3 & 0 \\ 0 & R_4 \end{smallmatrix}\right] V_2^T \end{bmatrix}$$

where the quantities on the right are being evaluated along a subsequence where V_1 and V_2 converge. As before, $R_1 = R(t_{1,n}, \epsilon_{1,n})$, $R_2 = R(t_{2,n}, \epsilon_{1,n})$, $R_3 = R(\tau_{1,n}, \epsilon_{2,n})$, and $R_4 = R(\tau_{2,n}, \epsilon_{2,n})$. Going to a subsequence we assume that R_2/R_4 converges, and by switching the roles of R_2 and R_4 if needed, that $\lim R_2/R_4 = a < \infty$. Notice that $a \geqslant 0$. Let $V = V_2^t V_1 = \left[\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right]$, where $c = \cos(\theta)$ and $s = \sin(\theta)$ for some θ. We now conjugate by V_2 to work in a basis where T_2 is diagonal. Then we find

$$\begin{aligned} \begin{bmatrix} V_2(S_1+iT_1)V_2^t \\ V_2(S_2+iT_2)V_2^t \end{bmatrix} &\in \lim \operatorname{Ran} \begin{bmatrix} V \left[\begin{smallmatrix} R_1 & 0 \\ 0 & R_2 \end{smallmatrix}\right] V^t \\ \left[\begin{smallmatrix} R_3 & 0 \\ 0 & R_4 \end{smallmatrix}\right] \end{bmatrix} \\ &= \lim \operatorname{Ran} \begin{bmatrix} V \left[\begin{smallmatrix} R_1 & 0 \\ 0 & R_2 \end{smallmatrix}\right] \\ \left[\begin{smallmatrix} R_3 & 0 \\ 0 & R_4 \end{smallmatrix}\right] V \end{bmatrix} \end{aligned}$$

$$= \lim \operatorname{Ran} \begin{bmatrix} R_1 c & -R_2 s \\ R_1 s & R_2 c \\ R_3 c & -R_3 s \\ R_4 s & R_4 c \end{bmatrix} \begin{bmatrix} R_1^{-1} & 0 \\ 0 & \epsilon^{-2} \end{bmatrix}$$

$$= \lim \operatorname{Ran} \begin{bmatrix} c & -s/|t_2| \\ s & c/|t_2| \\ ac & -R_3 s/\epsilon^2 \\ 0 & c/|\tau_2| \end{bmatrix}.$$

Suppose $\lim R_3 s/\epsilon^2 = \infty$. Then the limiting range on the right is equal to the range of

$$\begin{bmatrix} c & 0 \\ s & 0 \\ ac & 1 \\ 0 & 0 \end{bmatrix}.$$

This is not possible because the last row of the matrix on the left has imaginary part $\tau_2 < 0$ and is therefore non-zero. Hence we may assume $R_3 s/\epsilon^2 \to b < \infty$. In particular, this implies that $s \to 0$, since $\epsilon^2/R_3 \to 0$. Thus $V = I$ and we have shown that $V_1 = V_2$.

Next we will show that $S_1 = S_2$ and $T_1 = T_2$. Returning to the range condition, write

$$V(S_1 + iT_1)V^t = \begin{bmatrix} s_{1,1} & 0 \\ s_{2,1} & s_{2,2} \end{bmatrix} + i \begin{bmatrix} 0 & 0 \\ 0 & t_2 \end{bmatrix},$$
$$V(S_2 + iT_2)V^t = \begin{bmatrix} \sigma_{1,1} & 0 \\ \sigma_{2,1} & \sigma_{2,2} \end{bmatrix} + i \begin{bmatrix} 0 & 0 \\ 0 & \tau_2 \end{bmatrix},$$

where now $V = V_1 = V_2$. The zero in the top right corner follows from $R_2/R_1 \to 0$ and $R_4/R_3 \to 0$. Then

$$\begin{bmatrix} s_{1,1} & 0 \\ s_{2,1} & s_{2,2} + it_2 \\ \sigma_{1,1} & 0 \\ \sigma_{2,1} & \sigma_{2,2} + i\tau_2 \end{bmatrix} \in \operatorname{Ran} \begin{bmatrix} 1 & 0 \\ 0 & 1/|t_2| \\ a & b \\ 0 & 1/|\tau_2| \end{bmatrix}.$$

In particular the second column of the matrix on the left must be a non-zero multiple of the second column of the matrix on the right. This is possible only if $b = 0$, so we may assume this. The resulting range condition is equivalent to

$$\begin{bmatrix} \sigma_{1,1} & 0 \\ \sigma_{2,1} & \sigma_{2,2} + i\tau_2 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & 1/|\tau_2| \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & |t_2| \end{bmatrix} \begin{bmatrix} s_{1,1} & 0 \\ s_{2,1} & s_{2,2} + it_2 \end{bmatrix}.$$

Taking the imaginary part of this equation yields $t_2 = \tau_2$. The real part reads

$$\begin{bmatrix} \sigma_{1,1} & 0 \\ \sigma_{2,1} & \sigma_{2,2} \end{bmatrix} = \begin{bmatrix} as_{1,1} & 0 \\ s_{2,1} & s_{2,2} \end{bmatrix}.$$

So $s_{2,1} = \sigma_{2,1}$, $s_{2,2} = \sigma_{2,2}$ and $\sigma_{1,1} = a s_{1,1}$ with $a \geqslant 0$. Finally, (2.5) implies that $s_{1,1}^2 = \sigma_{1,1}^2$ so it must be that $a = 1$ and $s_{1,1} = \sigma_{1,1}$.

Thus we have shown that $S_1 = S_2$ and $T_1 = T_2$. Let us call the common values S and T. It follows from (2.8) that $T_a^2 = T^2$ and from (2.7) that $T_a \leqslant 0$. Thus $T_a = T$. To see that $S_a = S$ too, notice that in the basis where T is diagonal S_a will also have a zero in the top right corner. Thus we can write

$$V S_a V^t = \begin{bmatrix} a_{1,1} & 0 \\ a_{2,1} & a_{2,2} \end{bmatrix}$$

and then (2.10) implies

$$\begin{bmatrix} a_{1,1} & a_{2,1} \\ 0 & a_{2,2} - it \end{bmatrix} \begin{bmatrix} a_{1,1} & 0 \\ a_{2,1} & a_{2,2} + it \end{bmatrix} = \begin{bmatrix} s_{1,1} & s_{2,1} \\ 0 & s_{2,2} - it \end{bmatrix} \begin{bmatrix} s_{1,1} & 0 \\ s_{2,1} & s_{2,2} + it \end{bmatrix}.$$

This gives $a_{2,1} = s_{2,1}$, $a_{1,1}^2 = s_{1,1}^2$ and $a_{2,2}^2 = s_{2,2}^2$. But the equation $X_a = (X_1 + X_2)/2$, written as $R(T_{a,n}, \epsilon_{a,n}) S_{a,n} = (R(T_{1,n}, \epsilon_{1,n}) S_{1,n} + R(T_{2,n}, \epsilon_{2,n}) S_{2,n})/2$ implies that $a_{1,1}$ has the same sign as $s_{1,1}$ and that $a_{2,2}$ has the same sign as $s_{2,2}$. Thus $S_a = S$.

Suppose that $(S + i|T|)$ is invertible. Then $W_1 = W_2 = W_a = (S + i|T|)^{-1}(S + iT)$ and we have proved case (i) of this proposition.

It remains to deal with the case where $(S + i|T|)$ is not invertible. In this case the values of S and T do not completely determine the limiting value of Z (or W). We will show that the possible limiting values are described by case (iii) of this proposition.

The matrix $(S + i|T|)$ is not invertible whenever $s_{1,1} = 0$. So we wish to consider the situation where we have a sequence of positive numbers $\epsilon_n \to 0$ and sequences of matrices $T_n = \left[\begin{smallmatrix} t_{1,n} & 0 \\ 0 & t_{2,n} \end{smallmatrix}\right]$ with $t_{1,n} \to 0$ and $t_{2,n} \to t_2 < 0$ and $S_n = \left[\begin{smallmatrix} s_{1,1,n} & s_{2,1,n} R(t_{2,n}, \epsilon_n)/R(t_{1,n}, \epsilon_n) \\ s_{2,1,n} & s_{2,2,n} \end{smallmatrix}\right]$ with $s_{1,1,n} \to 0$, $s_{2,1,n} \to s_{2,1}$ and $s_{2,2,n} \to s_{2,2}$. Since $R(t_{2,n}, \epsilon_n) \sim \epsilon^2/|t_2|$ we find that

$$\begin{aligned}
\lim_{n\to\infty} Z_n &= \lim_{n\to\infty} \frac{1}{\epsilon_n^2} \bigl(R(T_n, \epsilon_n) S_n + i R(T_n, \epsilon_n)^2 \bigr) \\
&= \lim_{n\to\infty} \frac{1}{\epsilon_n^2} V \left(\begin{bmatrix} R(t_{1,n}, \epsilon_n) & 0 \\ 0 & R(t_{2,n}, \epsilon_n) \end{bmatrix} \begin{bmatrix} s_{1,1,n} & s_{2,1,n} R(t_{2,n}, \epsilon_n)/R(t_{1,n}, \epsilon_n) \\ s_{2,1,n} & s_{2,2,n} \end{bmatrix} \right. \\
&\quad \left. + i \begin{bmatrix} R(t_{1,n}, \epsilon_n)^2 & 0 \\ 0 & R(t_{2,n}, \epsilon_n)^2 \end{bmatrix} \right) V^t \\
&= \lim_{n\to\infty} V \begin{bmatrix} \frac{s_{1,1,n}}{\epsilon_n} R(\frac{t_{1,n}}{\epsilon_n}, 1) + i R(\frac{t_{1,n}}{\epsilon_n}, 1)^2 & \frac{s_{2,1}}{|t_2|} \\ \frac{s_{2,1}}{|t_2|} & \frac{s_{2,2}}{|t_2|} \end{bmatrix} V^t.
\end{aligned}$$

The top left entry can have any limiting value in $\overline{\mathbb{H}}$, depending on the relative rates at which $t_{1,n}$, $s_{1,1,n}$ and ϵ_n converge to zero. This shows that case (iii) of this proposition holds.

Case 0-   - -

Following the calculation above we find in this case that

$$\begin{bmatrix} V_2(S_1+iT_1)V_2^t \\ V_2(S_2+iT_2)V_2^t \end{bmatrix} \in \lim \operatorname{Ran} \begin{bmatrix} R_1c & -R_2s \\ R_1s & R_2c \\ R_3c & -R_3s \\ R_4s & R_4c \end{bmatrix} \begin{bmatrix} R_1^{-1} & 0 \\ 0 & \epsilon^{-2} \end{bmatrix}$$
$$= \lim \operatorname{Ran} \begin{bmatrix} c & -s/|t_2| \\ s & c/|t_2| \\ 0 & -s/|\tau_1| \\ 0 & c/|\tau_2| \end{bmatrix}.$$

This contradicts the fact that S_2+iT_2 is invertible in this case. So this case is not possible.

Case - -   - -

In this case both S_1+iT_1 and S_2+iT_2 are invertible, so the condition (2.5) implies that

$$\begin{bmatrix} S_1+iT_1 \\ S_2+iT_2 \end{bmatrix} \in \operatorname{Ran} \begin{bmatrix} A \\ I \end{bmatrix}$$

for some invertible real matrix A. Then we find that $(S_1+iT_1) = A(S_2+iT_2)$ so that $S_1 = AS_2$ and $T_1 = AT_2$. Then $A = T_1T_2^{-1}$ and so $T_1^{-1}S_1 = T_1^{-1}AS_2 = T_2^{-1}S_2 = B$ for some matrix B. Notice that $B+i = T_1^{-1}(S_1+iT_1)$ is invertible. Now $(S_1+iT_1)^*(S_1+iT_1) = (B+i)^*T_1^2(B+i)$ and similarly $(S_1+iT_1)^*(S_1+iT_1) = (B+i)^*T_2^2(B+i)$. So (2.6) implies $T_1^2 = T_2^2$ which implies $T_1 = T_2$ since both eigenvalues are negative in each case. Then we find $A = I$ and so $S_1 = S_2$ too.

Now we find, using the asymptotics of $R(T_{1,n}, \epsilon_{1,n})$ that $Y_1 = 0$ and $Z_1 = X_1 = |T_1|^{-1}S_1$. Similarly $Y_2 = 0$ and $Z_2 = X_2 = |T_2|^{-1}S_2$. Therefore $Z_1 = Z_2 = (Z_1+Z_2)/2 = Z_a$. This completes the proof. □

Proof of Proposition 1.8. (i) The only fixed point for Ψ_λ in $\overline{\mathbb{SH}}_2$ is $Z = iI$, and this is not on the boundary.

(ii) It follows from (1.7) that

$$\tilde{\Psi}_\lambda = \begin{bmatrix} e^{-i\Theta_\lambda} & 0 \\ 0 & e^{i\Theta_\lambda} \end{bmatrix}$$

where

$$\begin{aligned} e^{-i\Theta_\lambda} &= \cos(\Theta_\lambda) - i\sin(\Theta_\lambda) \\ &= (\lambda - \Delta_G)/(2\sqrt{2}) - i\sqrt{1-(\lambda-\Delta_G)^2/8} \\ &= V_1 \begin{bmatrix} \omega_1(\lambda) & 0 \\ 0 & \omega_2(\lambda) \end{bmatrix} V_1^t. \end{aligned}$$

Here V_1 is the rotation matrix diagonalizing Δ_G and $\omega_1(\lambda) = (\lambda-1)/(2\sqrt{2}) - i\sqrt{1-(\lambda-1)^2/8}$, $\omega_2(\lambda) = (\lambda+1)/(2\sqrt{2}) - i\sqrt{1-(\lambda+1)^2/8}$ lie on the unit circle for $\lambda \in J$.

The equation $\tilde{\Psi}_\lambda \cdot (V\left[\begin{smallmatrix}1&0\\0&\alpha\end{smallmatrix}\right]V^t) = V\left[\begin{smallmatrix}1&0\\0&\beta\end{smallmatrix}\right]V^t$ that we are trying to rule out can now be written $V^t e^{-i\Theta_\lambda} V\left[\begin{smallmatrix}1&0\\0&\alpha\end{smallmatrix}\right]V^t(e^{i\Theta_\lambda})^{-1}V = \left[\begin{smallmatrix}1&0\\0&\beta\end{smallmatrix}\right]$. Since $(e^{i\Theta_\lambda})^{-1} = e^{-i\Theta_\lambda}$ this is equivalent to

$$V_2\begin{bmatrix}\omega_1(\lambda) & 0\\ 0 & \omega_2(\lambda)\end{bmatrix}V_2^t\begin{bmatrix}1&0\\0&\alpha\end{bmatrix}V_2\begin{bmatrix}\omega_1(\lambda) & 0\\ 0 & \omega_2(\lambda)\end{bmatrix}V_2^t = \begin{bmatrix}1&0\\0&\beta\end{bmatrix} \tag{2.16}$$

where $V_2 = V^t V_1$. To show this is impossible for any rotation matrix $V_2 = \left[\begin{smallmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{smallmatrix}\right]$, observe that the matrix

$$U = V_2\begin{bmatrix}\omega_1(\lambda) & 0\\ 0 & \omega_2(\lambda)\end{bmatrix}V_2^t \tag{2.17}$$

is unitary. We obtain from (2.16)

$$U\begin{bmatrix}1&0\\0&\alpha\end{bmatrix} = \begin{bmatrix}1&0\\0&\beta\end{bmatrix}U^*.$$

In particular, the upper left matrix entries have to agree. This gives

$$U_{1,1} = \overline{U}_{1,1}.$$

Thus $\operatorname{Im} U_{1,1} = 0$. On the other hand, using that V_2 is real, it follows from (2.17) that

$$\operatorname{Im} U_{1,1} = \sin^2(\theta)\operatorname{Im}\omega_1(\lambda) + \cos^2(\theta)\operatorname{Im}\omega_2(\lambda).$$

But the right side cannot be zero for $\lambda \in (-2\sqrt{2}+1, 2\sqrt{2}-1)$, in view of the definition of $\omega_i(\lambda)$. Thus (2.16) cannot hold.

(iii) We wish to show that the equation

$$\Psi_\lambda\left(V\begin{bmatrix}z & r\\ r & p\end{bmatrix}V^t\right) = V\begin{bmatrix}z' & r\\ r & p\end{bmatrix}V^t \tag{2.18}$$

cannot hold.

If $z = z' = i\infty$ then we must first transfer (2.18) to the ball model. The point $\left[\begin{smallmatrix}i\infty & r\\ r & p\end{smallmatrix}\right] \in \mathbb{SH}_2$ corresponds to the point $\left[\begin{smallmatrix}1 & 0\\ 0 & (p-i)/(p+i)\end{smallmatrix}\right]$ in the ball model. So, in this case (2.18) asserts that Ψ_λ has a fixed point on the boundary. This is false, so we have ruled out the case $z = z' = i\infty$.

If $z = i\infty$ and $z' \in \mathbb{R}$, then we may compute the left side of (2.18) as follows. Recall from (1.7) that

$$\Psi_\lambda = \begin{bmatrix}\cos(\Theta_\lambda) & -\sin(\Theta_\lambda)\\ \sin(\Theta_\lambda) & \cos(\Theta_\lambda)\end{bmatrix}$$

where

$$\cos(\Theta_\lambda) = V_1\begin{bmatrix}c_1 & 0\\ 0 & c_2\end{bmatrix}V_1^t, \qquad \sin(\Theta_\lambda) = V_1\begin{bmatrix}s_1 & 0\\ 0 & s_2\end{bmatrix}V_1^t.$$

Here V_1 is a real rotation matrix and $s_1, s_2 > 0$. Using this notation and the representation

$$V_2 = V^t V_1 = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

with $c = \cos(\theta)$ and $s = \sin(\theta)$, we can calculate an expression for the left side of (2.18). Upon substituting $z = -1/w$ and setting $w = 0$, (2.18) results in a matrix equation whose bottom right entry can be written

$$c^2 s^2 \big(s_1^2 + s_2^2 + (c_1 - c_2)^2\big) + s_1 s_2 \big(c^4 + s^4\big) + s_1 s_2 p^2 = 0.$$

Since s_1 and s_2 are both strictly positive this equation cannot hold. Thus we have ruled out the case $z = i\infty$ and $z' \in \mathbb{R}$.

The equation above also cannot hold when s_1 and s_2 are replaced with $-s_1$ and $-s_2$, and this can be used to rule out the case $z \in \mathbb{R}$ and $z' = i\infty$.

Finally, if $z_1, z_2 \in \mathbb{R}$ then (2.18) can be written

$$\begin{aligned} &V_2 \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix} V_2^t \begin{bmatrix} z & r \\ r & p \end{bmatrix} - V_2 \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} V_2^t \\ &\quad = \begin{bmatrix} z' & r \\ r & p \end{bmatrix} V_2 \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} V_2^t \begin{bmatrix} z & r \\ r & p \end{bmatrix} + \begin{bmatrix} z' & r \\ r & p \end{bmatrix} V_2 \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix} V_2^t. \end{aligned}$$

The bottom right entry of this equation reads

$$s_1 \big(s^2 + (rc + ps)^2\big) + s_2 \big(c^2 + (rs - pc)^2\big) = 0.$$

Again, since s_1 and s_2 are strictly positive, this equation cannot hold. We have ruled out (2.18) in all cases so the proof of (iii) is complete. □

Acknowledgments

It is pleasure to acknowledge conversations with Jonathan Breuer who suggested that we consider this model, and the Banff International Research Station, where these conversations took place.

Appendix A

Proof of Lemma 1.2. (i) It is enough to prove this statement for Γ of the form $\Gamma = \left[\begin{smallmatrix} I & B \\ 0 & I \end{smallmatrix}\right]$ with $B^T = B$, $\Gamma = \left[\begin{smallmatrix} 0 & -I \\ I & 0 \end{smallmatrix}\right]$ or $\Gamma = \left[\begin{smallmatrix} A & 0 \\ 0 & A^{T-1} \end{smallmatrix}\right]$, since these generate $\mathrm{Sp}(4, \mathbb{R})$. If $\Gamma = \left[\begin{smallmatrix} I & B \\ 0 & I \end{smallmatrix}\right]$ with $B^T = B$ then $\Gamma \cdot (X + iY) = (X + B) + iY$. So $Z_1 - Z_2$, Y_1 and Y_2 are invariant under the action of Γ which implies that $w_p(Z_1, Z_2) = \|Y_2^{-1/2}(Z_1 - Z_2)^* Y_1^{-1}(Z_1 - Z_2)Y_2^{-1/2}\|_{1+p}^{1+p}$ is invariant too. If $\Gamma = \left[\begin{smallmatrix} 0 & -I \\ I & 0 \end{smallmatrix}\right]$ then $\Gamma \cdot Z = -Z^{-1}$. The invariance of w_p follows from the identities $-Z_1^{-1} + Z_2^{-1} = Z_1^{-1}(Z_1 - Z_2)Z_2^{-1}$ and $\operatorname{Im} Z_i^{-1} = Z_i^{-1} Y_i Z_i^{*-1}$, together with the fact that $\|C^* C\|_{1+p} = \|CC^*\|_{1+p}$. The proof for the case $\Gamma = \left[\begin{smallmatrix} A & 0 \\ 0 & A^{T-1} \end{smallmatrix}\right]$ is similar.

(ii) Since $t > 0$ we have $(Y+t)^{-1} \leqslant Y^{-1}$. Thus the required inequality follows from

$$\begin{aligned}
&\left\|(Y_2+t)^{-1/2}(Z_1-Z_2)^*(Y_1+t)^{-1}(Z_1-Z_2)(Y_2+t)^{-1/2}\right\|_{1+p}^{1+p} \\
&\quad\leqslant \left\|(Y_2+t)^{-1/2}(Z_1-Z_2)^*Y_1^{-1}(Z_1-Z_2)(Y_2+t)^{-1/2}\right\|_{1+p}^{1+p} \\
&\quad= \left\|Y_1^{-1/2}(Z_1-Z_2)(Y_2+t)^{-1}(Z_1-Z_2)^*Y_1^{-1/2}\right\|_{1+p}^{1+p} \\
&\quad\leqslant \left\|Y_1^{-1/2}(Z_1-Z_2)Y_2^{-1}(Z_1-Z_2)^*Y_1^{-1/2}\right\|_{1+p}^{1+p} \\
&\quad= \left\|Y_2^{-1/2}(Z_1-Z_2)^*Y_1^{-1}(Z_1-Z_2)Y_2^{-1/2}\right\|_{1+p}^{1+p}.
\end{aligned}$$

(iii) We follow [5]. For $\lambda \in R_\epsilon$, Y_λ is bounded above and below by positive constants. Thus, $\operatorname{Im} G = Y_\lambda^{1/2} \operatorname{Im} Z\, Y_\lambda^{1/2} < C \operatorname{Im} Z$ with constants uniform in λ. Since all norms are equivalent for 2×2 matrices, and by the convexity of $|\cdot|^{1+p}$, it suffices to show that for $Z = X + iY$, $\|Y\|_1 \leqslant \|(Z-iI)^*Y^{-1}(Z-iI)\|_1 + 4$. Because Y is positive definite,

$$\begin{aligned}
\|Y\|_1 &= \operatorname{tr}(Y) \\
&\leqslant \operatorname{tr}\bigl(Y + Y^{-1} - 2I\bigr) + 4 \\
&= \operatorname{tr}\bigl((Y-I)Y^{-1}(Y-I)\bigr) + 4 \\
&\leqslant \operatorname{tr}\bigl((Y-I)Y^{-1}(Y-I) + XY^{-1}X\bigr) + 4 \\
&= \operatorname{tr}\bigl(\bigl(X - i(Y-I)\bigr)Y^{-1}\bigl(X + i(Y-I)\bigr)\bigr) + 4 \\
&= \left\|(Z-iI)^*Y^{-1}(Z-iI)\right\|_1 + 4.
\end{aligned} \tag{A.1}$$

This completes the proof. For future reference, notice that (A.1) also holds with $\|Y^{-1}\|_1$ on the left side.

(iv) Using $\|AB\|_{1+p} \leqslant \|A\|_{2(1+p)}\|B\|_{2(1+p)}$ and $\|A\|_{2(1+p)}^2 = \|A^*A\|_{1+p}$, together with the comment following (A.1) we find that for any $\epsilon > 0$

$$\begin{aligned}
&\left\|(Z+Q-iI)^*Y^{-1}(Z+Q-iI)\right\|_{1+p} \\
&\quad\leqslant \left\|(Z-iI)^*Y^{-1}(Z-iI)\right\|_{1+p} + 2\left\|QY^{-1/2}\right\|_{2(1+p)}\left\|Y^{-1/2}(Z-iI)\right\|_{2(1+p)} \\
&\qquad+ \left\|QY^{-1}Q\right\|_{1+p} \\
&\quad\leqslant (1+\epsilon)\left\|(Z-iI)^*Y^{-1}(Z-iI)\right\|_{1+p} + (1+1/\epsilon)\|Q\|^2\left\|Y^{-1}\right\|_{1+p} \\
&\quad\leqslant \bigl(1+\epsilon+C_\epsilon\|Q\|^2\bigr)\left\|(Z-iI)^*Y^{-1}(Z-iI)\right\|_{1+p} + C_\epsilon\|Q\|^2.
\end{aligned}$$

Now the result follows from the fact that for any $\epsilon > 0$, there is C_ϵ such that $|a+b|^{1+p} \leqslant (1+\epsilon)|a|^{1+p} + C_\epsilon|b|^{1+p}$ for positive a and b. $\square$

Lemma A.1. *Let $Z = X + iY$ be a complex $n \times n$ matrix with X and Y real and symmetric. Moreover assume that $Y \geqslant t_1 > 0$. Then Z is bijective and $\|Z^{-1}\| \leqslant t_1^{-1}$.*

Proof. For all $\varphi \in \mathbb{C}^n$,

$$t_1 \|\varphi\|^2 \leqslant (\varphi, Y\varphi) = \operatorname{Im}(\varphi, Z\varphi) \leqslant \left|(\varphi, Z\varphi)\right| \leqslant \|\varphi\| \|Z\varphi\|,$$

and hence

$$\|\varphi\| \leqslant t_1^{-1} \|Z\varphi\|. \tag{A.2}$$

This implies that Z is injective and hence bijective since n is finite. Inserting $\varphi = Z^{-1}\psi$ into (A.2), we find

$$\left\|Z^{-1}\psi\right\| \leqslant t_1^{-1} \|\psi\|$$

for all $\psi \in \mathbb{C}^n$. This yields the claim. □

References

[1] M. Aizenman, R. Sims, S. Warzel, Stability of the absolutely continuous spectrum of random Schrödinger operators on tree graphs, Probab. Theory Related Fields 136 (3) (2006) 363–394, arXiv:math-ph/0502006v3.
[2] P.J. Freitas, On the action of the symplectic group on the Siegel upper half plane, PhD thesis, University of Illinois at Chicago, 1999, http://ptmat.fc.ul.pt/~pedro/thesis.pdf.
[3] R. Froese, D. Hasler, W. Spitzer, Absolutely continuous spectrum for the Anderson model on a tree: A geometric proof of Klein's theorem, Comm. Math. Phys. 269 (1) (2007) 239–257, mp-arc:05-388.
[4] R. Froese, D. Hasler, W. Spitzer, Absolutely continuous spectrum for a random potential on a tree with strong transverse correlations and large weighted loops, Rev. Math. Phys. 21 (6) (2009) 709–733, arXiv:0809.4197v1 [math-ph].
[5] R. Froese, D. Hasler, W. Spitzer, On the AC spectrum of one-dimensional random Schrödinger operators with matrix-valued potentials, Math. Phys. Anal. Geom. 13 (3) (2010) 219–233, arXiv:0912.0294v1 [math-ph].
[6] R. Froese, D. Hasler, W. Spitzer, A geometric approach to absolutely continuous spectrum for discrete Schrödinger operators, in: Progr. Probab., vol. 64, 2011, pp. 201–226.
[7] F. Halasan, Absolutely continuous spectrum for the Anderson model on some tree-like graphs, arXiv:0810.2516v3 [math-ph], see also U.B.C. thesis "Absolutely continuous spectrum for the Anderson model on trees", 2009, available at http://circle.ubc.ca/handle/2429/18857.
[8] M. Keller, D. Lenz, S. Warzel, Absolutely continuous spectrum for random operators on trees of finite cone type, arXiv:1108.0057v1 [math-ph].
[9] A. Klein, Extended states in the Anderson model on the Bethe lattice, Adv. Math. 133 (1998) 163–184, mp-arc:94-236.
[10] A. Klein, C. Sadel, Absolutely continuous spectrum for random Schrödinger operators on the Bethe strip, Math. Nachr., doi:10.1002/mana.201100019, in press.