Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 89, 1–23.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP340

Random perturbations of hyperbolic dynamics

Florian Dorsch* Hermann Schulz-Baldes†

Abstract

A sequence of large invertible matrices given by a small random perturbation around a fixed diagonal and positive matrix induces a random dynamics on a high-dimensional sphere. For a certain class of rotationally invariant random perturbations it is shown that the dynamics approaches the stable fixed points of the unperturbed matrix up to errors even if the strength of the perturbation is large compared to the relative increase of nearby diagonal entries of the unperturbed matrix specifying the local hyperbolicity.

Keywords: Furstenberg measure; random matrices; random dynamical systems.
AMS MSC 2010: 37H10; 37H15; 37A50; 60B20.
Submitted to EJP on November 5, 2018, final version accepted on July 7, 2019.
Supersedes arXiv:1805.11064v2.

1 Model, main results and comments

Let us consider the random dynamics on the L-dimensional sphere $\mathbb{S}^{\mathsf{L}}$, $\mathsf{L} \geq 2$, given by

$$v_n \;=\; \mathcal{T}_n \cdot v_{n-1}\,, \qquad n \in \mathbb{N}\,, \tag{1.1}$$

where the action $\cdot\; : \mathrm{GL}(\mathsf{L}+1,\mathbb{R}) \times \mathbb{S}^{\mathsf{L}} \to \mathbb{S}^{\mathsf{L}}$ of the general linear group is

$$\mathcal{T} \cdot v \;=\; \frac{\mathcal{T} v}{\|\mathcal{T} v\|}\,, \tag{1.2}$$

and the random matrices $\mathcal{T}_n$ are of the form

$$\mathcal{T}_n \;=\; \mathcal{R}\,(\mathbf{1} \,+\, \lambda r_n U_n)\,, \qquad n \in \mathbb{N}\,. \tag{1.3}$$

Here $\mathcal{R} = \mathrm{diag}(\kappa_{\mathsf{L}+1},\ldots,\kappa_1)$ is a fixed *unperturbed* positive diagonal matrix whose entries satisfy $\kappa_1 \geq \cdots \geq \kappa_{\mathsf{L}+1} > 0$ and a *random perturbation* $\lambda r_n U_n$ is given by a coupling constant $\lambda \in [0,1)$, a scalar *radial randomness* r_n and an *angular randomness*

*Department Mathematik, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany.
E-mail: dorsch@math.fau.de
†Department Mathematik, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany.
E-mail: schuba@mi.uni-erlangen.de

induced by orthogonal matrices U_n. The main assumption is that both the U_n and r_n are independent and identically distributed (i.i.d.) with $\mathrm{O}(\mathsf{L}+1)$-valued and Haar distributed U_n and $[0,1]$-valued $r_n \not\equiv 0$. This includes, *e.g.*, the case $r_n \equiv 1$. Hence the object of study is a particular Markov process on the continuous state space $\mathbb{S}^{\mathsf{L}}$.

The above is the standard set-up of the theory of products of random matrices [3] except that usually the action is studied on the projective space and not its double cover by $\mathbb{S}^{\mathsf{L}}$, but for sake of simplicity we suppress this difference. By Furstenberg's Theorem the random action has a unique invariant probability measure $\mu_{r,\lambda}$ on $\mathbb{S}^{\mathsf{L}}$ if $\lambda \neq 0$ (see [3], Part A, Theorem III.4.3). This paper is about obtaining further quantitative information about this invariant measure in the special case described above. Hence the paper is thematically located at the interface between random matrix theory, the theory of products of random matrices and random dynamical systems. One of the key technical elements in the proofs is a stochastic order underlying the process (1.1) with $\mathcal{R} = \mathbf{1}$, see Proposition 2.11 below.

Let us begin by describing the dynamics (1.1) heuristically. The unperturbed deterministic dynamics $\mathcal{R}\cdot$ induced by $\mathcal{R}$ is maximally hyperbolic if the deterministic local expansion rates

$$\delta\mathcal{R}_i \;=\; \frac{\kappa_i - \kappa_{i+1}}{\kappa_{i+1}}$$

are strictly positive for all $i = 1, \ldots, \mathsf{L}$. Then there is a simple stable fixed point given by the unit vector $e_{\mathsf{L}+1}$ corresponding to the last component (the fixed point is unique only on projective space). The deterministic dynamics $\mathcal{R}^N \cdot v_0$ converges to the unit vector e_j if j is the largest index such that the jth component of the initial condition v_0 does not vanish. However, e_j is an unstable fixed point of $\mathcal{R}\cdot$ if $j \leq \mathsf{L}$. All these facts are elementary to check. In the following, we also speak of the unit eigenvector $e_{\mathsf{L}+2-j}$ of the eigenvalue κ_j as the jth channel specified by the unperturbed dynamics. We will *not* assume maximal hyperbolicity in the following.

If now the strength of the perturbation is non-zero and satisfies $\lambda < 2^{-4}\min\{\delta\mathcal{R}_1, \frac{1}{2}\}$, one can prove that the random dynamics leaves any unstable fixed point and is driven to the vicinity of the stable fixed point in which it then remains. Thus in this case the Furstenberg invariant measure $\mu_{r,\lambda}$ is supported only by a strict subset of $\mathbb{S}^{\mathsf{L}}$, which is a neighborhood of the stable fixed point. More generally, the theorem below states that if $\lambda < 2^{-4}\min\{\delta\mathcal{R}_i, \frac{1}{2}\}$ for some i, then $\mathrm{supp}(\mu_{r,\lambda})$ is a strict subset of $\mathbb{S}^{\mathsf{L}}$. From the proof one can infer that the support is a small (in a quantitative manner) neighborhood of $\{0\}^{\mathsf{L}+1-i} \times \mathbb{S}^{i-1}$. The main interest of this paper is, however, to analyze the situation where several of the $\delta\mathcal{R}_i$ vanish or are at least all smaller than λ. Hence the unperturbed dynamics may be merely partially hyperbolic. In this situation the random perturbation is *not* small compared to the local hyperbolicity of $\mathcal{R}$. Intuitively, it is clear that the random dynamics may then visit all points on $\mathbb{S}^{\mathsf{L}}$ because the randomness can overcome the hyperbolic character of $\mathcal{R}$ and lead to significant escapes from anywhere. This just means that the support of the invariant measure is the whole sphere $\mathbb{S}^{\mathsf{L}}$. This last fact is precisely part of the following first result.

Theorem 1.1. *Suppose that $\lambda \in (0,1)$, that the i.i.d. $r_n \not\equiv 0$ are $[0,1]$-valued and that the i.i.d. U_n are Haar distributed on $\mathrm{O}(\mathsf{L}+1)$. Then the Furstenberg measure $\mu_{r,\lambda}$ is absolutely continuous w.r.t. the normalized surface measure ν_{L}. If $\mathbb{P}(r=0)=0$ holds, then the random variables $v_N \in \mathbb{S}^{\mathsf{L}}$ are distributed absolutely continuously w.r.t. ν_{L} on $\mathbb{S}^{\mathsf{L}}$ even for any $N \geq 1$ and initial condition v_0. Provided that $\lambda < 2^{-4}\min\{\delta\mathcal{R}_i, \frac{1}{2}\}$ for some $i = 1, \ldots, \mathsf{L}$, the support of $\mu_{r,\lambda}$ is a strict subset of $\mathbb{S}^{\mathsf{L}}$. If $\lambda > \delta\mathcal{R}_i$ for all $i = 1, \ldots, \mathsf{L}$ and $1 \in \mathrm{supp}(r)$, then the support of $\mu_{r,\lambda}$ is the whole sphere $\mathbb{S}^{\mathsf{L}}$.*

Now let us suppose that the randomness, while being large compared to the local expansion rates $\lambda > \delta\mathcal{R}_i$, is small compared to the expansion rates

$$\delta\mathcal{R}_{i,j} \;=\; \frac{\kappa_i - \kappa_j}{\kappa_j}\;,$$

from channel i to channel j for some $j > i$. Then if $\lambda < \delta\mathcal{R}_{i,j}$, there is some contraction hyperbolicity on this larger scale, even though the local hyperbolicity is dominated by the randomness. Hence a finer analysis of the interplay between the randomness and the hyperbolic unperturbed dynamics is needed. Intuitively, one certainly expects the random dynamics to spend little time in the channel j and this should lead to a small weight of the Furstenberg measure on this channel. Roughly this is what we actually prove below. To state our main result more precisely, we need some further notations. Let us partition the channels into three parts $(\mathsf{L}_\mathfrak{a}, \mathsf{L}_\mathfrak{b}, \mathsf{L}_\mathfrak{c}) \in \mathbb{N}\times\mathbb{N}\times\mathbb{N}$, namely $\mathsf{L}_\mathfrak{a} + \mathsf{L}_\mathfrak{b} + \mathsf{L}_\mathfrak{c} = \mathsf{L}+1$. Each vector $v = (v_1, \ldots, v_{\mathsf{L}+1})^\mathsf{T} \in \mathbb{R}^{\mathsf{L}+1}$ is split into its upper part $\mathfrak{a}(v) \in \mathbb{R}^{\mathsf{L}_\mathfrak{a}}$, middle part $\mathfrak{b}(v) \in \mathbb{R}^{\mathsf{L}_\mathfrak{b}}$ and lower part $\mathfrak{c}(v) \in \mathbb{R}^{\mathsf{L}_\mathfrak{c}}$ via

$$\mathfrak{a}(v) \;=\; (v_1, \ldots, v_{\mathsf{L}_\mathfrak{a}})^\mathsf{T}\;, \quad \mathfrak{b}(v) \;=\; (v_{\mathsf{L}_\mathfrak{a}+1}, \ldots, v_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}})^\mathsf{T}\;, \quad \mathfrak{c}(v) \;=\; (v_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, v_{\mathsf{L}+1})^\mathsf{T}\;.$$

Moreover, let us introduce the *macroscopic gap* $\gamma = \gamma\,(\mathcal{R}, \mathsf{L}_\mathfrak{b}, \mathsf{L}_\mathfrak{c})$ between the upper and lower parts by

$$\gamma \;=\; \min\left\{1\,,\;\frac{\kappa^2_{\mathsf{L}_\mathfrak{c}}}{\kappa^2_{\mathsf{L}_\mathfrak{b}+\mathsf{L}_\mathfrak{c}+1}} \;-\; 1\right\} \;\in\; [0,1]\;. \tag{1.4}$$

Note that the macroscopic gap γ is positive provided that $\kappa_{\mathsf{L}_\mathfrak{c}} > \kappa_{\mathsf{L}_\mathfrak{b}+\mathsf{L}_\mathfrak{c}+1}$. Now the deviation of the random path $(v_n)_{n\in\mathbb{N}}$ defined by (1.1) and (1.3) from the attractive part $\{0\}^{\mathsf{L}_\mathfrak{a}} \times \mathbb{S}^{\mathsf{L}_\mathfrak{b}+\mathsf{L}_\mathfrak{c}-1}$ of the phase space can be measured as the norm of the upper part $\|\mathfrak{a}(v_N)\|$. The main result provides a quantitative bound on the expectation value of $\|\mathfrak{a}(v_N)\|^2$ for sufficiently large N when the expectation is taken over the randomness contained in $\mathcal{T}_n$ for $n = 1, \ldots, N$.

Theorem 1.2. *Suppose that the i.i.d. $r_n \not\equiv 0$ are $[0,1]$-valued and that the i.i.d. U_n are Haar distributed on $\mathsf{O}(\mathsf{L}+1)$. Furthermore suppose $(\mathsf{L}_\mathfrak{a}, \mathsf{L}_\mathfrak{b}) \neq (1,1)$ and $\gamma > 0$. Then, for all $0 < \lambda \leq \frac{1}{4}$ there exist $N_0 = N_0(\mathsf{L}, \mathsf{L}_\mathfrak{c}, \lambda) \in \mathbb{N}$ such that*

$$\mathbb{E}\,\|\mathfrak{a}(v_N)\|^2 \;\leq\; 2\left(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}\right)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-2}{\mathsf{L}_\mathfrak{c}+2}} \left(\frac{6}{\gamma}\,\frac{\mathsf{L}_\mathfrak{a}}{\mathsf{L}_\mathfrak{c}}\,\lambda^2\right)^{\frac{\mathsf{L}_\mathfrak{c}}{2+\mathsf{L}_\mathfrak{c}}} \tag{1.5}$$

for all $N \geq N_0$ and $v_0 \in \mathbb{S}^\mathsf{L}$.

Using the invariance property of the Furstenberg measure $\mu_{r,\lambda}$, one deduces the following

Corollary 1.3. *Under the same hypothesis as in Theorem 1.2,*

$$\int \mathrm{d}\mu_{r,\lambda}(v)\;\|\mathfrak{a}(v)\|^2 \;\leq\; 2\left(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}\right)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-2}{\mathsf{L}_\mathfrak{c}+2}} \left(\frac{6}{\gamma}\,\frac{\mathsf{L}_\mathfrak{a}}{\mathsf{L}_\mathfrak{c}}\,\lambda^2\right)^{\frac{\mathsf{L}_\mathfrak{c}}{2+\mathsf{L}_\mathfrak{c}}} \tag{1.6}$$

The estimates (1.5) and (1.6) strongly differ from the behavior for $\mathcal{R} = \mathbf{1}$ where no hyperbolicity is present. Then $\mathbb{E}\|\mathfrak{a}(v_N)\|^2 \sim \mathsf{L}_\mathfrak{a}\mathsf{L}^{-1}$ holds for large N independent of $\lambda > 0$ which just reflects the equidistribution of the random dynamics on all channels (this follows from Proposition 2.14 below). To us, the most interesting regime is that of large $\mathsf{L}_\mathfrak{a}$, $\mathsf{L}_\mathfrak{b}$ and $\mathsf{L}_\mathfrak{c}$, say all a fraction of L, and of γ of the order of 1 (but possibly less than 1). Then the r.h.s. in (1.5) and (1.6) is approximately proportional to λ^2 which is the expected behavior. Indeed, the random kicks of order λ are uniform and thus do not distinguish between channels, and hence the drift into each channel is given by their variance or λ^2, so that $\mathbb{E}\|\mathfrak{a}(v)\|^2$ should be of the order λ^2 times the proportion $\mathsf{L}_{\mathfrak{a}}\mathsf{L}^{-1}$ of channels in $\mathfrak{a}(v)$.

Our main motivation for the present study are potential applications to the field of discrete random Schrödinger operators like the Anderson model, see [3, 4, 1] for general mathematical background information. Little is known rigorously about the so-called *weak localization regime* of such operators in space dimension higher than or equal to 3. In this regime, the eigenfunctions are not expected to be exponentially localized and the quantum dynamics is believed to be diffusive like in a Brownian motion. Furthermore, random matrix theory is expected to provide a good description of the eigenvalues and eigenfunctions locally in space and within a suitable range of energies. In infinite volume the spectral measures likely have an absolutely continuous component. The approach to this problem closest to the present study is the transfer matrix method. It allows to construct (generalized) eigenfunctions and Green functions of finite volume approximations. Best understood is then the quasi-one-dimensional limit in which one has strong Anderson localization, that is, pure-point spectrum with exponentially localized eigenfunctions with a rate called the *inverse localization length* [3, 4, 5]. In a perturbative regime of small coupling of the randomness, one can calculate this localization length [14, 10] and, more generally, the whole Lyapunov spectrum [11, 12] provided the random dynamics of the transfer matrices is well understood. For such systems, one can also derive flow equations for the finite volume growth exponents, the so-called DMPK-equations [2, 15, 13]. Beneath these works, only [14, 13] address the hyperbolic character of the unperturbed dynamics (corresponding to the $\mathcal{R}$ above), however, only in the regime $\lambda \ll \delta\mathcal{R}_i$ of very small randomness [14] or even a randomness vanishing at a suitable rate in the system size, namely the number of random matrices $\mathcal{T}_n$ involved [13].

In order to apply the results of this paper (notably Theorem 1.2) to the transfer matrices of the Anderson model and extract relevant information on its eigenfunctions, several non-trivial extensions have to be worked out. First of all, the transfer matrices at real energies have a symplectic symmetry that has to be implemented and then leads, in particular, to a supplementary symmetry in the Lyapunov spectrum. This can be done as in [3, 14, 8]. Then one has to consider the dynamics not only on unit vectors, but rather on the whole flag manifold [3, 14]. Furthermore, while the transfer matrices can be brought in the form (1.3) [14], the random matrices U_n stemming from the Anderson model are not Haar distributed and contain much fewer random entries. In the quasi-one-dimensional regime, this can be dealt with using commutator methods, see [5] and [12] for a perturbative result when $\mathcal{R}$ is elliptic, that is, of unit norm.

Theorem 1.2 also has some short-comings by itself. First of all, it and its proof do not provide a good quantitative estimate on N_0. Furthermore, the proof does not readily transpose to the case where $\mathbf{1} + \lambda r U$ is replaced by $\exp(\lambda r U)$. Actually, many of the arguments below depend heavily on geometric considerations and explicit calculations exploring formulas for averages over the Haar measure.

2 Outline of the proofs of Theorems 1.1 and 1.2

Throughout the remainder of the paper we assume that $\lambda \in (0,1)$ and that $(r_n)_{n\in\mathbb{N}}$ are i.i.d., $[0,1]$-valued and satisfy $\mathbb{P}(r_n = 0) < 1$. Furthermore $(U_n)_{n\in\mathbb{N}}$ are supposed to be $\mathrm{O}(\mathsf{L}+1)$-valued and i.i.d. according to the Haar measure. The one-dimensional Lebesgue measure will be denoted by $\mathbf{x}$. We also abbreviate absolutely continuous and absolute continuity by a.c..

Lemma 2.1. *The random variable $\langle v, (\mathbf{1} + \lambda r U) \cdot v\rangle$ is $[\sqrt{1-\lambda^2}, 1]$-valued for all $v \in \mathbb{S}^{\mathsf{L}}$. Moreover, its distribution is independent of $v \in \mathbb{S}^{\mathsf{L}}$.*

Let us denote the Borel probability distribution of $\langle v, (\mathbf{1}+\lambda r U)\cdot v\rangle$ by

$$\varpi_{r,\lambda}(A) \;=\; \mathbb{P}(\langle v, (\mathbf{1}+\lambda r U)\cdot v\rangle \in A)\;, \qquad A\in\mathscr{B}([\sqrt{1-\lambda^2},1])\;.$$

The aim of the next lemma is to analyze its canonical decomposition into pure-point, singular continuous and absolutely continuous component:

$$\varpi_{r,\lambda} \;=\; \varpi^{\rm pp}_{r,\lambda} \;+\; \varpi^{\rm sc}_{r,\lambda} \;+\; \varpi^{\rm ac}_{r,\lambda}\;.$$

Lemma 2.2. *One has $\varpi^{\rm pp}_{r,\lambda}=\mathbb{P}(r=0)\,\delta_1$ and $\varpi^{\rm sc}_{r,\lambda}=0$.*

The next lemma states an elementary invariance property.

Lemma 2.3. *For all $v\in\mathbb{S}^{\mathsf{L}}$ the random variable $(\mathbf{1}+\lambda rU)\cdot v$ is distributed axially symmetrically w.r.t. v. More precisely, for all Borel subsets $A\in\mathscr{B}(\mathbb{S}^{\mathsf{L}})$, any orthogonal $V\in\mathsf{O}(\mathsf{L}+1)$ and all pairs $(\mathscr{V},v)\in\mathsf{O}(\mathsf{L}+1)\times\mathbb{S}^{\mathsf{L}}$ with $\mathscr{V}v=v$, one has*

$$\begin{aligned}\mathbb{P}\left((\mathbf{1}+\lambda rU)\cdot v\in A\right) \;&=\; \mathbb{P}\left((\mathbf{1}+\lambda rU)\cdot Vv\in VA\right) &(2.1)\\ &=\; \mathbb{P}\left((\mathbf{1}+\lambda rU)\cdot v\in \mathscr{V}A\right)\;. &(2.2)\end{aligned}$$

Lemmata 2.2 and 2.3 allow to consider the Borel probability distribution on $\mathbb{S}^{\mathsf{L}}$ of the random variable $(\mathbf{1}+\lambda rU)\cdot v$:

$$\varrho_{r,\lambda,v}(A) \;=\; \mathbb{P}((\mathbf{1}+\lambda rU)\cdot v\in A)\;, \qquad A\in\mathscr{B}(\mathbb{S}^{\mathsf{L}})\;.$$

The next lemma analyzes its canonical decomposition $\varrho_{r,\lambda,v}=\varrho^{\rm pp}_{r,\lambda,v}+\varrho^{\rm sc}_{r,\lambda,v}+\varrho^{\rm ac}_{r,\lambda,v}$ w.r.t. ν_{L}.

Lemma 2.4. *Let $v,w\in\mathbb{S}^{\mathsf{L}}$. Then*

$$\varrho^{\rm pp}_{r,\lambda,v} \;=\; \mathbb{P}(r=0)\,\delta_v\;, \qquad \varrho^{\rm sc}_{r,\lambda,v}\;=\;0\;, \tag{2.3}$$

and the Radon-Nikodym derivative of the absolutely continuous part $\varrho^{\rm ac}_{r,\lambda,v}$ w.r.t. ν_{L} obeys the following symmetry property:

$$\frac{{\rm d}\varrho^{\rm ac}_{r,\lambda,v}}{{\rm d}\nu_{\mathsf{L}}}(w) \;=\; \frac{{\rm d}\varrho^{\rm ac}_{r,\lambda,w}}{{\rm d}\nu_{\mathsf{L}}}(v)\,. \tag{2.4}$$

The final preparatory result involves the deterministic hyperbolic part $\mathcal{R}\cdot$ of the dynamics.

Lemma 2.5. *The absolute continuity of Borel measures on $\mathbb{S}^{\mathsf{L}}$ w.r.t. ν_{L} is preserved under $(\mathcal{R}\cdot)_*$.*

Once all these lemmata are proved (once again, see Section 3), it is possible to complete the proof of the first part of Theorem 1.1, namely to prove the absolute continuity stated therein.

Proof of Theorem 1.1. For $n\in\mathbb{N}$, let us denote the distribution of v_n for some given initial condition $v_0\in\mathbb{S}^{\mathsf{L}}$ by ς_n. It can be computed iteratively by

$$\varsigma_n \;=\; \int_{\mathbb{S}^{\mathsf{L}}} {\rm d}\varsigma_{n-1}(w)\;((\mathcal{R}\cdot)_*(\varrho_{r,\lambda,w}))(\cdot)\;, \qquad \varsigma_0\;=\;\delta_{v_0}\;. \tag{2.5}$$

Now, let $\aleph\in\mathscr{B}(\mathbb{S}^{\mathsf{L}})$ be a ν_{L}-nullset. Then, $\aleph$ is also an $((\mathcal{R}\cdot)_*(\varrho^{\rm ac}_{r,\lambda,w}))$-nullset by Lemma 2.5. Therefore (2.5) combined with (2.3) implies that

$$\begin{aligned}\varsigma_n(\aleph) \;&=\; \int_{\mathbb{S}^{\mathsf{L}}} {\rm d}\varsigma_{n-1}(w)\;((\mathcal{R}\cdot)_*(\varrho_{r,\lambda,w}))(\aleph)\\ &=\; \int_{\mathbb{S}^{\mathsf{L}}} {\rm d}\varsigma_{n-1}(w)\;((\mathcal{R}\cdot)_*(\varrho^{\rm pp}_{r,\lambda,w}))(\aleph)\\ &=\; \mathbb{P}(r=0)\;\int_{\mathbb{S}^{\mathsf{L}}} {\rm d}\varsigma_{n-1}(w)\;((\mathcal{R}\cdot)_*(\delta_w))(\aleph)\\ &=\; \mathbb{P}(r=0)\;((\mathcal{R}\cdot)_*(\varsigma_{n-1}))(\aleph)\;.\end{aligned} \tag{2.6}$$

By iteratively applying (2.6) from $n = 1$ to some $N \geq 1$ and inserting $\varsigma_0 = \delta_{v_0}$ one obtains

$$\varsigma_N(\aleph) \;=\; \mathbb{P}(r=0)^N \;((\mathcal{R}^N\cdot)_*(\varsigma_0))(\aleph) \;=\; \mathbb{P}(r=0)^N \;((\mathcal{R}^N\cdot)_*(\delta_{v_0}))(\aleph)\,. \tag{2.7}$$

An iterative application of the invariance property of the Furstenberg measure $\mu_{r,\lambda}$ yields

$$\mu_{r,\lambda} \;=\; \int_{\mathbb{S}^{\mathsf{L}}} \mathrm{d}\mu_{r,\lambda}(v)\; \varsigma_N\big|_{v_0=v}\,,$$

which implies together with (2.7) that

$$\mu_{r,\lambda}(\aleph) \;=\; \int_{\mathbb{S}^{\mathsf{L}}} \mathrm{d}\mu_{r,\lambda}(v)\; \varsigma_N(\aleph)\big|_{v_0=v} \;=\; \mathbb{P}(r=0)^N \;((\mathcal{R}^N\cdot)_*(\mu_{r,\lambda}))(\aleph) \;\leq\; \mathbb{P}(r=0)^N\,. \tag{2.8}$$

Due to the assumption $r \not\equiv 0$, *i.e.*, $\mathbb{P}(r=0) < 1$, the absolute continuity of $\mu_{r,\lambda}$ w.r.t. ν_{L} follows from (2.8) in the limit $N \to \infty$. If $\mathbb{P}(r=0) = 0$ holds, then (2.7) implies that even the distribution of v_N is absolutely continuous w.r.t. ν_{L} for all $N \geq 1$.

The penultimate statement of Theorem 1.1 is Lemma 2.6 below. Let us now focus on the last claim, namely the fact that the support of the Furstenberg measure is the whole sphere $\mathbb{S}^{\mathsf{L}}$ if $\lambda > \delta\mathcal{R}_i$ for all i. More precisely, we show that $\mu_{r,\lambda}(B_\epsilon(w)) > 0$ holds for every ball of radius $\epsilon > 0$ around any arbitrary point $w \in \mathbb{S}^{\mathsf{L}}$. For this purpose, let us pick some $u \in \mathrm{supp}(\mu_{r,\lambda})$. In view of Lemma 2.7 (see also below), there exists a path of finite length from u to w, *i.e.*, there exists $N \in \mathbb{N}$ and $\{s_n\}_{n=1}^N \subset \mathrm{supp}(r)$ and $\{\mathscr{U}_n\}_{n=1}^N \subset \mathrm{O}(\mathsf{L}+1) = \mathrm{supp}(U)$ such that

$$w \;=\; \prod_{n=1}^{N} \mathcal{R}(\mathbf{1} + \lambda s_n \mathscr{U}_n)\cdot u$$

holds. Obviously, the event

$$\Big\| \prod_{n=1}^{N} \mathcal{R}(\mathbf{1} + \lambda r_n U_n) - \prod_{n=1}^{N} \mathcal{R}(\mathbf{1} + \lambda s_n \mathscr{U}_n) \Big\| \;<\; \zeta$$

has positive probability for all $\zeta > 0$ and the map $(A, v) \mapsto A \cdot v$ is continuous. Therefore, there exists some $\xi > 0$ such that $\mathbb{P}(v_N \in B_\epsilon(w)) > 0$ for all $v_0 \in B_\xi(u)$. Now every ball $B_\xi(u)$ of radius $\xi > 0$ around $u \in \mathrm{supp}(\mu_{r,\lambda})$ satisfies $\mu_{r,\lambda}(B_\xi(u)) > 0$. This allows to infer that

$$\mu_{r,\lambda}(B_\epsilon(w)) \;=\; \int_{\mathbb{S}^{\mathsf{L}}} \mathrm{d}\mu_{r,\lambda}(v_0)\; \mathbb{P}(v_N \in B_\epsilon(w)) \;\geq\; \int_{B_\xi(u)} \mathrm{d}\mu_{r,\lambda}(v_0)\; \mathbb{P}(v_N \in B_\epsilon(w)) \;>\; 0\,,$$

which proves the claim. □

Lemma 2.6. *If* $\lambda < 2^{-4}\min\{\delta\mathcal{R}_i, \frac{1}{2}\}$ *for some* $i = 1, \dots, \mathsf{L}$, *then* $\mathrm{supp}(\mu_{r,\lambda}) \neq \mathbb{S}^{\mathsf{L}}$.

Lemma 2.7. *Suppose that* $\lambda > \max_{i=1,\dots,\mathsf{L}} \delta\mathcal{R}_i$ *and that* $1 \in \mathrm{supp}(r)$. *Then for every couple* $u, w \in \mathbb{S}^{\mathsf{L}}$ *there exist* $N \in \mathbb{N}$ *and* $s_1, \dots, s_N \in \mathrm{supp}(r)$ *and* $\mathscr{U}_1, \dots, \mathscr{U}_N \in \mathrm{O}(\mathsf{L}+1)$ *such that*

$$w \;=\; \prod_{n=1}^{N} \mathcal{R}(\mathbf{1} + \lambda s_n \mathscr{U}_n)\cdot u\,.$$

Next let us outline the proof of Theorem 1.2. It will be useful to split each $\mathcal{T}_n$ into the unperturbed, deterministic action $\mathcal{R}$ and a random perturbation $\mathbf{1} + \lambda r U$, and analyze the action of both factors separately. The unperturbed action $\mathcal{R}\cdot$ leads to a decrease of the norm of the upper part and an increase of the norm of the lower part. More precisely, provided that $\kappa_{\mathsf{L}_\mathfrak{c}} > \kappa_{\mathsf{L}_\mathfrak{b}+\mathsf{L}_\mathfrak{c}+1}$, one has for any $v \in \mathbb{S}^\mathsf{L}$ obeying $\|\mathfrak{a}(v)\| \neq 0 \neq \|\mathfrak{c}(v)\|$ the bounds

$$\|\mathfrak{a}(\mathcal{R}\cdot v)\| \;<\; \|\mathfrak{a}(v)\| \;, \qquad \|\mathfrak{c}(\mathcal{R}\cdot v)\| \;>\; \|\mathfrak{c}(v)\| \;. \tag{2.9}$$

The former inequality is now strengthened.

Lemma 2.8. *For all $v \in \mathbb{S}^\mathsf{L}$,*

$$\|\mathfrak{a}\,(\mathcal{R}\cdot v)\|^2 \;\leq\; \Big(1-\|\mathfrak{c}(v)\|^2\,\frac{\gamma}{2}\Big)\,\|\mathfrak{a}(v)\|^2 \;. \tag{2.10}$$

This implies that the unperturbed dynamics obeys

$$\lim_{N\to\infty}\,\mathfrak{a}(\mathcal{R}^N\cdot v_0) \;=\; 0$$

if $\|\mathfrak{c}(v_0)\| > 0$ and $\kappa_{\mathsf{L}_\mathfrak{c}} > \kappa_{\mathsf{L}_\mathfrak{b}+\mathsf{L}_\mathfrak{c}+1}$. The random perturbation, on the other hand, may augment $\|\mathfrak{a}(v)\|$. However, in expectation this growth is bounded by a term of order $\mathcal{O}(\lambda^2)$.

Lemma 2.9. *Let $\lambda \in (0,\frac{1}{4}]$ and $\mathsf{L} \geq 3$. Then for all $v \in \mathbb{S}^\mathsf{L}$,*

$$\mathbb{E}\,\|\mathfrak{a}\,((\mathbf{1}+\lambda rU)\cdot v)\|^2 \;\leq\; \|\mathfrak{a}(v)\|^2 \;+\; \lambda^2\,\frac{3\,\mathsf{L}_\mathfrak{a}}{\mathsf{L}+1}\;. \tag{2.11}$$

At first glance, it may now appear straightforward to prove upper bounds on $\mathbb{E}\|\mathfrak{a}(v_N)\|^2$ for large N by combining Lemmata 2.8 and 2.9. An iterative application turns out to be more involved, however. The core task is to deal with the expectation value of products $\|\mathfrak{a}(v_n)\|^2\,\|\mathfrak{c}(v_n)\|^2$ in (2.10). This is tackled by the following elementary lemma.

Lemma 2.10. *$\mathbb{S}^\mathsf{L}$-valued random variables u with arbitrary distribution satisfy*

$$\mathbb{E}\;\|\mathfrak{a}(u)\|^2\,\|\mathfrak{c}(u)\|^2 \;\geq\; \delta\,\Big[\mathbb{E}\;\|\mathfrak{a}(u)\|^2 \;-\; \mathbb{P}\big(\|\mathfrak{c}(u)\|^2<\delta\big)\Big] \tag{2.12}$$

for all $\delta \in [0,1]$.

Consequently the next aim is to bound

$$\mathbb{P}\big(\,\|\mathfrak{c}\,((\mathbf{1}+\lambda rU)\cdot v_n)\|^2 \leq \delta\big) \tag{2.13}$$

from above so that inequalities (2.10) and (2.12) can be used. This turns out to be possible by comparing the random dynamics (1.1) generated by (1.3) with the random dynamics generated by $\mathbf{1}+\lambda r_n U_n$ instead of $\mathcal{T}_n$, that is, the case of $\mathcal{R}=\mathbf{1}$ which has no hyperbolicity. The comparison of the cumulative distribution function (2.13) under these two random dynamics is based on the next result.

Proposition 2.11. *Let $(\mathsf{L}_\mathfrak{a},\mathsf{L}_\mathfrak{b}) \neq (1,1)$ and $v,w \in \mathbb{S}^\mathsf{L}$ be such that $\|\mathfrak{c}(v)\| \geq \|\mathfrak{c}(w)\|$. For all $\epsilon \in [0,1]$ and $\lambda \in (0,\frac{1}{4}]$, one then has*

$$\mathbb{P}\big(\,\|\mathfrak{c}\,((1+\lambda rU)\cdot v)\| \leq \epsilon\big) \;\leq\; \mathbb{P}\big(\,\|\mathfrak{c}\,((1+\lambda rU)\cdot w)\| \leq \epsilon\big)\;. \tag{2.14}$$

Remark 2.12. Since $\|\mathfrak{c}\,((1+\lambda rU)\cdot v)\|$ and $\|\mathfrak{c}\,((1+\lambda rU)\cdot w)\|$ are $\mathbb{R}$-valued, the validity of (2.14) for all $\epsilon \in [0,1]$ is equivalent to the stochastic order

$$\mathbb{P}\big(\,\|\mathfrak{c}\,((1+\lambda rU)\cdot v)\| \in \cdot\,\big) \;\geq_{\mathrm{st}}\; \mathbb{P}\big(\,\|\mathfrak{c}\,((1+\lambda rU)\cdot w)\| \in \cdot\,\big)\;, \tag{2.15}$$

as defined, *e.g.*, in Section 17.7 of [6].

Now one can iteratively combine the second part of (2.9) and Proposition 2.11. For ordered products, we use the following notation:

$$\prod_{i=j}^{k} F_i \;=\; \begin{cases} F_k\cdots F_j\;, & j\leq k\;,\\ \mathbf{1}\;, & j>k\;.\end{cases}$$

Corollary 2.13. *Let* $(\mathsf{L}_\mathfrak{a},\mathsf{L}_\mathfrak{b})\neq(1,1)$*. Then for all* $v\in\mathbb{S}^{\mathsf{L}}$*,* $\epsilon\in[0,1]$*,* $N\in\mathbb{N}$ *and* $\lambda\in(0,\frac14]$*,*

$$\mathbb{P}\Big(\Big\|\mathfrak{c}\big((\mathbf{1}+\lambda r_N U_N)\prod_{n=1}^{N-1}\mathcal{R}(\mathbf{1}+\lambda r_n U_n)\cdot v\big)\Big\|\leq\epsilon\Big)\;\leq\;\mathbb{P}\Big(\Big\|\mathfrak{c}\big(\prod_{n=1}^{N}(\mathbf{1}+\lambda r_n U_n)\cdot v\big)\Big\|\leq\epsilon\Big)\,. \quad (2.16)$$

Corollary 2.13 allows to bound (2.13) by the r.h.s. of (2.16) with $\delta=\epsilon^2$. This r.h.s. can readily be estimated if one knows the invariant probability measure on $\mathbb{S}^{\mathsf{L}}$ under the dynamics $(\mathbf{1}+\lambda rU)\cdot$ (it is again unique and given by the Furstenberg measure). The following proposition shows that this invariant measure is equal to the normalized invariant surface measure ν_{L} on $\mathbb{S}^{\mathsf{L}}$. In the terminology of [12, 11] this means that the dynamics $(\mathbf{1}+\lambda rU)\cdot$ has the so-called *random phase property*.

Proposition 2.14. *For all* $\lambda\in(0,\frac14]$ *and* $h\in\mathsf{L}^\infty(\mathbb{S}^{\mathsf{L}})$*, one has*

$$\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\,\mathbb{E}\,h\left((\mathbf{1}+\lambda rU)\cdot v\right)\;=\;\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\,h(v)\,. \quad (2.17)$$

At large N, the r.h.s. of (2.16) therefore approaches $\nu_{\mathsf{L}}(\{v\in\mathbb{S}^{\mathsf{L}}:\|\mathfrak{c}(v)\|^2<\delta\})$ (see [3], Part A, Theorem 4.3). Therefore the following geometric identity will be needed.

Lemma 2.15. *For all* $\delta\in[0,1]$*,*

$$\nu_{\mathsf{L}}\left(\Big\{v\in\mathbb{S}^{\mathsf{L}}:\|\mathfrak{c}(v)\|^2<\delta\Big\}\right)\;=\;\frac{\Gamma(\frac{\mathsf{L}+1}{2})}{\Gamma(\frac{\mathsf{L}_\mathfrak{c}}{2})\Gamma(\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}{2})}\int_0^\delta\mathrm{d}\mathbf{x}(x)\;x^{\frac{\mathsf{L}_\mathfrak{c}}{2}-1}(1-x)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}{2}-1}\,, \quad (2.18)$$

which just means that $\|\mathfrak{c}(v)\|^2$ *is distributed according to the beta distribution with parameters* $(\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}{2},\frac{\mathsf{L}_\mathfrak{c}}{2})$*. For* $(\mathsf{L}_\mathfrak{a},\mathsf{L}_\mathfrak{b})\neq(1,1)$ *this can, moreover, be bounded as follows:*

$$\nu_{\mathsf{L}}\left(\Big\{v\in\mathbb{S}^{\mathsf{L}}:\|\mathfrak{c}(v)\|^2<\delta\Big\}\right)\;\leq\;\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}\Big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}{2}-1}\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{c}}\,\delta\Big)^{\frac{\mathsf{L}_\mathfrak{c}}{2}}\Big(1-\frac{\delta}{6}\Big)\,. \quad (2.19)$$

The following Corollary 2.16 combines Proposition 2.14 and Lemma 2.15 and concludes the transient focus on the special case of $\mathcal{R}=\mathbf{1}$.

Corollary 2.16. *Let* $(\mathsf{L}_\mathfrak{a},\mathsf{L}_\mathfrak{b})\neq(1,1)$ *and* $\delta\in(0,1)$*. Then there exist* $\tilde N_0=\tilde N_0(\mathsf{L},\mathsf{L}_\mathfrak{c},\delta)\in\mathbb{N}$ *and* $\eta=\eta(\mathsf{L},\mathsf{L}_\mathfrak{c},\delta)>0$ *such that*

$$\mathbb{P}\Big(\Big\|\mathfrak{c}\Big(\prod_{n=1}^{N}(\mathbf{1}+\lambda r_n U_n)\cdot v\Big)\Big\|^2<\delta\Big)\;\leq\;\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}\Big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}{2}-1}\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{c}}\,\delta\Big)^{\frac{\mathsf{L}_\mathfrak{c}}{2}}\;-\;\eta \quad (2.20)$$

holds for all $N\geq\tilde N_0$ *and* $v\in\mathbb{S}^{\mathsf{L}}$*.*

Lemmata 2.8, 2.9, 2.10 and Corollaries 2.13 and 2.16 now allow to conclude.

Proof of Theorem 1.2. Let $\delta\in(0,1)$ and $\eta=\eta(\mathsf{L},\mathsf{L}_\mathfrak{c},\delta)>0$ and $\tilde N_0=\tilde N_0(\mathsf{L},\mathsf{L}_\mathfrak{c},\delta)\in\mathbb{N}$ be as in Corollary 2.16. Moreover, let us choose $\tilde N\geq\tilde N_0$. Then Lemmata 2.8, 2.9, 2.10 and Corollaries 2.13 and 2.16 imply the estimate

$$\begin{aligned}
\mathbb{E}\,\|\mathfrak{a}(v_{\tilde N+1})\|^2\;&=\;\mathbb{E}\,\|\mathfrak{a}(\mathcal{R}(\mathbf{1}+\lambda r_{\tilde N+1}U_{\tilde N+1})\cdot v_{\tilde N})\|^2\\
&\leq\;\big(1-\frac{\gamma\delta}{2}\big)\,\mathbb{E}\,\|\mathfrak{a}((\mathbf{1}+\lambda r_{\tilde N+1}U_{\tilde N+1})\cdot v_{\tilde N})\|^2+\frac{\gamma\delta}{2}\,\mathbb{P}\big(\|\mathfrak{c}((\mathbf{1}+\lambda r_{\tilde N+1}U_{\tilde N+1})\cdot v_{\tilde N})\|^2<\delta\big)\\
&\leq\;\big(1-\frac{\gamma\delta}{2}\big)\Big[\mathbb{E}\,\|\mathfrak{a}(v_{\tilde N})\|^2+\lambda^2\,\frac{3\,\mathsf{L}_\mathfrak{a}}{\mathsf{L}+1}\Big]\;+\;\Big[\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}\Big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}{2}-1}\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{c}}\,\delta\Big)^{\frac{\mathsf{L}_\mathfrak{c}}{2}}\;-\;\eta\Big]\frac{\gamma\delta}{2}\\
&\leq\;\big(1-\frac{\gamma\delta}{2}\big)\,\mathbb{E}\,\|\mathfrak{a}(v_{\tilde N})\|^2\;+\;M_\delta\;-\;\frac{\gamma\delta\eta}{2}\,,
\end{aligned}$$

where

$$M_\delta \;=\; \lambda^2\,\frac{3\,\mathsf{L}_\mathfrak{a}}{\mathsf{L}+1}\;+\;\frac{\gamma\mathsf{L}_\mathfrak{c}}{2(\mathsf{L}+1)}\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}\Big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}{2}-1}\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{c}}\,\delta\Big)^{\frac{\mathsf{L}_\mathfrak{c}}{2}+1}\,.$$

An iterative application of this inequality from $\tilde N=\tilde N_0$ to $N-1$ yields

$$\begin{aligned}\mathbb{E}\|\mathfrak{a}(v_N)\|^2 \;&\le\; \big(1-\frac{\gamma\delta}{2}\big)^{N-\tilde N_0}\mathbb{E}\|\mathfrak{a}(v_{\tilde N_0})\|^2\;+\;\big[M_\delta\;-\;\frac{\gamma\delta\eta}{2}\big]\sum_{\tilde N=\tilde N_0}^{N-1}\big(1-\frac{\gamma\delta}{2}\big)^{\tilde N-\tilde N_0}\\ &\le\;\big(1-\frac{\gamma\delta}{2}\big)^{N-\tilde N_0}\;+\;\frac{2}{\gamma\delta}\,\big[M_\delta\;-\;\frac{\gamma\delta\eta}{2}\big]\end{aligned}$$

for all $N\ge\tilde N_0$. Thus for all

$$N\;\ge\;\tilde N_0\;+\;\frac{\log(\eta)}{\log\big(1-\frac{\gamma\delta}{2}\big)}$$

one has

$$\mathbb{E}\|\mathfrak{a}(v_N)\|^2\;\le\;\frac{2\,M_\delta}{\gamma\delta}\,. \tag{2.21}$$

Now, the right side of (1.5) is larger than 1 if

$$\mathtt{d}\;=\;\frac{\mathsf{L}_\mathfrak{c}}{\mathsf{L}+1}\,\Big(\frac{6\lambda^2\mathsf{L}_\mathfrak{a}}{\gamma\mathsf{L}_\mathfrak{c}}\Big)^{\frac{2}{\mathsf{L}_\mathfrak{c}+2}}\,\Big(\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}{\mathsf{L}+1}\Big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-2}{\mathsf{L}_\mathfrak{c}+2}}$$

satisfies $\mathtt{d}\ge 1$. If this is violated, the choice $\delta=\mathtt{d}$ is possible and optimizes the order of the right side of (2.21) in λ and proves (1.5). □

3 Details of the proof of Theorem 1.1

Proof of Lemma 2.1. The first item is obvious. As for the dependence of the distribution of $\langle v,(\mathbf{1}+\lambda rU)\cdot v\rangle$ on $v\in\mathbb{S}^L$, let $w\in\mathbb{S}^\mathsf{L}$ and $\mathscr{W}\in\mathsf{O}(\mathsf{L}+1)$ be such that $\mathscr{W}w=v$. For all $s\in[0,1]$ and $\mathscr{U}\in\mathsf{O}(\mathsf{L}+1)$ one has

$$\begin{aligned}\langle v,(\mathbf{1}+\lambda s\mathscr{U})\cdot v\rangle\;&=\;\|(\mathbf{1}+\lambda s\mathscr{U})v\|^{-1}\;\langle v,(\mathbf{1}+\lambda s\mathscr{U})v\rangle\\ &=\;\|(\mathbf{1}+\lambda s\mathscr{W}^*\mathscr{U}\mathscr{W})w\|^{-1}\;\langle w,(\mathbf{1}+\lambda s\mathscr{W}^*\mathscr{U}\mathscr{W})w\rangle\\ &=\;\langle w,(\mathbf{1}+\lambda s\mathscr{W}^*\mathscr{U}\mathscr{W})\cdot w\rangle\;,\end{aligned}$$

but $\mathscr{W}^*U\mathscr{W}$ is distributed identically to U due to the invariance of the Haar measure. □

Proof of Lemma 2.2. The normalized surface measure ν_L on $\mathbb{S}^\mathsf{L}$ is equal to the push-forward $(\tau_v)_*(\theta_\mathsf{L})=\theta_\mathsf{L}\circ(\tau_v)^{-1}$ of the Haar measure θ_L on $\mathsf{O}(\mathsf{L}+1)$ under the map $\tau_v:\mathsf{O}(\mathsf{L}+1)\to\mathbb{S}^\mathsf{L}$ given by $\tau_v(U)=Uv$, independently of the choice of $v\in\mathbb{S}^\mathsf{L}$ (see [9], Chapter 3). Considering, moreover, the projection $\varsigma_v:\mathbb{S}^\mathsf{L}\to\mathbb{R}$ into the direction v given by $\varsigma_v(w)=\langle v,w\rangle$, it is also known that the push-forward $(\varsigma_v)_*(\nu_\mathsf{L})=\nu_\mathsf{L}\circ(\varsigma_v)^{-1}$ of ν_L is a.c. w.r.t. $\mathbf{x}$ with a Radon-Nikodym density given by

$$\frac{\Gamma(\frac{\mathsf{L}+1}{2})}{\sqrt{\pi}\,\Gamma(\frac{\mathsf{L}}{2})}\,\big(1-(\cdot)^2\big)^{\frac{\mathsf{L}}{2}-1}\,\chi_{[-1,1]}\,. \tag{3.1}$$

The action (1.2) can be spelled out explicitly in terms of the random variable $Y=\langle v,Uv\rangle$ as

$$\langle v,(\mathbf{1}+\lambda rU)\cdot v\rangle\;=\;\frac{1+\lambda rY}{\sqrt{1+2\lambda rY+\lambda^2r^2}}\,. \tag{3.2}$$

Thus let us denote the r.h.s. of (3.2) by $G(r,Y)$. As Y is distributed according to $(\varsigma_v)_*\,(\nu_{\mathsf{L}})$, it is a.c. w.r.t. $\mathbf{x}$ on $[-1,1]$. For $s\in(0,1]$, let $H^s_\pm:\big[\sqrt{1-\lambda^2s^2},1\big]\to\mathbb{R}$ be the two inverse branches of $y\mapsto G(s,y)$. They are given by

$$H^s_\pm(z)\;=\;\Big[z^2-1\pm z\big(z^2-1+\lambda^2s^2\big)^{\frac12}\Big](\lambda s)^{-1}\;.$$

For every $\mathbf{x}$-nullset $\aleph$ one thus has

$$G(s,\cdot)^{-1}(\aleph)\;=\;\bigcup_{\sigma=\pm}H^s_\sigma\big(\aleph\cap\big[\sqrt{1-\lambda^2s^2},1\big]\big)\,.$$

As a consequence,

$$\begin{aligned}G(s,\cdot)^{-1}(\aleph)\;&=\;H^s_\pm\big(\aleph\cap\big\{\sqrt{1-\lambda^2s^2}\big\}\big)\;\cup\;\bigcup_{\sigma=\pm}H^s_\sigma\big(\aleph\cap\big(\sqrt{1-\lambda^2s^2},1\big]\big)\\&=\;\big(H^s_\pm(\aleph)\cap\{-\lambda s\}\big)\;\cup\;\bigcup_{\sigma=\pm}H^s_\sigma\big(\aleph\cap\big(\sqrt{1-\lambda^2s^2},1\big]\big)\;.\end{aligned}$$

Now $H^s_\pm$ are locally Lipschitz continuous on $\big(\sqrt{1-\lambda^2s^2},1\big]$. This implies that also $G(s,\cdot)^{-1}(\aleph)$ is an $\mathbf{x}$-nullset. Due to the absolutely continuous distribution of Y w.r.t. $\mathbf{x}$, one therefore has

$$\begin{aligned}\mathbb{P}\left(G(r,Y)\in\aleph\right)&=\int\mathrm{d}\mathbb{P}(r\in\cdot)(s)\;\mathbb{P}\left(G(s,Y)\in\aleph\right)\\&=\;\int_{(0,1]}\mathrm{d}\mathbb{P}(r\in\cdot)(s)\;\mathbb{P}\left(Y\in G(s,\cdot)^{-1}(\aleph)\right)+\chi_\aleph(1)\;\mathbb{P}(r=0)\\&=\;\chi_\aleph(1)\;\mathbb{P}(r=0)\\&=\;\mathbb{P}(r=0)\;\delta_1(\aleph)\;,\end{aligned}$$

and this concludes the proof. □

Proof of Lemma 2.3. Let $(s,\mathscr{U})\in[0,1]\times\mathsf{O}(\mathsf{L}+1)$. Then,

$$\begin{aligned}(\mathbf{1}+\lambda s\mathscr{U})\cdot v\;&=\;\|(\mathbf{1}+\lambda s\mathscr{U})v\|^{-1}\,(\mathbf{1}+\lambda s\mathscr{U})v\\&=\;\|\mathscr{V}(\mathbf{1}+\lambda s\mathscr{U})v\|^{-1}\,\mathscr{V}^*\mathscr{V}(\mathbf{1}+\lambda s\mathscr{U})v\\&=\;\|(\mathbf{1}+\lambda s\mathscr{V}\mathscr{U})v\|^{-1}\,\mathscr{V}^*(\mathbf{1}+\lambda s\mathscr{V}\mathscr{U})v\\&=\;\mathscr{V}^*(\mathbf{1}+\lambda s\mathscr{V}\mathscr{U})\cdot v\end{aligned}$$

holds. But $\mathscr{V}U$ is distributed identically to U and this implies (2.2). As $(\mathbf{1}+\lambda s\mathscr{U})\cdot Vv=V(\mathbf{1}+\lambda sV^*\mathscr{U}V)\cdot v$, the proof of (2.1) follows in a similar manner. □

Proof of Lemma 2.4. By Lemma 2.2, the pure point part of the probability distribution $\varpi_{r,\lambda}$ of the $[\sqrt{1-\lambda^2},1]$-valued random variable $Z=\langle v,(\mathbf{1}+\lambda rU)\cdot v\rangle$ is given by $\varpi^{\mathrm{pp}}_{r,\lambda}=\mathbb{P}(r=0)\,\delta_1$. This implies the first equality in (2.3), since $Z=1$ is equivalent to $(\mathbf{1}+\lambda rU)\cdot v=v$.

As for the continuous part of $\varrho_{r,\lambda,v}$, let us write

$$(\mathbf{1}+\lambda rU)\cdot v\;=\;Z\,v\;+\;(1-Z^2)^{\frac12}\,v_\perp\;,\tag{3.3}$$

where $v_\perp\in\mathbb{S}^{\mathsf{L}}$ is a random unit vector orthogonal to v. By Lemma 2.3, the distribution of $v_\perp$ is invariant under the fixed point group of v, namely the action of $\{\mathscr{V}\in\mathsf{O}(\mathsf{L}+1)\,:\,\mathscr{V}v=v\}$. Thus the distribution of $v_\perp$ is given by the push-forward of $(i_v)_*(\nu_{\mathsf{L}-1})$ under a natural embedding $i_v:\mathbb{S}^{\mathsf{L}-1}\to\{w\in\mathbb{S}^{\mathsf{L}}\,:\,w\perp v\}$. Furthermore, Z

and $v_\perp$ are independent. Indeed, by (3.2) Z only depends on the component $Y = \langle v, Uv\rangle$ of the vector Uv in the direction of v, while for $Z \neq 1$

$$v_\perp \;=\; \frac{P_\perp\big((\mathbf{1}+\lambda rU)\cdot v\big)}{(1-Z^2)^{\frac{1}{2}}} \;=\; \frac{P_\perp Uv}{\|P_\perp Uv\|}\;,$$

with $P_\perp$ being the projection onto the orthogonal complement of the span of v, so that $v_\perp$ only depends on the direction of the component of Uv orthogonal to v, which is independent of the component parallel to v.

Now by the above and Lemma 2.2 the distribution of $(Z, (i_v)^{-1}(v_\perp))$ is equal to $(\varpi^{\mathrm{pp}}_{r,\lambda} + \varpi^{\mathrm{ac}}_{r,\lambda}) \otimes \nu_{\mathsf{L}-1}$ and therefore

$$\varrho_{r,\lambda,v} \;=\; (F_v)_*\big(\varpi^{\mathrm{pp}}_{r,\lambda}\otimes\nu_{\mathsf{L}-1}\big) \;+\; (F_v)_*\big(\varpi^{\mathrm{ac}}_{r,\lambda}\otimes\nu_{\mathsf{L}-1}\big)\;, \tag{3.4}$$

where the function

$$F_v \;:\; \big[\sqrt{1-\lambda^2},1\big]\times\mathbb{S}^{\mathsf{L}-1} \;\to\; \mathbb{S}^{\mathsf{L}}\;, \qquad (z,w)\;\mapsto\; z\,v \;+\; (1-z^2)^{\frac{1}{2}}\,i_v(w)$$

maps the set $\{1\}\times\mathbb{S}^{\mathsf{L}-1}$ to the point v and the set $[\sqrt{1-\lambda^2},1)\times\mathbb{S}^{\mathsf{L}-1}$ bijectively onto $\big\{u\in\mathbb{S}^{\mathsf{L}} : \langle u,v\rangle \in [\sqrt{1-\lambda^2},1)\big\}$. Using

$$(F_v)_*\big(\varpi^{\mathrm{pp}}_{r,\lambda}\otimes\nu_{\mathsf{L}-1}\big)(\{v\}) \;=\; \mathbb{P}(r=0)$$

and

$$\begin{aligned}(F_v)_*\big(\varpi^{\mathrm{pp}}_{r,\lambda}\otimes\nu_{\mathsf{L}-1}\big)(\mathbb{S}^{\mathsf{L}}\setminus\{v\}) \;&=\; \big(\varpi^{\mathrm{pp}}_{r,\lambda}\otimes\nu_{\mathsf{L}-1}\big)\big([\sqrt{1-\lambda^2},1)\times\mathbb{S}^{\mathsf{L}-1}\big)\\ &=\; \varpi^{\mathrm{pp}}_{r,\lambda}\big([\sqrt{1-\lambda^2},1)\big)\;\nu_{\mathsf{L}-1}\big(\mathbb{S}^{\mathsf{L}-1}\big)\\ &=\; 0\;,\end{aligned}$$

combined with the first identity in (2.3), one infers that $(F_v)_*\big(\varpi^{\mathrm{pp}}_{r,\lambda}\otimes\nu_{\mathsf{L}-1}\big)$ and $\varrho^{\mathrm{pp}}_{r,\lambda,v}$ coincide. This, in turn, implies together with (3.4) that

$$(F_v)_*\big(\varpi^{\mathrm{ac}}_{r,\lambda}\otimes\nu_{\mathsf{L}-1}\big) \;=\; \varrho^{\mathrm{ac}}_{r,\lambda,v} \;+\; \varrho^{\mathrm{sc}}_{r,\lambda,v} \tag{3.5}$$

holds and (3.5) is continuous. Now, since the restriction of F_v to any compact subset of $[\sqrt{1-\lambda^2},1)\times\mathbb{S}^{\mathsf{L}-1}$ is bi-Lipschitz, the preimage of any ν_{L}-nullset contained in $\mathbb{S}^{\mathsf{L}}\setminus\{v\}$ under F_v is an $\mathbf{x}\otimes\nu_{\mathsf{L}-1}$-nullset and hence, in particular, a $\varpi^{\mathrm{ac}}_{r,\lambda}\otimes\nu_{\mathsf{L}-1}$-nullset. Therefore, (3.5) is even absoluely continuous, *i.e.*, the second identity in (2.3) holds.

As for the proof of (2.4), one may assume that $v\neq w$, as (2.4) is trivial otherwise. Now let $V\in\mathsf{O}(\mathsf{L}+1)$ such that $Vv=w$. Since $\langle v,Vv\rangle = \langle v,V^*v\rangle$, the vectors Vv and V^*v have the same projection in the direction of v. Hence there exists a $\mathscr{V}\in\mathsf{O}(\mathsf{L}+1)$ satisfying $\mathscr{V}v=v$ such that $\mathscr{V}V^*v=Vv$. Now by applying both (2.1) and (2.2) one deduces that for every ball $B_\epsilon(v)\subset\mathbb{S}^{\mathsf{L}}$ of radius $\epsilon>0$ around v

$$\begin{aligned}\varrho_{r,\lambda,w}(B_\epsilon(v)) \;&=\; \mathbb{P}\left((\mathbf{1}+\lambda rU)\cdot w\in B_\epsilon(v)\right)\\ &=\; \mathbb{P}\left((\mathbf{1}+\lambda rU)\cdot v\in \mathscr{V}V^*B_\epsilon(v)\right)\\ &=\; \mathbb{P}\left((\mathbf{1}+\lambda rU)\cdot v\in B_\epsilon(w)\right)\\ &=\; \varrho_{r,\lambda,v}(B_\epsilon(w))\,.\end{aligned}$$

If $\epsilon<\|v-w\|$ and due to (2.3), this is equivalent to

$$\varrho^{\mathrm{ac}}_{r,\lambda,w}(B_\epsilon(v)) \;=\; \varrho^{\mathrm{ac}}_{r,\lambda,v}(B_\epsilon(w))\;.$$

Taking the Radon-Nikodym derivatives now implies (2.4). □

Proof of Lemma 2.5. The map $\mathcal{R}^{-1}\cdot$ is Lipschitz because for all $v_1, v_2 \in \mathbb{S}^{\mathsf{L}}$ one has

$$\begin{aligned}\left\|\mathcal{R}^{-1}\cdot v_1 - \mathcal{R}^{-1}\cdot v_2\right\| \;&=\; \frac{1}{\|\mathcal{R}^{-1}v_2\|}\left\|\Big(\left\|\mathcal{R}^{-1}v_2\right\| - \left\|\mathcal{R}^{-1}v_1\right\|\Big)\frac{\mathcal{R}^{-1}v_1}{\|\mathcal{R}^{-1}v_1\|} + \mathcal{R}^{-1}(v_1-v_2)\right\| \\ &\leq\; \|\mathcal{R}\|\,\Big(\Big|\left\|\mathcal{R}^{-1}v_2\right\| - \left\|\mathcal{R}^{-1}v_1\right\|\Big| + \left\|\mathcal{R}^{-1}(v_1-v_2)\right\|\Big) \\ &\leq\; 2\,\|\mathcal{R}\|\,\left\|\mathcal{R}^{-1}(v_1-v_2)\right\| \\ &\leq\; 2\,\|\mathcal{R}\|\,\left\|\mathcal{R}^{-1}\right\|\,\|v_1-v_2\|\;.\end{aligned}$$

Thus $\mathcal{R}^{-1}\cdot\aleph$ is a ν_{L}-nullset for any ν_{L}-nullset $\aleph$, which implies the claim. □

Proof of Lemma 2.6. Let $\lambda < 2^{-4}\delta'\mathcal{R}_i$ where $\delta'\mathcal{R}_i = \min\{\delta\mathcal{R}_i, \frac{1}{2}\}$. Let us denote the orthogonal projections onto $\mathbb{R}^{\mathsf{L}+1-i}\times\{0\}^i$ and $\{0\}^{\mathsf{L}+1-i}\times\mathbb{R}^i$ by $\mathscr{P}_i^{\uparrow}$ and $\mathscr{P}_i^{\downarrow}$, respectively. One has the estimates

$$\begin{aligned}\|\mathscr{P}_i^{\uparrow}(\mathcal{R}\cdot w)\|^2 \;&=\; \Big(1 + \|\mathscr{P}_i^{\downarrow}\mathcal{R}w\|^2\|\mathscr{P}_i^{\uparrow}\mathcal{R}w\|^{-2}\Big)^{-1} \\ &\leq \Big(1 + \kappa_i^2\kappa_{i+1}^{-2}\|\mathscr{P}_i^{\downarrow}w\|^2\|\mathscr{P}_i^{\uparrow}w\|^{-2}\Big)^{-1} \\ &=\; \|\mathscr{P}_i^{\uparrow}w\|^2\,\Big(1 + \big(\kappa_i\kappa_{i+1}^{-1}+1\big)\,\delta\mathcal{R}_i\|\mathscr{P}_i^{\downarrow}w\|^2\Big)^{-1} \\ &\leq\; \|\mathscr{P}_i^{\uparrow}w\|^2\,\Big(1 + 2\delta\mathcal{R}_i\|\mathscr{P}_i^{\downarrow}w\|^2\Big)^{-1} \\ &\leq\; \|\mathscr{P}_i^{\uparrow}w\|^2\,\Big(1 - \delta'\mathcal{R}_i\|\mathscr{P}_i^{\downarrow}w\|^2\Big)\;,\end{aligned}$$

and

$$\begin{aligned}\|\mathscr{P}_i^{\uparrow}((\mathbf{1}+\lambda s\mathscr{U})\cdot w)\|^2 \;&=\; 1\;-\;\|\mathscr{P}_i^{\downarrow}(\mathbf{1}+\lambda s\mathscr{U})w\|^2\|(\mathbf{1}+\lambda s\mathscr{U})w\|^{-2} \\ &\leq\; 1\;-\;\|\mathscr{P}_i^{\downarrow}(\mathbf{1}+\lambda s\mathscr{U})w\|^2\left[2 - \|(\mathbf{1}+\lambda s\mathscr{U})w\|^2\right] \\ &=\; \|\mathscr{P}_i^{\uparrow}w\|^2\;+\;2\lambda s\left\langle\Big(\|\mathscr{P}_i^{\downarrow}(\mathbf{1}+\lambda s\mathscr{U})w\|^2 - \mathscr{P}_i^{\downarrow}\Big)w, \mathscr{U}w\right\rangle \\ &\qquad +\;\lambda^2s^2\left\langle\mathscr{P}_i^{\downarrow}w, \Big(\mathscr{P}_i^{\downarrow} + 2\lambda s\mathscr{U}\Big)w\right\rangle \\ &\leq\; \|\mathscr{P}_i^{\uparrow}w\|^2\;+\;2\lambda s(1+\lambda s)^2\;+\;\lambda^2s^2(1+2\lambda s) \\ &\leq\; \|\mathscr{P}_i^{\uparrow}w\|^2\;+\;\frac{7}{2}\lambda\end{aligned}$$

for all $w\in\mathbb{S}^{\mathsf{L}}$, $s\in[0,1]$ and $\mathscr{U}\in\mathsf{O}(\mathsf{L}+1)$. Combining these estimates leads to

$$\|\mathscr{P}_i^{\uparrow}((\mathbf{1}+\lambda s\mathscr{U})\mathcal{R}\cdot v)\|^2\;-\;(1-\lambda/2)\|\mathscr{P}_i^{\uparrow}v\|^2\;\leq\;\delta'\mathcal{R}_i\|\mathscr{P}_i^{\uparrow}v\|^4 - \delta'\mathcal{R}_i\|\mathscr{P}_i^{\uparrow}v\|^2 + 4\lambda\;, \tag{3.6}$$

holding for all $v\in\mathbb{S}^{\mathsf{L}}$, $s\in[0,1]$ and $\mathscr{U}\in\mathsf{O}(\mathsf{L}+1)$. For these parameters, (3.6) now implies the following statements:

(i) $\|\mathscr{P}_i^{\uparrow}v\|^2 \in \frac{1}{2}\left(1+[-1,1]\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right) \implies \|\mathscr{P}_i^{\uparrow}((\mathbf{1}+\lambda s\mathscr{U})\mathcal{R}\cdot v)\|^2 \leq (1-\frac{\lambda}{2})\|\mathscr{P}_i^{\uparrow}v\|^2\,,$

(ii) $\|\mathscr{P}_i^{\uparrow}v\|^2 < \frac{1}{2}\left(1-\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right) \implies \|\mathscr{P}_i^{\uparrow}((\mathbf{1}+\lambda s\mathscr{U})\mathcal{R}\cdot v)\|^2 < \frac{1}{2}\big(1-\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\big).$

As for the dynamics $\{v_n\}_{n\in\mathbb{N}}$ defined by (1.1) and (1.3), statements (i) and (ii) guarantee the existence of $N\in\mathbb{N}$ such that all $n\in\mathbb{N}$ satisfy

$$\mathbb{P}\left(\|\mathscr{P}_i^{\uparrow}v_{n+N}\|^2 < \tfrac{1}{2}\left(1-\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right)\;\middle|\;\|\mathscr{P}_i^{\uparrow}v_n\|^2 \leq \tfrac{1}{2}\left(1+\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right)\right)\;=\;1\;,$$

and thus

$$\mathbb{P}\left(\|\mathscr{P}_i^{\uparrow}v_{n+N}\|^2 < \tfrac{1}{2}\left(1-\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right)\right)\;\geq\;\mathbb{P}\left(\|\mathscr{P}_i^{\uparrow}v_n\|^2 \leq \tfrac{1}{2}\left(1+\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right)\right)\;.$$

Therefore,

$$\begin{aligned}
&\mu_{r,\lambda}\left(\left\{v\in\mathbb{S}^{\mathsf{L}}:\|\mathscr{P}_i^{\uparrow}v\|^2<\tfrac{1}{2}\left(1-\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right)\right\}\right)\\
&\quad=\int\limits_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\mu_{r,\lambda}(v_0)\,\mathbb{P}\left(\|\mathscr{P}_i^{\uparrow}v_N\|^2<\tfrac{1}{2}\left(1-\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right)\right)\\
&\quad\geq\int\limits_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\mu_{r,\lambda}(v_0)\,\chi_{\left\{v\in\mathbb{S}^{\mathsf{L}}:\|\mathscr{P}_i^{\uparrow}v\|^2\leq\frac{1}{2}\left(1+\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right)\right\}}(v_0)\\
&\quad=\mu_{r,\lambda}\left(\left\{v\in\mathbb{S}^{\mathsf{L}}:\|\mathscr{P}_i^{\uparrow}v\|^2\leq\tfrac{1}{2}\left(1+\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right)\right\}\right)\,,
\end{aligned}$$

which is equivalent to

$$\mu_{r,\lambda}\left(\left\{v\in\mathbb{S}^{\mathsf{L}}:\tfrac{1}{2}\left(1-\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right)<\|\mathscr{P}_i^{\uparrow}v\|^2\leq\tfrac{1}{2}\left(1+\sqrt{1-16\lambda(\delta'\mathcal{R}_i)^{-1}}\right)\right\}\right)\;=\;0\,,$$

which proves $\mathrm{supp}(\mu_{r,\lambda})\neq\mathbb{S}^{\mathsf{L}}$. □

Proof of Lemma 2.7. The aim is to construct $(s_n)_{n=1,\ldots,N}$ in $\mathrm{supp}(r)$ and $(\mathscr{U}_n)_{n=1,\ldots,N}$ in $\mathrm{O}(\mathsf{L}+1)$ such that for a given couple $u,w\in\mathbb{S}^{\mathsf{L}}$

$$\prod_{n=1}^{N}\mathcal{R}\left(\mathbf{1}+\lambda s_n\mathscr{U}_n\right)\cdot u\;=\;w\;,$$

where u is an initial condition which we may choose to be the stable fixed point $e_{\mathsf{L}+1}$, as the motion from some arbitrary u towards this stable fixed point $e_{\mathsf{L}+1}$ via a finite path is somewhat straightforward and is thus left to the reader. To accommodate notations, let us use the unit vectors $\tilde{e}_j=e_{\mathsf{L}+2-j}$ so that $\mathcal{R}\tilde{e}_j=\kappa_j\tilde{e}_j$. Then $w=\sum_{j=1}^{\mathsf{L}+1}w_j\tilde{e}_j=(w_1,\ldots,w_{\mathsf{L}+1})^{\intercal}$. Further let us introduce $\mathsf{K}=\max\{\mathsf{J}\in\{1,\ldots,\mathsf{L}+1\}:w_{\mathsf{J}}\neq0\}$.

<u>Step 1.</u> *There exist* $N_1\in\mathbb{N}_0$ *and* $(\mathscr{U}_n^{\pm})_{n=1,\ldots,N_1}$ *in* $\mathrm{O}(\mathsf{L}+1)$ *such that*

$$\prod_{n=1}^{N_1}\mathcal{R}\left(\mathbf{1}+\lambda\mathscr{U}_n^{\pm}\right)\cdot\tilde{e}_1\;=\;\pm\,\tilde{e}_{\mathsf{K}}\;.$$

One can assume $\mathsf{K}\neq1$ as the statement is trivial otherwise. Let us set

$$\mathscr{U}_1^{\pm}\;=\;\begin{pmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&&0&\pm1\\&&&1&0\end{pmatrix}.$$

Then $\mathcal{R}\left(\mathbf{1}+\lambda\mathscr{U}_1^{\pm}\right)\tilde{e}_1=\pm\kappa_2\lambda\tilde{e}_2+\kappa_1\tilde{e}_1$. Next for $n\in\{2,\ldots,N_1-1\}$ with N_1 to be chosen later, we choose $\mathscr{U}_n^{\pm}=\mathrm{diag}(1,\ldots,1,-1)$. It follows that $\mathcal{R}\left(\mathbf{1}+\lambda\mathscr{U}_n^{\pm}\right)=\mathrm{diag}(\kappa_{\mathsf{L}+1}(1+\lambda),\ldots,\kappa_2(1+\lambda),\kappa_1(1-\lambda))$ so that

$$\prod_{n=1}^{N_1-1}\mathcal{R}\left(\mathbf{1}+\lambda\mathscr{U}_n^{\pm}\right)\tilde{e}_1\;=\;\pm\left[\kappa_2(1+\lambda)\right]^{N_1-1}\kappa_2\lambda\tilde{e}_2+\left[\kappa_1(1-\lambda)\right]^{N_1-1}\kappa_1\tilde{e}_1\;.$$

The assumption on λ guarantees that $\kappa_2(1+\lambda)>\kappa_1(1-\lambda)$ and therefore one can choose N_1 such that

$$\left[\kappa_1(1-\lambda)\right]^{N_1-1}\kappa_1\;\leq\;\lambda\left(\left[\kappa_2(1+\lambda)\right]^{N_1-1}\kappa_2\lambda\right)\;.$$

Hence, there exists some $\epsilon \le \lambda$ such that the proportionality relation

$$\prod_{n=1}^{N_1-1} \mathcal{R}\left(\mathbf{1} + \lambda \mathscr{U}_n^{\pm}\right)\tilde{e}_1 \;\propto\; \pm\,\tilde{e}_2 \,+\, \epsilon\tilde{e}_1$$

holds. Now, one can choose $\mathscr{U}_{N_1}^{\pm}$ in such a way that

$$\left\langle \lambda\mathscr{U}_{N_1}^{\pm}(\pm\tilde{e}_2 + \epsilon\tilde{e}_1), \tilde{e}_1\right\rangle \;=\; -\,\epsilon$$

and

$$\left\langle \lambda\mathscr{U}_{N_1}^{\pm}(\pm\tilde{e}_2 + \epsilon\tilde{e}_1), \tilde{e}_{\mathsf{J}}\right\rangle \;=\; 0\;, \qquad \forall\;\; \mathsf{J} \in \{3, \ldots, \mathsf{L}+1\}\;,$$

are satisfied. It follows that

$$\left(\mathbf{1} + \lambda\mathscr{U}_{N_1}^{\pm}\right)\prod_{n=1}^{N_1-1} \mathcal{R}\left(\mathbf{1} + \lambda \mathscr{U}_n^{\pm}\right)\tilde{e}_1 \;\propto\; \pm\,\tilde{e}_2\;,$$

and thus

$$\prod_{n=1}^{N_1} \mathcal{R}\left(\mathbf{1} + \lambda \mathscr{U}_n^{\pm}\right)\cdot\tilde{e}_1 \;=\; \pm\,\tilde{e}_2$$

holds. In the same vein, one may construct paths from $\tilde{e}_{\mathsf{J}-1}$ to $\tilde{e}_{\mathsf{J}}$ for $\mathsf{J} \in \{3, \ldots, \mathsf{K}\}$. This finishes the proof of Step 1.

Next let us set $\tilde{\mathsf{K}} = \min\{\mathsf{J} \in \{1, \ldots, \mathsf{K}\} : \kappa_{\mathsf{J}} = \kappa_{\mathsf{K}}\}$.

Step 2. *There exist* $N_2 \in \mathbb{N}_0$*, sequences* $(s_n)_{n=1,\ldots,N_2}$ *in* $\mathrm{supp}(r)$ *and* $(\mathscr{U}_n)_{n=1,\ldots,N_2}$ *in* $\mathrm{O}(\mathsf{L}+1)$ *such that*

$$\prod_{n=1}^{N_2} \mathcal{R}\left(\mathbf{1} + \lambda s_n \mathscr{U}_n\right)\cdot\tilde{e}_{\mathsf{K}} \;=\; \sum_{\mathsf{J}=\tilde{\mathsf{K}}}^{\mathsf{K}} w_{\mathsf{J}}\tilde{e}_{\mathsf{J}}\Big\|\sum_{\mathsf{J}=\tilde{\mathsf{K}}}^{\mathsf{K}} w_{\mathsf{J}}\tilde{e}_{\mathsf{J}}\Big\|^{-1}\;. \tag{3.7}$$

Let $\tilde{U}_{\mathsf{K}-\tilde{\mathsf{K}}+1}$ be an $\mathrm{O}(\mathsf{K}-\tilde{\mathsf{K}}+1)$-valued random variable distributed according to the Haar measure. It induces an $\mathrm{O}(\mathsf{L}+1)$-valued random variable by $\tilde{U}_{\tilde{\mathsf{K}},\mathsf{K}} = \mathbf{1}_{\mathsf{L}+1-\mathsf{K}}\oplus\tilde{U}_{\mathsf{K}-\tilde{\mathsf{K}}+1}\oplus\mathbf{1}_{\tilde{\mathsf{K}}-1}$. Since $\kappa_{\mathsf{K}} = \cdots = \kappa_{\tilde{\mathsf{K}}}$, the action $\mathcal{R}\cdot$ is trivial on the submanifold $\mathscr{S}_{\tilde{\mathsf{K}},\mathsf{K}} = \{0\}^{\mathsf{L}+1-\mathsf{K}} \times \mathbb{S}^{\mathsf{K}-\tilde{\mathsf{K}}} \times \{0\}^{\tilde{\mathsf{K}}-1}$ and commutes with $(\mathbf{1} + \lambda r\tilde{U}_{\tilde{\mathsf{K}},\mathsf{K}})\cdot$ which acts transitively on $\mathscr{S}_{\tilde{\mathsf{K}},\mathsf{K}}$ (see Proposition 2.14 for a detailed proof). This shows Step 2. Combined with the above, the next step concludes the proof.

Step 3. *There exist* $N_3 \in \mathbb{N}_0$ *and* $(\mathscr{U}_n)_{n=1,\ldots,N_3}$ *in* $\mathrm{O}(\mathsf{L}+1)$ *such that*

$$\prod_{n=1}^{N_3} \mathcal{R}\left(\mathbf{1} + \lambda \mathscr{U}_n\right)\cdot\sum_{\mathsf{J}=\tilde{\mathsf{K}}}^{\mathsf{K}} w_{\mathsf{J}}\tilde{e}_{\mathsf{J}}\Big\|\sum_{\mathsf{J}=\tilde{\mathsf{K}}}^{\mathsf{K}} w_{\mathsf{J}}\tilde{e}_{\mathsf{J}}\Big\|^{-1} \;=\; w\;. \tag{3.8}$$

One can assume $\left(w_{\tilde{\mathsf{K}}-1}, \ldots, w_1\right)^{\mathsf{T}} \neq 0$, as the statement is trivial otherwise. Let us abbreviate

$$y \;=\; (w_{\mathsf{K}}, \ldots, w_{\tilde{\mathsf{K}}})^{\mathsf{T}}\Big\|\sum_{\mathsf{J}=\tilde{\mathsf{K}}}^{\mathsf{K}} w_{\mathsf{J}}\tilde{e}_{\mathsf{J}}\Big\|^{-1} \;\in\; \mathbb{R}^{\mathsf{K}-\tilde{\mathsf{K}}+1}$$

and use the notation $(x_{\mathsf{L}+1}, \ldots, x_1)^{\mathsf{T}} := \mathscr{U}_1(0, \ldots, 0, y, 0, \ldots 0)^{\mathsf{T}}$ with $\mathscr{U}_1$ to be chosen later. Set $\mathscr{U}_n = \mathbf{1}$ for $n \in \{2, \ldots, N_3\}$, where N_3 will also be chosen later. Now the l.h.s. of (3.8) is proportional to

$$\mathcal{R}^{N_3}\left(\mathbf{1} + \lambda\mathscr{U}_1\right)(0, \ldots, 0, y, 0, \ldots 0)^{\mathsf{T}} \;=\; \begin{pmatrix} \lambda(\kappa_{\mathsf{L}+1}^{N_3}x_{\mathsf{L}+1}, \ldots, \kappa_{\mathsf{K}+1}^{N_3}x_{\mathsf{K}+1})^{\mathsf{T}} \\ \kappa_{\mathsf{K}}^{N_3}\left(y + \lambda(x_{\mathsf{K}}, \ldots, x_{\tilde{\mathsf{K}}})^{\mathsf{T}}\right) \\ \lambda(\kappa_{\tilde{\mathsf{K}}-1}^{N_3}x_{\tilde{\mathsf{K}}-1}, \ldots, \kappa_1^{N_3}x_1)^{\mathsf{T}} \end{pmatrix}\;,$$

which in turn has to be proportional to w so that, for some $c \in (0,\infty)$,

$$\begin{pmatrix} \lambda(\kappa_{\mathsf{L}+1}^{N_3} x_{\mathsf{L}+1}, \dots, \kappa_{\mathsf{K}+1}^{N_3} x_{\mathsf{K}+1})^\intercal \\ \kappa_{\mathsf{K}}^{N_3} \left(y + \lambda(x_{\mathsf{K}}, \dots, x_{\tilde{\mathsf{K}}})^\intercal\right) \\ \lambda(\kappa_{\tilde{\mathsf{K}}-1}^{N_3} x_{\tilde{\mathsf{K}}-1}, \dots, \kappa_1^{N_3} x_1)^\intercal \end{pmatrix} \;=\; c \begin{pmatrix} (w_{\mathsf{L}+1}, \dots, w_{\mathsf{K}+1})^\intercal \\ (w_{\mathsf{K}}, \dots, w_{\tilde{\mathsf{K}}})^\intercal \\ (w_{\tilde{\mathsf{K}}-1}, \dots, w_1)^\intercal \end{pmatrix} \,. \tag{3.9}$$

Now $(w_{\mathsf{L}+1}, \dots, w_{\mathsf{K}+1})^\intercal = (0,\dots,0)^\intercal$ requires the choice $(x_{\mathsf{L}+1}, \dots, x_{\mathsf{K}+1})^\intercal = (0,\dots,0)^\intercal$. Moreover, since y is proportional to $(w_{\mathsf{K}}, \dots, w_{\tilde{\mathsf{K}}})^\intercal$, the middle part of (3.9) forces us to set

$$(x_{\mathsf{K}}, \dots, x_{\tilde{\mathsf{K}}})^\intercal \;=\; y\big(1 - x_1^2 - \dots - x_{\tilde{\mathsf{K}}-1}^2\big)^{\frac{1}{2}} \,,$$

where $x_{\tilde{\mathsf{K}}-1}, \dots, x_1$ are given by the lower part of (3.9) as

$$x_{\tilde{\mathsf{K}}-1} \;=\; \frac{c}{\lambda}\,\frac{w_{\tilde{\mathsf{K}}-1}}{\kappa_{\tilde{\mathsf{K}}-1}^{N_3}} \,, \quad \dots \,, \quad x_1 \;=\; \frac{c}{\lambda}\,\frac{w_1}{\kappa_1^{N_3}} \,,$$

where c and N_3 have still to be chosen appropriately in order to satisfy the remaining middle part, which is now of the (scalar) form

$$\kappa_{\mathsf{K}}^{N_3}\Big[1 + \lambda\Big(1 - \frac{c^2}{\lambda^2}\sum_{\mathsf{J}=1}^{\tilde{\mathsf{K}}-1}\Big(\frac{w_{\mathsf{J}}}{\kappa_{\mathsf{J}}^{N_3}}\Big)^2\Big)^{\frac{1}{2}}\Big] \;=\; c\,\big\|(w_{\mathsf{K}}, \dots, w_{\tilde{\mathsf{K}}})^\intercal\big\| \,.$$

It hence suffices to demonstrate the existence of some $N_3 \in \mathbb{N}$ such that the function

$$c \in (0,\infty) \;\mapsto\; f_{N_3}(c) \;=\; c\,\big\|(w_{\mathsf{K}}, \dots, w_{\tilde{\mathsf{K}}})^\intercal\big\| \;-\; \kappa_{\mathsf{K}}^{N_3}\Big[1 + \lambda\Big(1 - \frac{c^2}{\lambda^2}\sum_{\mathsf{J}=1}^{\tilde{\mathsf{K}}-1}\Big(\frac{w_{\mathsf{J}}}{\kappa_{\mathsf{J}}^{N_3}}\Big)^2\Big)^{\frac{1}{2}}\Big]$$

has a zero. As $f_{N_3}(\cdot)$ is continuous, it suffices to demonstrate that it attains both negative and positive values. It is obvious that $f_{N_3}(0) < 0$. Setting

$$c_{\max}(N_3) \;=\; \lambda\Big(\sum_{\mathsf{J}=1}^{\tilde{\mathsf{K}}-1}\Big(\frac{w_{\mathsf{J}}}{\kappa_{\mathsf{J}}^{N_3}}\Big)^2\Big)^{-1/2}$$

one observes that

$$f_{N_3}(c_{\max}(N_3)) \;=\; \kappa_{\mathsf{K}}^{N_3}\Big[\lambda\,\big\|(w_{\mathsf{K}}, \dots, w_{\tilde{\mathsf{K}}})^\intercal\big\|\,\Big(\sum_{\mathsf{J}=1}^{\tilde{\mathsf{K}}-1}\Big(w_{\mathsf{J}}\,\frac{\kappa_{\mathsf{K}}^{N_3}}{\kappa_{\mathsf{J}}^{N_3}}\Big)^2\Big)^{-1/2} - 1\Big] \,.$$

Since $\frac{\kappa_{\mathsf{K}}}{\kappa_{\mathsf{J}}} < 1$ for $\mathsf{J} \in \{1, \dots, \tilde{\mathsf{K}} - 1\}$, positive values are reached for sufficiently large N_3. □

4 Details of the proof of Theorem 1.2

This section contains the proofs of the preparatory lemmas for the proof of Theorem 1.2.

Proof of Lemma 2.8. Inequality (2.10) is obviously satisfied if $\mathfrak{a}(v) = 0$, as in this case $\mathfrak{a}(\mathcal{R} \cdot v) = 0$ holds. Now, let $\mathfrak{a}(v) \neq 0$. Then, its validity is demonstrated by the estimate

$$\begin{aligned}\|\mathfrak{a}(\mathcal{R}\cdot v)\|^2 &= \left(1+\frac{\|\mathfrak{b}(\mathcal{R}v)\|^2+\|\mathfrak{c}(\mathcal{R}v)\|^2}{\|\mathfrak{a}(\mathcal{R}v)\|^2}\right)^{-1}\\
&\le \left(1+\frac{(\kappa_{\mathsf{L}_\mathfrak{b}+\mathsf{L}_\mathfrak{c}+1})^2\|\mathfrak{b}(v)\|^2+(\kappa_{\mathsf{L}_\mathfrak{c}})^2\|\mathfrak{c}(v)\|^2}{(\kappa_{\mathsf{L}_\mathfrak{b}+\mathsf{L}_\mathfrak{c}+1})^2\|\mathfrak{a}(v)\|^2}\right)^{-1}\\
&= \|\mathfrak{a}(v)\|^2\left(1+\|\mathfrak{c}(v)\|^2\left[\left(\frac{\kappa_{\mathsf{L}_\mathfrak{c}}}{\kappa_{\mathsf{L}_\mathfrak{b}+\mathsf{L}_\mathfrak{c}+1}}\right)^2-1\right]\right)^{-1}\\
&\le \|\mathfrak{a}(v)\|^2\left(1+\|\mathfrak{c}(v)\|^2\min\left\{1,\left(\frac{\kappa_{\mathsf{L}_\mathfrak{c}}}{\kappa_{\mathsf{L}_\mathfrak{b}+\mathsf{L}_\mathfrak{c}+1}}\right)^2-1\right\}\right)^{-1}\\
&\le \|\mathfrak{a}(v)\|^2\left(1-\frac{\|\mathfrak{c}(v)\|^2}{2}\min\left\{1,\left(\frac{\kappa_{\mathsf{L}_\mathfrak{c}}}{\kappa_{\mathsf{L}_\mathfrak{b}+\mathsf{L}_\mathfrak{c}+1}}\right)^2-1\right\}\right),\end{aligned}$$

in which we used that $\|\mathfrak{a}(v)\|^2+\|\mathfrak{b}(v)\|^2+\|\mathfrak{c}(v)\|^2=1$ in the third step. Due to the definition (1.4) this implies the result. $\square$

Proof of Lemma 2.9. Let $\mathscr{U}\in\mathsf{O}(\mathsf{L}+1)$ and $s\in[0,1]$. We apply the bound $(1+x)^{-1}\le 1-x+2x^2$ for $x\ge-\frac{1}{2}$ to $x=2\lambda s\,\langle v,Uv\rangle+\lambda^2s^2$ where $\lambda\le\frac{1}{4}$. This yields the estimate

$$\begin{aligned}\|(\mathbf{1}+\lambda s\mathscr{U})\,v\|^{-2} &= \left(1+2\lambda s\,\langle v,\mathscr{U}v\rangle+\lambda^2s^2\right)^{-1}\\
&\le 1-\left(2\lambda s\,\langle v,\mathscr{U}v\rangle+\lambda^2s^2\right)+2\left(2\lambda s\,\langle v,\mathscr{U}v\rangle+\lambda^2s^2\right)^2\\
&= \left[1-\lambda^2s^2+2\lambda^4s^4\right]-2\lambda s(1-4\lambda^2s^2)\,\langle v,\mathscr{U}v\rangle+8\lambda^2s^2\,\langle v,\mathscr{U}v\rangle^2\;.\end{aligned}$$

As any term of odd order in the entries of U is centered, this implies for the average over U

$$\begin{aligned}\mathbb{E}\,\|\mathfrak{a}((\mathbf{1}+\lambda sU)\cdot v)\|^2 \;&=\; \mathbb{E}\,\|\mathfrak{a}\left((\mathbf{1}+\lambda sU)\,v\right)\|^2\,\|(\mathbf{1}+\lambda sU)\,v\|^{-2}\\
&\le\; \left[1-\lambda^2s^2+2\lambda^4s^4\right]\left(\|\mathfrak{a}(v)\|^2+\lambda^2s^2\,\mathbb{E}\,\|\mathfrak{a}(Uv)\|^2\right)\\
&\quad-4\,\lambda^2s^2(1-4\lambda^2s^2)\,\mathbb{E}\,\langle\mathfrak{a}(Uv),\mathfrak{a}(v)\rangle\,\langle v,Uv\rangle\\
&\quad+\;8\lambda^2s^2\left(\|\mathfrak{a}(v)\|^2\,\mathbb{E}\,\langle v,Uv\rangle^2+\lambda^2s^2\,\mathbb{E}\,\|\mathfrak{a}(Uv)\|^2\,\langle v,Uv\rangle^2\right)\;.\end{aligned}$$

The averages on the r.h.s. can now be evaluated explicitly, *e.g.*, using Lemma 2 in [8],

$$\begin{aligned}\mathbb{E}\,\|\mathfrak{a}(Uv)\|^2 \;&=\; \mathbb{E}\,\mathrm{tr}\left[U^*\begin{pmatrix}\mathbf{1}_{\mathsf{L}_\mathfrak{a}} & 0\\ 0 & 0\end{pmatrix}U|v\rangle\langle v|\right] \;=\; \frac{\mathsf{L}_\mathfrak{a}}{\mathsf{L}+1}\;,\\
\mathbb{E}\,\langle v,Uv\rangle^2 \;&=\; \mathbb{E}\,\mathrm{tr}\,[U^*|v\rangle\langle v|U|v\rangle\langle v|] \;=\; \frac{1}{\mathsf{L}+1}\,,\\
\mathbb{E}\,\|\mathfrak{a}(Uv)\|^2\,\langle v,Uv\rangle^2 \;&=\; \mathbb{E}\,\mathrm{tr}\left[U^*\begin{pmatrix}\mathbf{1}_{\mathsf{L}_\mathfrak{a}} & 0\\ 0 & 0\end{pmatrix}U|v\rangle\langle v|U^*|v\rangle\langle v|U|v\rangle\langle v|\right] \;=\; \frac{\mathsf{L}_\mathfrak{a}+2\,\|\mathfrak{a}(v)\|^2}{(\mathsf{L}+1)\,(\mathsf{L}+3)}\,,\\
\mathbb{E}\,\langle\mathfrak{a}(Uv),\mathfrak{a}(v)\rangle\,\langle v,Uv\rangle \;&=\; \mathbb{E}\,\mathrm{tr}\left[U^*\begin{pmatrix}\mathbf{1}_{\mathsf{L}_\mathfrak{a}} & 0\\ 0 & 0\end{pmatrix}|v\rangle\langle v|U|v\rangle\langle v|\right] \;=\; \frac{\|\mathfrak{a}(v)\|^2}{\mathsf{L}+1}\;.\end{aligned}$$

We obtain

$$\begin{aligned}
\mathbb{E}\,\|\mathfrak{a}((\mathbf{1}+\lambda sU)\cdot v)\|^2 \;&\leq\; \left[1-\lambda^2 s^2+2\lambda^4 s^4\right]\left(\|\mathfrak{a}(v)\|^2+\lambda^2 s^2\,\frac{\mathsf{L}_\mathfrak{a}}{\mathsf{L}+1}\right)\\
&\quad -4\lambda^2 s^2(1-4\lambda^2 s^2)\,\frac{\|\mathfrak{a}(v)\|^2}{\mathsf{L}+1}\\
&\quad +8\lambda^2 s^2\left(\frac{\|\mathfrak{a}(v)\|^2}{\mathsf{L}+1}+\lambda^2 s^2\,\frac{\mathsf{L}_\mathfrak{a}+2\,\|\mathfrak{a}(v)\|^2}{(\mathsf{L}+1)\,(\mathsf{L}+3)}\right)\\
&= \left(1-\lambda^2 s^2\,\frac{\mathsf{L}-3}{\mathsf{L}+1}+\lambda^4 s^4\,\frac{16(\mathsf{L}+4)}{(\mathsf{L}+1)(\mathsf{L}+3)}\right)\|\mathfrak{a}(v)\|^2\\
&\quad +\left(1-\lambda^2 s^2\,\frac{\mathsf{L}-5}{\mathsf{L}+3}+2\lambda^4 s^4\right)\lambda^2 s^2\,\frac{\mathsf{L}_\mathfrak{a}}{\mathsf{L}+1}\,.
\end{aligned}$$

This, in turn, implies (2.11), since $\lambda \leq \frac{1}{4}$ and $\mathsf{L} \geq 3$. □

Proof of Lemma 2.10. Using conditional expectations, one obtains the estimate

$$\begin{aligned}
\mathbb{E}\,\|\mathfrak{a}(v)\|^2\,\|\mathfrak{c}(v)\|^2 \;&\geq\; \mathbb{E}\big(\|\mathfrak{a}(v)\|^2\,\|\mathfrak{c}(v)\|^2 \;\big|\; \|\mathfrak{c}(v)\|^2 \geq \delta\big)\;\mathbb{P}\big(\|\mathfrak{c}(v)\|^2 \geq \delta\big)\\
&\geq\; \delta\;\mathbb{E}\big(\|\mathfrak{a}(v)\|^2 \;\big|\; \|\mathfrak{c}(v)\|^2 \geq \delta\big)\;\mathbb{P}\big(\|\mathfrak{c}(v)\|^2 \geq \delta\big)\\
&=\; \delta\left[\mathbb{E}\,\|\mathfrak{a}(v)\|^2-\mathbb{E}\big(\|\mathfrak{a}(v)\|^2 \;\big|\; \|\mathfrak{c}(v)\|^2 < \delta\big)\;\mathbb{P}\big(\|\mathfrak{c}(v)\|^2 < \delta\big)\right]\\
&\geq\; \delta\left[\mathbb{E}\,\|\mathfrak{a}(v)\|^2-\mathbb{P}\big(\|\mathfrak{c}(v)\|^2 < \delta\big)\right]\;.
\end{aligned}$$

This proves (2.12). □

Proof of Proposition 2.11. The proof is split into two intermediate steps. The first one is similar to Lemma 2.3:

Step 1. *Let $v,w \in \mathbb{S}^{\mathsf{L}}$ satisfy $\|\mathfrak{c}(v)\| = \|\mathfrak{c}(w)\|$. Then the random variables $\|\mathfrak{c}\,((\mathbf{1}+\lambda rU)\cdot v)\|$ and $\|\mathfrak{c}\,((\mathbf{1}+\lambda rU)\cdot w)\|$ are distributed identically, that is*

$$\mathbb{P}\left(\|\mathfrak{c}\,((\mathbf{1}+\lambda rU)\cdot v)\| \in \cdot\right) \;=\; \mathbb{P}\left(\|\mathfrak{c}\,((\mathbf{1}+\lambda rU)\cdot w)\| \in \cdot\right)\;. \tag{4.1}$$

For the proof, let us first note that the assumption of $\|\mathfrak{c}(v)\| = \|\mathfrak{c}(w)\|$ guarantees the existence of $(\mathscr{W}_1,\mathscr{W}_2) \in \mathsf{O}(\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}) \times \mathsf{O}(\mathsf{L}_\mathfrak{c})$ such that $\mathscr{W} = \mathscr{W}_1 \oplus \mathscr{W}_2 \in \mathsf{O}(\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}) \oplus \mathsf{O}(\mathsf{L}_\mathfrak{c})$ satisfies $w = \mathscr{W} v$. Next let $\tilde{\mathsf{L}}$ be either equal to $\mathsf{L}_\mathfrak{c}$ or equal to $\mathsf{L}+1$. Furthermore let $\mathscr{P}_{\tilde{\mathsf{L}}}$ denote the orthogonal projection onto $\{0\}^{\mathsf{L}+1-\tilde{\mathsf{L}}} \times \mathbb{R}^{\tilde{\mathsf{L}}}$. It is obvious that $\mathscr{W}^*$ commutes with $\mathscr{P}_{\tilde{\mathsf{L}}}$. Hence, all $(s,\mathscr{U}) \in [0,1] \times \mathsf{O}(\mathsf{L}+1)$ obey

$$\left\|\mathscr{P}_{\tilde{\mathsf{L}}}(\mathbf{1}+\lambda s\mathscr{U})w\right\|^2 \;=\; \left\|\mathscr{W}^*\mathscr{P}_{\tilde{\mathsf{L}}}(\mathbf{1}+\lambda s\mathscr{U})\mathscr{W} v\right\|^2 \;=\; \left\|\mathscr{P}_{\tilde{\mathsf{L}}}(\mathbf{1}+\lambda s\mathscr{W}^*\mathscr{U}\mathscr{W})v\right\|^2\;.$$

This identity is now used in the third equality in the following calculation:

$$\begin{aligned}
\|\mathfrak{c}\,((\mathbf{1}+\lambda s\mathscr{U})\cdot w)\| \;&=\; \|\mathfrak{c}\,((\mathbf{1}+\lambda s\mathscr{U})w)\|\;\|(\mathbf{1}+\lambda s\mathscr{U})w\|^{-1}\\
&=\; \|\mathscr{P}_{\mathsf{L}_\mathfrak{c}}(\mathbf{1}+\lambda s\mathscr{U})w\|\;\|\mathscr{P}_{\mathsf{L}+1}(\mathbf{1}+\lambda s\mathscr{U})w\|^{-1}\\
&=\; \|\mathscr{P}_{\mathsf{L}_\mathfrak{c}}(\mathbf{1}+\lambda s\mathscr{W}^*\mathscr{U}\mathscr{W})v\|\;\|\mathscr{P}_{\mathsf{L}+1}(\mathbf{1}+\lambda s\mathscr{W}^*\mathscr{U}\mathscr{W})v\|^{-1}\\
&=\; \|\mathfrak{c}\,((\mathbf{1}+\lambda s\mathscr{W}^*\mathscr{U}\mathscr{W})v)\|\;\|(\mathbf{1}+\lambda s\mathscr{W}^*\mathscr{U}\mathscr{W})v\|^{-1}\\
&=\; \|\mathfrak{c}\,((\mathbf{1}+\lambda s\mathscr{W}^*\mathscr{U}\mathscr{W})\cdot v)\|\;.
\end{aligned}$$

But $\mathscr{W}^*U\mathscr{W}$ is distributed identically to U so that (4.1) and thus Step 1 follows.

In view of Step 1, (2.14) is equivalent to the existence of a path $\phi : [0,1] \to \mathbb{S}^{\mathsf{L}}$ such that $\|\mathfrak{c}(\cdot)\| \circ \phi : [0,1] \to [0,1]$ is non-decreasing and surjective and that for all $\epsilon \in [0,1]$ the map $t \mapsto \mathbb{P}\big(\|\mathfrak{c}\,((\mathbf{1} + \lambda r U) \cdot \phi(t))\| \leq \epsilon\big)$ is non-increasing. Hence the proof of the lemma is completed by the following

Step 2. *The map $f_\epsilon : [0, \frac{\pi}{2}] \mapsto [0,1]$ defined by*

$$f_\epsilon(t) \;=\; \mathbb{P}\big(\|\mathfrak{c}\,((\mathbf{1} + \lambda r U) \cdot (\cos(t), 0, \ldots, 0, \sin(t))^{\mathsf{T}})\| \leq \epsilon\big) \tag{4.2}$$

is non-increasing for all $\epsilon \in [0,1]$.

To prove this monotonicity property, it is not necessary to calculate the probability explicitly, but only proportionality is needed. As mentioned in the proof of Theorem 1.1, the normalized surface measure ν_{L} is distributed identically to the pushforward $(h_v)_*\,(\theta_{\mathsf{L}})$ of the Haar measure θ_{L} on $\mathrm{O}(\mathsf{L}+1)$ under the map $h_v : \mathrm{O}(\mathsf{L}+1) \to \mathbb{S}^{\mathsf{L}}$ given by $h_v(U) = Uv$ for any $v \in \mathbb{S}^{\mathsf{L}}$. Thus, $(\mathbf{1} + \lambda r U) \cdot (\cos(t), 0, \ldots, 0, \sin(t))^{\mathsf{T}}$ is distributed identically to

$$\frac{(\cos(t) + \lambda r z_1, \lambda r z_2, \ldots, \lambda r z_{\mathsf{L}}, \sin(t) + \lambda r z_{\mathsf{L}+1})^{\mathsf{T}}}{\|(\cos(t) + \lambda r z_1, \lambda r z_2, \ldots, \lambda r z_{\mathsf{L}}, \sin(t) + \lambda r z_{\mathsf{L}+1})^{\mathsf{T}}\|}\,,$$

where $(z_1, \ldots, z_{\mathsf{L}+1})^{\mathsf{T}}$ is assumed to be distributed according to ν_{L}. It follows that

$$\mathbb{P}\big(\|\mathfrak{c}\,((\mathbf{1} + \lambda r U) \cdot (\cos(t), 0, \ldots, 0, \sin(t))^{\mathsf{T}})\| \leq \epsilon\big) = \mathbb{P}\left(W^t_\epsilon(r, z_1, z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_{\mathsf{L}+1}) \leq 0\right)\,,$$

where

$$\begin{aligned} W^t_\epsilon(r, z_1, z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_{\mathsf{L}+1}) \;&= \lambda^2 r^2 z^2_{\mathsf{L}+1} + 2\lambda r(1-\epsilon^2)\sin(t) z_{\mathsf{L}+1} \;+\; \sin^2(t) \\ &\quad + \lambda^2 r^2 \|(z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_{\mathsf{L}})^{\mathsf{T}}\|^2 - \big(1 + \lambda^2 r^2 + 2\lambda r \cos(t) z_1\big)\,\epsilon^2\,. \end{aligned}$$

Now $W^t_\epsilon(r, z_1, z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_{\mathsf{L}+1})$ is a parabola in $z_{\mathsf{L}+1}$ with unique minimum. It attains non-positive values if and only if

$$\big[1 + \lambda^2 r^2 + 2\lambda r \cos(t) z_1 - (2-\epsilon^2)\sin^2(t)\big]\,\epsilon^2 \;\geq\; \lambda^2 r^2 \|(z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_{\mathsf{L}})^{\mathsf{T}}\|^2\,. \tag{4.3}$$

Let us use the notation $n_\mathfrak{c}(z) = \|(z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_{\mathsf{L}})^{\mathsf{T}}\|^2$. If (4.3) holds, then the inequality $W^t_\epsilon(r, z_1, z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_{\mathsf{L}+1}) \leq 0$ is equivalent to

$$a^t_{\epsilon,-}(r, z_1, n_\mathfrak{c}(z)) \;\leq\; \lambda r z_{\mathsf{L}+1} \;\leq\; a^t_{\epsilon,+}(r, z_1, n_\mathfrak{c}(z))\,, \tag{4.4}$$

where the two roots of the polynomial are

$$a^t_{\epsilon,\pm}(r, z_1, n_\mathfrak{c}(z)) \;=\; (\epsilon^2-1)\sin(t) \pm \Big[(1+\lambda^2 r^2 + 2\lambda r\cos(t) z_1 - (2-\epsilon^2)\sin^2(t))\epsilon^2 - \lambda^2 r^2 n_\mathfrak{c}(z)\Big]^{\frac{1}{2}}.$$

For later use, let us note that $a^t_{\epsilon,+}(r, z_1, n_\mathfrak{c}(z))$ is non-increasing in t.

Next let $s, \tilde{u} \in [0,1]$ and $u \in [-1,1]$ and set $\rho_{s,u,\tilde{u}} = \lambda s\sqrt{1-u^2-\tilde{u}}$. Now r and $(z_1, z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_{\mathsf{L}+1})^{\mathsf{T}}$ are independent, and, provided that $\|(z_1, z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_{\mathsf{L}})^{\mathsf{T}}\|^2 = u^2 + \tilde{u}$ is fixed, $z_{\mathsf{L}+1}$ is distributed equally to one component of a uniformly distributed vector on the sphere $\sqrt{1-u^2-\tilde{u}}\;\mathbb{S}^{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-1}$ of radius $\sqrt{1-u^2-\tilde{u}}$. Therefore one has the proportionality for the derivative of the conditional distribution

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}\Big(\lambda r z_{\mathsf{L}+1} \leq x \,\Big|\, (r, z_1, n_\mathfrak{c}(z)) = (s, u, \tilde{u})\Big) \;\propto\; \big(1 - u^2 - \tilde{u} - \tfrac{x^2}{\lambda^2 s^2}\big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-3}{2}}\,\chi_{[-\rho_{s,u,\tilde{u}},\rho_{s,u,\tilde{u}}]}(x)\,.$$

This is similar to (3.1) for $v = e_{\mathsf{L}+1}$. Combining this proportionality relation with (4.4), still under the assumption that (4.3) is satisfied, one deduces that

$$\begin{aligned} &\mathbb{P}\Big(W^t_\epsilon(r, z_1, z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_{\mathsf{L}+1}) \leq 0 \,\Big|\, (r, z_1, n_\mathfrak{c}(z)) = (s, u, \tilde{u})\Big) \\ &\qquad \propto \int_{a^t_{\epsilon,-}(s,u,\tilde{u})}^{a^t_{\epsilon,+}(s,u,\tilde{u})} \mathbf{d}\mathbf{x}(x)\,\big(1 - u^2 - \tilde{u} - \tfrac{x^2}{\lambda^2 s^2}\big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-3}{2}}\,\chi_{[-\rho_{s,u,\tilde{u}},\rho_{s,u,\tilde{u}}]}(x)\,. \end{aligned} \tag{4.5}$$

On the other hand, the probability on the l.h.s. vanishes if (4.3) is violated. By using that the l.h.s. of (4.3) is non-increasing in t as long as it is non-negative, this is equivalent to $t \in [b_\epsilon(r, z_1, z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_\mathsf{L}), \frac{\pi}{2}]$ for some $b_\epsilon(r, z_1, z_{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}+1}, \ldots, z_\mathsf{L}) \in [0, \frac{\pi}{2}]$. Hence it suffices to demonstrate that the r.h.s. of (4.5) is non-increasing in t under the condition that (4.3) holds. As one has the inequality $\big|a^t_{\epsilon,+}\big(s, u, \tilde{u}\big)\big| \leq -a^t_{\epsilon,-}\big(s, u, \tilde{u}\big)$, it is sufficient to consider the cases

$$\begin{array}{ll} \text{(i)} & \rho_{s,u,\tilde{u}} \in \big[0, |a^t_{\epsilon,+}(s,u,\tilde{u})|\big]\,, \\ \text{(ii)} & \rho_{s,u,\tilde{u}} \in \big(|a^t_{\epsilon,+}(s,u,\tilde{u})|, -a^t_{\epsilon,-}(s,u,\tilde{u})\big]\,, \\ \text{(iii)} & \rho_{s,u,\tilde{u}} \in \big(-a^t_{\epsilon,-}(s,u,\tilde{u}), \lambda s\big]\,. \end{array}$$

In these cases the r.h.s. of (4.5) reads respectively:

$$\begin{array}{ll} \text{(i)} & \chi_{[0,\infty)}\big(a^t_{\epsilon,+}(s,u,\tilde{u})\big) \displaystyle\int_{-\rho_{s,u,\tilde{u}}}^{\rho_{s,u,\tilde{u}}} \mathbf{d}\mathbf{x}(x)\, \big(1 - u^2 - \tilde{u} - \tfrac{x^2}{\lambda^2 s^2}\big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-3}{2}}\,, \\ \text{(ii)} & \displaystyle\int_{-\rho_{s,u,\tilde{u}}}^{a^t_{\epsilon,+}(s,u,\tilde{u})} \mathbf{d}\mathbf{x}(x)\, \big(1 - u^2 - \tilde{u} - \tfrac{x^2}{\lambda^2 s^2}\big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-3}{2}}\,, \\ \text{(iii)} & \displaystyle\int_{a^t_{\epsilon,-}(s,u,\tilde{u})}^{a^t_{\epsilon,+}(s,u,\tilde{u})} \mathbf{d}\mathbf{x}(x)\, \big(1 - u^2 - \tilde{u} - \tfrac{x^2}{\lambda^2 s^2}\big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-3}{2}}\,. \end{array}$$

Now, still under the condition that (4.3) holds, $a^t_{\epsilon,+}(s, u, \tilde{u})$ is non-increasing in t and, thus so is (4.5) in the cases (i) and (ii). Moreover, one has the inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\, a^t_{\epsilon,+}(s,u,\tilde{u}) \;\leq\; \frac{\mathrm{d}}{\mathrm{d}t}\, a^t_{\epsilon,-}(s,u,\tilde{u}) \;\leq\; 0\,.$$

If $\mathsf{L}_\mathfrak{a} + \mathsf{L}_\mathfrak{b} \geq 3$, the case (iii) is therefore dealt with by

$$\big(1 - u^2 - \tilde{u} - \tfrac{a^t_{\epsilon,+}(s,u,\tilde{u})^2}{\lambda^2 s^2}\big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-3}{2}} \;\geq\; \big(1 - u^2 - \tilde{u} - \tfrac{a^t_{\epsilon,-}(s,u,\tilde{u})^2}{\lambda^2 s^2}\big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}-3}{2}}\,.$$

In conclusion, (4.5) is non-increasing in t for all $s, \tilde{u} \in [0, 1]$ and $u \in [-1, 1]$. Due to $\mathsf{L}_\mathfrak{a} + \mathsf{L}_\mathfrak{b} \geq 3$, this finishes the proof of Step 2 and hence also the propostion. □

Proof of Corollary 2.13. For $w \in \mathbb{S}^\mathsf{L}$, $N \in \mathbb{N}$, $\tilde{N} \in \{1, \ldots, N\}$ and $M \geq \tilde{N} + 1$, let us consider the stochastic order

$$\mathbb{P}\Big(\big\|\mathfrak{c}\big(\big[\prod_{n=\tilde{N}+1}^{M} (\mathbf{1} + \lambda r_n U_n)\big]\mathcal{R}\cdot w\big)\big\| \in \cdot\Big) \;\geq_{\mathrm{st}}\; \mathbb{P}\Big(\big\|\mathfrak{c}\big(\big[\prod_{n=\tilde{N}+1}^{M} (\mathbf{1} + \lambda r_n U_n)\big]\cdot w\big)\big\| \in \cdot\Big)\,. \tag{4.6}$$

For $M = \tilde{N} + 1$, it follows from Proposition 2.11 and the estimate

$$\begin{aligned} \|\mathfrak{c}(w)\| \;&=\; \Big[1 + (\kappa_{\mathsf{L}_\mathfrak{c}})^2 \Big(\|\mathfrak{a}(w)\|^2 + \|\mathfrak{b}(w)\|^2\Big) (\kappa_{\mathsf{L}_\mathfrak{c}})^{-2}\, \|\mathfrak{c}(w)\|^{-2}\Big]^{-\frac{1}{2}} \\ &\leq\; \Big[1 + \Big(\|\mathfrak{a}(\mathcal{R}w)\|^2 + \|\mathfrak{b}(\mathcal{R}w)\|^2\Big) \|\mathfrak{c}(\mathcal{R}w)\|^{-2}\Big]^{-\frac{1}{2}} \\ &=\; \|\mathfrak{c}\,(\mathcal{R}\cdot w)\|\;, \end{aligned}$$

holding true whenever $\|\mathfrak{c}(w)\| > 0$. Next we show by an iterative argument that (4.6) also holds for larger M. This is based on the general fact that the expectations of any non-decreasing function of a pair of stochastically ordered random variable is ordered (see *e.g.* [6], p. 385). Due to (4.1), the map $g_\epsilon : [0, 1] \to [0, 1]$ given by

$$g_\epsilon(x) \;=\; \mathbb{P}\big(\|\mathfrak{c}((\mathbf{1} + \lambda r U)\cdot \tilde{w})\| > \epsilon \,\big|\, \|\mathfrak{c}(\tilde{w})\| = x\big)$$

is well-defined for all $\epsilon \in [0,1]$. Moreover, it is non-decreasing by Proposition 2.11 and can be extended to a non-decreasing function on $\mathbb{R}$. Thus if (4.6) holds for some $M \in \{\tilde{N}+1,\dots,N-1\}$, then all $\epsilon \in [0,1]$ satisfy

$$\mathbb{E}\, g_\epsilon\Big(\big\|\mathfrak{c}\big([\prod_{n=\tilde{N}+1}^{M}(\mathbf{1}+\lambda r_n U_n)]\mathcal{R}\cdot w\big)\big\|\Big) \;\geq\; \mathbb{E}\, g_\epsilon\Big(\big\|\mathfrak{c}\big(\prod_{n=\tilde{N}+1}^{M}(\mathbf{1}+\lambda r_n U_n)\cdot w\big)\big\|\Big)\,,$$

or, equivalently,

$$\mathbb{P}\Big(\big\|\mathfrak{c}\big([\prod_{n=\tilde{N}+1}^{M+1}(\mathbf{1}+\lambda r_n U_n)]\mathcal{R}\cdot w\big)\big\|\leq\epsilon\Big) \;\leq\; \mathbb{P}\Big(\big\|\mathfrak{c}\big(\prod_{n=\tilde{N}+1}^{M+1}(\mathbf{1}+\lambda r_n U_n)\cdot w\big)\big\|\leq\epsilon\Big)\,,$$

namely (4.6) remains valid if M is replaced by $M+1$ so that it also holds for $M=N$. As $w \in \mathbb{S}^{\mathsf{L}}$ is arbitrary in the above, one infers that all $(v,\epsilon) \in \mathbb{S}^{\mathsf{L}} \times [0,1]$ obey

$$\begin{aligned}
&\mathbb{P}\Big(\Big\|\mathfrak{c}\big(\prod_{\tilde{n}=\tilde{N}+1}^{N}(\mathbf{1}+\lambda r_{\tilde{n}} U_{\tilde{n}})\prod_{n=1}^{\tilde{N}}\mathcal{R}(\mathbf{1}+\lambda r_n U_n)\cdot v\big)\Big\|\leq\epsilon\Big)\\
&\;=\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\mathbb{P}\Big((\mathbf{1}+\lambda r_{\tilde{N}} U_{\tilde{N}})\textstyle\prod_{n=1}^{\tilde{N}-1}\mathcal{R}\,(\mathbf{1}+\lambda r_n U_n)\cdot v\in\cdot\Big)(w)\,\mathbb{P}\Big(\Big\|\mathfrak{c}\,\Big(\Big(\prod_{n=\tilde{N}+1}^{N}(\mathbf{1}+\lambda r_n U_n)\Big)\mathcal{R}\cdot w\Big)\Big\|\leq\epsilon\Big)\\
&\;\leq\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\mathbb{P}\Big((\mathbf{1}+\lambda r_{\tilde{N}} U_{\tilde{N}})\textstyle\prod_{n=1}^{\tilde{N}-1}\mathcal{R}\,(\mathbf{1}+\lambda r_n U_n)\cdot v\in\cdot\Big)(w)\,\mathbb{P}\Big(\Big\|\mathfrak{c}\,\Big(\prod_{n=\tilde{N}+1}^{N}(\mathbf{1}+\lambda r_n U_n)\cdot w\Big)\Big\|\leq\epsilon\Big)\\
&\;=\mathbb{P}\Big(\Big\|\mathfrak{c}\big(\prod_{\tilde{n}=\tilde{N}}^{N}(\mathbf{1}+\lambda r_{\tilde{n}} U_{\tilde{n}})\prod_{n=1}^{\tilde{N}-1}\mathcal{R}(\mathbf{1}+\lambda r_n U_n)\cdot v\big)\Big\|\leq\epsilon\Big)\,.
\end{aligned}$$

An iterative application of this bound yields (2.16). □

Proof of Proposition 2.14. Let $h \in \mathrm{L}^\infty(\mathbb{S}^{\mathsf{L}})$. Using Tonelli's theorem in the second step as well as (2.4) in the penultimate step, one finds

$$\begin{aligned}
\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\varrho^{\mathrm{ac}}_{r,\lambda,v}(w)\,h(w) &= \int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(w)\,\frac{\mathrm{d}\varrho^{\mathrm{ac}}_{r,\lambda,v}}{\mathrm{d}\nu_{\mathsf{L}}}(w)\,h(w)\\
&= \int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(w)\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\,\frac{\mathrm{d}\varrho^{\mathrm{ac}}_{r,\lambda,v}}{\mathrm{d}\nu_{\mathsf{L}}}(w)\,h(w)\\
&= \int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(w)\;h(w)\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\,\frac{\mathrm{d}\varrho^{\mathrm{ac}}_{r,\lambda,v}}{\mathrm{d}\nu_{\mathsf{L}}}(w)\\
&= \int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(w)\;h(w)\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\,\frac{\mathrm{d}\varrho^{\mathrm{ac}}_{r,\lambda,w}}{\mathrm{d}\nu_{\mathsf{L}}}(v)\\
&= \varrho^{\mathrm{ac}}_{r,\lambda,v}(\mathbb{S}^{\mathsf{L}})\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(w)\;h(w)\,.
\end{aligned}$$

Combining this with (2.3) yields

$$\begin{aligned}
\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\,\mathbb{E}\,h\left((\mathbf{1}+\lambda rU)\cdot v\right) &= \int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\varrho_{r,\lambda,v}(w)\,h(w)\\
&= \int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\varrho^{\mathrm{pp}}_{r,\lambda,v}(w)\,h(w)+\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(v)\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\varrho^{\mathrm{ac}}_{r,\lambda,v}(w)\,h(w)\\
&= \varrho^{\mathrm{pp}}_{r,\lambda,v}(\mathbb{S}^{\mathsf{L}})\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(w)\;h(w)+\varrho^{\mathrm{ac}}_{r,\lambda,v}(\mathbb{S}^{\mathsf{L}})\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(w)\;h(w)\\
&= \varrho_{r,\lambda,v}(\mathbb{S}^{\mathsf{L}})\int_{\mathbb{S}^{\mathsf{L}}}\mathrm{d}\nu_{\mathsf{L}}(w)\;h(w)\,.
\end{aligned}$$

Since $\varrho_{r,\lambda,v}$ is normalized, the claim (2.17) follows. □

Proof of Lemma 2.15. The measure ν_{L} is given by the normalized restriction of the L-dimensional Hausdorff measure in $\mathbb{R}^{\mathsf{L}+1}$ to the sphere $\mathbb{S}^{\mathsf{L}}$ (see *e.g.* [9]). Let $\mathsf{L} = \mathsf{L}^+ + \mathsf{L}^- + 1$. Later on, we will choose $\mathsf{L}^- = \mathsf{L}_{\mathfrak{c}} - 1$. Then let us decompose $v \in \mathbb{R}^{\mathsf{L}+1}$ as follows

$$v \;=\; r \begin{pmatrix} \cos(\theta) v^+ \\ \sin(\theta) v^- \end{pmatrix} \;,$$

where $r = \|v\|$, $\theta \in [0, \frac{\pi}{2}]$ and $v^\pm \in \mathbb{S}^{\mathsf{L}^\pm} \subset \mathbb{R}^{\mathsf{L}^\pm + 1}$ are unit vectors which are then described by angles $(\theta_1^\pm, \ldots, \theta_{\mathsf{L}^\pm}^\pm) \in [0, 2\pi) \times [0, \pi)^{\times \mathsf{L}^\pm - 1}$ using the standard spherical coordinates, namely $v^\pm = v^\pm(\theta_1^\pm, \ldots, \theta_{\mathsf{L}^\pm}^\pm)$ has the components

$$v_1^\pm \;=\; \prod_{n=1}^{\mathsf{L}^\pm} \sin(\theta_n^\pm) \;, \quad v_k^\pm \;=\; \cos(\theta_{k-1}^\pm) \prod_{n=k}^{\mathsf{L}^\pm} \sin(\theta_n^\pm) \;, \qquad v_{\mathsf{L}^\pm+1}^\pm \;=\; \cos(\theta_{\mathsf{L}^\pm}^\pm) \;.$$

This provides a bijection from $\mathbb{R}^{\mathsf{L}+1}$ to $(0, \infty) \times (0, \frac{\pi}{2}) \times (0, 2\pi)^{\times 2} \times (0, \pi)^{\times \mathsf{L}^+ + \mathsf{L}^- - 2}$, up to sets of zero measure. The Jacobian of the transformation is

$$J \;=\; \det \begin{pmatrix} \cos(\theta) v^+ & -r\sin(\theta) v^+ & r\cos(\theta) \partial_{\theta^+} v^+ & 0 \\ \sin(\theta) v^- & r\cos(\theta) v^- & 0 & r\sin(\theta) \partial_{\theta^-} v^- \end{pmatrix} \;,$$

which can be evaluated explicitly

$$J \;=\; r^{\mathsf{L}} \cos(\theta)^{\mathsf{L}^+} \sin(\theta)^{\mathsf{L}^-} \Big(\prod_{n=1}^{\mathsf{L}^+} \sin(\theta_n^+)^{n-1} \Big) \Big(\prod_{n=1}^{\mathsf{L}^-} \sin(\theta_n^-)^{n-1} \Big) \;.$$

Hence

$$\nu_{\mathsf{L}}(\{v \in \mathbb{S}^{\mathsf{L}} \;:\; \sin(\theta)^2 \leq \delta\}) \;=\; \frac{\int_0^{\arcsin(\delta^{\frac{1}{2}})} \mathbf{dx}(\theta) \;\sin(\theta)^{\mathsf{L}^-} \cos(\theta)^{\mathsf{L}^+}}{\int_0^{\frac{\pi}{2}} \mathbf{dx}(\theta) \;\sin(\theta)^{\mathsf{L}^-} \cos(\theta)^{\mathsf{L}^+}} \;.$$

Setting $\mathsf{L}^- = \mathsf{L}_{\mathfrak{c}} - 1$, substituting $x = \sin(\theta)^2$ and evaluating the integral in the numerator leads to the identity (2.18). The generalized binomial coefficient can be bounded as follows:

$$\frac{\Gamma(\frac{\mathsf{L}+1}{2})}{\Gamma(\frac{\mathsf{L}_{\mathfrak{c}}}{2})\Gamma(\frac{\mathsf{L}_{\mathfrak{a}}+\mathsf{L}_{\mathfrak{b}}}{2})} \;\leq\; \frac{\mathsf{L}_{\mathfrak{c}}}{2} \Big(\frac{\mathsf{L}+1}{\mathsf{L}_{\mathfrak{a}}+\mathsf{L}_{\mathfrak{b}}} \Big)^{\frac{\mathsf{L}_{\mathfrak{a}}+\mathsf{L}_{\mathfrak{b}}}{2} - 1} \Big(\frac{\mathsf{L}+1}{\mathsf{L}_{\mathfrak{c}}} \Big)^{\frac{\mathsf{L}_{\mathfrak{c}}}{2}} \;.$$

Furthermore, as $(\mathsf{L}_{\mathfrak{a}}, \mathsf{L}_{\mathfrak{b}}) \neq (1,1)$, the factor $(1-x)^{\frac{\mathsf{L}_{\mathfrak{a}}+\mathsf{L}_{\mathfrak{b}}}{2}-1}$ can be bounded by $(1 - \frac{x}{2})$ so that the numerator in (2.18) is bounded by

$$\int_0^\delta \mathbf{dx}(x)\; x^{\frac{\mathsf{L}_{\mathfrak{c}}}{2}-1} (1-x)^{\frac{\mathsf{L}_{\mathfrak{a}}+\mathsf{L}_{\mathfrak{b}}}{2}-1} \;\leq\; \delta^{\frac{\mathsf{L}_{\mathfrak{c}}}{2}} \Big[\frac{2}{\mathsf{L}_{\mathfrak{c}}} \;-\; \frac{\delta}{\mathsf{L}_{\mathfrak{c}}+2} \Big] \;\leq\; \frac{2}{\mathsf{L}_{\mathfrak{c}}} \,\delta^{\frac{\mathsf{L}_{\mathfrak{c}}}{2}} \Big[1 \;-\; \frac{\delta}{6} \Big] \;.$$

This proves (2.19). □

Proof of Corollary 2.16. Let $\delta \in (0,1)$ and $\epsilon \in (0, \min\{\delta, 1-\delta\})$. Clearly

$$\mathrm{supp}(\mathbf{1} + \lambda r U) \;=\; \big\{ \mathbf{1} + \lambda s \mathscr{U} \;\big|\; (s, \mathscr{U}) \in \mathrm{supp}(r) \times \mathrm{O}(\mathsf{L}+1) \big\}$$

is both contracting and strongly irreducible (see [3], Part A, Definition III. 1.3 & III. 2.1), since $r \not\equiv 0$ and $\lambda > 0$. By Furstenberg's theorem, it follows that there is a unique invariant measure which due to Proposition 2.14 is given by the Haar measure ν_{L} on $\mathbb{S}^{\mathsf{L}}$. Furthermore, by Theorem III.4.3 in [3], one has for any continuous function $h : \mathbb{S}^{\mathsf{L}} \to \mathbb{R}$ that

$$\lim_{N\to\infty} \sup_{v \in \mathbb{S}^{\mathsf{L}}} \Big| \mathbb{E}\, h\Big(\prod_{n=1}^{N} (\mathbf{1} + \lambda r_n U_n) \cdot v \Big) \;-\; \int_{\mathbb{S}^{\mathsf{L}}} \mathrm{d}\nu_{\mathsf{L}}(w)\; h(w) \Big| \;=\; 0 \;.$$

Let us choose

$$h_{\delta,\epsilon}(v) \;=\; \min\Big\{1, \epsilon^{-1}\left(\delta+\epsilon-\|\mathfrak{c}(v)\|^2\right)\Big\}\;\chi_{\{\|\mathfrak{c}(v)\|^2\le\delta+\epsilon\}}(v)\;.$$

By construction, $h_{\delta,\epsilon}$ is continuous. Thus there exists an $\tilde{N}_0=\tilde{N}_0(\mathsf{L},\mathsf{L}_\mathfrak{c},\delta,\epsilon)\in\mathbb{N}$ such that all $N\ge\tilde{N}_0$ and $v\in\mathbb{S}^\mathsf{L}$ satisfy

$$\left|\mathbb{E}\;h_{\delta,\epsilon}\big(\prod_{n=1}^N(\mathbf{1}+\lambda r_nU_n)\cdot v\big)\;-\;\int_{\mathbb{S}^\mathsf{L}}\mathrm{d}\nu_\mathsf{L}(v)\;h_{\delta,\epsilon}(v)\right|\;\le\;\epsilon\;. \tag{4.7}$$

Further, $h_{\delta,\epsilon}$ can be bounded from below and above by indicator functions:

$$\chi_{\{\|\mathfrak{c}(v)\|^2\le\delta\}}\;\le\;h_{\delta,\epsilon}\;\le\;\chi_{\{\|\mathfrak{c}(v)\|^2\le\delta+\epsilon\}}\;. \tag{4.8}$$

Now using (4.7), (4.8) as well as (2.19) with $\delta+\epsilon$ instead of δ it follows that

$$\begin{aligned}
\mathbb{P}\Big(\big\|\mathfrak{c}\big(\prod_{n=1}^N(\mathbf{1}+\lambda r_nU_n)\cdot v\big)\big\|^2<\delta\Big)\;&\le\;\mathbb{E}\;h_{\delta,\epsilon}\Big(\prod_{n=1}^N(\mathbf{1}+\lambda r_nU_n)\cdot v\Big)\\
&\le\;\int_{\mathbb{S}^\mathsf{L}}\mathrm{d}\nu_\mathsf{L}(v)\;h^\mathsf{L}_{\delta,\epsilon}(v)+\epsilon\\
&\le\;\nu_\mathsf{L}\big(\big\{\|\mathfrak{c}(v)\|^2\le\delta+\epsilon\big\}\big)\;+\;\epsilon\\
&\le\;\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}\Big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}{2}-1}\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{c}}\,(\delta+\epsilon)\Big)^{\frac{\mathsf{L}_\mathfrak{c}}{2}}\Big(1-\frac{\delta}{6}\Big)\;+\;\epsilon\\
&=\;\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}\Big)^{\frac{\mathsf{L}_\mathfrak{a}+\mathsf{L}_\mathfrak{b}}{2}-1}\Big(\frac{\mathsf{L}+1}{\mathsf{L}_\mathfrak{c}}\,\delta\Big)^{\frac{\mathsf{L}_\mathfrak{c}}{2}}\;-\;\eta(\epsilon,\delta)\;,
\end{aligned}$$

the last equation simply by definition of $\eta(\epsilon,\delta)$. Now one readily checks that $\lim_{\epsilon\downarrow 0}\eta(\epsilon,\delta)$ is positive and therefore, by continuity of $\eta(\epsilon,\delta)$, there exists a positive ϵ for which the bound (2.20) is satisfied for some positive $\eta=\eta(\mathsf{L},\mathsf{L}_\mathfrak{c},\delta)>0$. □

References

[1] M. Aizenman, S. Warzel, *Random operators*, (AMS, Providence, 2015).

[2] S. Bachmann, W. De Roeck, *From the Anderson model on a strip to the DMPK equation and random matrix theory*, J. Stat. Phys. **139**, 541–564 (2010).

[3] P. Bougerol, J. Lacroix, *Products of Random Matrices with Applications to Schrödinger Operators*, (Birkhäuser, Boston, 1985).

[4] R. Carmona, J. M. Lacroix, *Spectral theory of random Schrödinger operators*, (Birkhäuser, Basel, 1990).

[5] Ya. I. Goldsheid, A. G. Margulis, *Lyapunov exponents of a product of random matrices*, Russian Math. Surveys **44**, 11–71 (1989).

[6] A. Klenke, *Probability theory: a comprehensive course*, (Springer Universitext, London, 2013).

[7] O. Kallenberg, *Foundations of modern probability*, (Springer, New York, 2002).

[8] A. W. W. Ludwig, H. Schulz-Baldes, M. Stolz, *Lyapunov spectra for all ten symmetry classes of quasi-one-dimensional disordered systems of non-interacting fermions*, J. Stat. Phys. **152**, 275–304 (2013).

[9] P. Mattila, *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*, (Cambridge University Press, Cambridge, 1995).

[10] R. Römer, H. Schulz-Baldes, *Weak disorder expansion for localization lengths of quasi-1D systems*, Euro. Phys. Lett. **68**, 247–253 (2004).

[11] R. Römer, H. Schulz-Baldes, *Random phase property and the Lyapunov spectrum for disordered multi-channel systems*, J. Stat. Phys. **140**, 122–153 (2010).

[12] C. Sadel, H. Schulz-Baldes, *Random Lie group actions on compact manifolds: a perturbative analysis*, Annals of Probability **38**, 2224–2257 (2010).

[13] C. Sadel, B. Virág, *A central limit theorem for products of random matrices and GOE statistics for the Anderson model on long boxes*, Commun. Math. Phys. **343**, 881–919, (2016).

[14] H. Schulz-Baldes, *Perturbation theory for an Anderson model on a strip*, GAFA **14**, 1089–1117 (2004).

[15] B. Valkó, B. Virág, *Random Schrödinger operators on long boxes, noise explosion and the GOE*, Trans. AMS **366**, 3709–3728 (2014).

Acknowledgments. We thank Andreas Knauf for many discussions as well as helpful and constructive comments. F. D. received funding from the *Studienstiftung des deutschen Volkes*. This work was also supported by the DFG.