OPTIMIZATION GEOMETRY OF EQUIVALENT BROWNIAN RKHS REPRESENTATIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

Equivalent finite parameterizations can represent the same functions and intrinsic norm yet induce different optimization algorithms. We study this effect in a controlled finite Brownian RKHS with nodal, increment, and spectral coordinates. Classical finite-element, RKHS-interpolation, Brownian-covariance, and mixed-boundary DCT identities make the shared hypothesis class, Brownian energy, approximation operator, and coordinate maps explicit.

Our main results concern the optimization geometry of this fixed model. With mapped initialization, identical scalar steps, and identical minibatches, nodal and spectral GD/SGD have exactly the same mapped trajectories. Increment GD is an explicit Euler step for the constant Brownian/Sobolev metric, with factor $1/h$. For Brownian-regularized least squares, $\kappa_2(\mathbf{H}_{\mathrm{inc}}) \le 1 + A/\rho$, independently of grid resolution G for fixed A, $\rho > 0$, and the stated normalization. Under the stated standard-Adam convention, the universal orthogonal equivariance group is exactly the signed permutations; the block DCT-VIII transform is not one. Float64 tests over five grids numerically verify the finite identities, mapped one-layer and recursive trajectories, conditioning predictions, and theorem-matched Adam separation. Thus coordinate effects are isolated without changing the represented functions, intrinsic regularizer, or approximation space.

1 INTRODUCTION

Reproducing kernel Hilbert spaces describe a model through the functions it represents and the intrinsic norm controlling them (Aronszajn, 1950; Schölkopf & Smola, 2002; Berlinet & Thomas-Agnan, 2004; Steinwart & Christmann, 2008). Splines, Gaussian processes, random features, and neural tangent kernels all exploit this viewpoint (Wahba, 1990; Rasmussen & Williams, 2006; Rahimi & Recht, 2007; Jacot et al., 2018; Lee et al., 2019). A complementary line studies the geometry of optimization: natural gradient, mirror descent, implicit-bias analysis, and reparameterization theory show that equivalent descriptions of one model can induce different algorithms (Amari, 1998; Beck & Teboulle, 2003; Gunasekar et al., 2018; Amid & Warmuth, 2020; Kristiadi et al., 2023; Li et al., 2022; Martens, 2020). Coordinatewise adaptive methods add a basis dependence that ordinary Euclidean gradient descent does not have under orthogonal changes (Kingma & Ba, 2015; Ling et al., 2022; Zhang et al., 2025).

These viewpoints are difficult to separate because a neural reparameterization often also changes initialization, implicit regularization, architecture, or the represented class. A controlled comparison instead needs several coordinate systems for one fixed Hilbert space. Brownian profile layers provide this setting: BKLs and VBKLs recursively use one-dimensional Brownian RKHS profiles (Mohammadigohari et al., 2026; Mohammadigohari, 2026), and the VBKL companion already employs anchored piecewise-linear profiles and their increment energy.

The required ingredients are largely classical: piecewise-linear stiffness energies come from finite elements (Strang & Fix, 1973; Brenner & Scott, 2008); nodal interpolation and power functions from spline and kernel approximation (Wahba, 1990; Fasshauer, 2007; Wendland, 2005); Brownian covariance identities from Gaussian processes (Rasmussen & Williams, 2006); and the mixed-boundary spectrum from DCT theory (Martucci, 1994; Strang, 1999; Masera et al., 2017). We assemble them in one two-sided anchored normalization so that the represented space, norm, ap-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

proximation error, and coordinate maps are simultaneously fixed and explicit, rather than claiming these ingredients separately as new.

In this controlled model, the spectral map is orthogonal whereas $\mathbf{D}_0^\top \mathbf{D}_0 = h\mathbf{K}_0$. We derive the resulting mapped optimizer trajectories, the constant-metric Brownian interpretation of increment descent, and its blockwise recursive consequence. We also obtain the fixed-A, ρ grid-resolution-independent least-squares bound and the maximal universal orthogonal equivariance group of the stated standard-Adam update.

Our contributions are:

1. one controlled finite Brownian RKHS with exactly equivalent nodal, increment, and block DCT-VIII coordinates, preserving the hypothesis class, intrinsic norm, and approximation space;
2. exact mapped GD/SGD identities: nodal and spectral trajectories coincide, while increment descent is a constant-metric Brownian/Sobolev Euler step with the precise mesh scaling;
3. the blockwise recursive consequence and $\kappa_2(\mathbf{H}_{\mathrm{inc}}) \le 1 + A/\rho$, independent of G under the stated fixed quantities; and
4. the signed-permutation maximality theorem for standard Adam, with source-traced float64 verification. EuroSAT and Salinas remain secondary external-validity studies, not coordinate-equivalence tests.

Proofs and supplementary material are in the appendix; Table 3 summarizes the results.

2 Notation and Preliminaries

We write $[n] = \{1, \dots, n\}$, $\langle \cdot, \cdot \rangle$ and $\| \cdot \|_2$ for the Euclidean inner product and norm, $\mathbf{I}_n$ for the identity, $\mathbf{e}_j$ for a canonical basis vector, δ_{ij} for the Kronecker delta, and $\mathrm{blkdiag}$ for block-diagonal assembly. Vectors are bold lowercase, matrices bold uppercase, scalars italic; grid quantities are indexed from 0 and spectral quantities from 1. Three symbols are disambiguated relative to the underlying frameworks: increment coordinates are $\mathbf{w}$, reserving δ for the Kronecker delta and the evaluation functional δ_t; eigenvalues of the block $\mathbf{T}_m$ are ν_k, reserving μ for measures; and eigenvalues of the anchored stiffness matrix are λ_k, $k \in [m]$, each of multiplicity two. Notation used only inside proofs is collected in Appendix A.

BKLs construct recursive reproducing-kernel representations from the one-dimensional Brownian kernel (Mohammadigohari et al., 2026),

$$k_{\mathrm{B}}(x, x') = \frac{|x| + |x'| - |x - x'|}{2} = \min(|x|, |x'|) \mathbb{1}\{xx' \ge 0\}, \qquad x, x' \in \mathbb{R},$$

the covariance of two-sided Brownian motion anchored at the origin. Its RKHS on I, denoted $\mathcal{H}_{\mathrm{B}}$, is the Cameron–Martin space $\mathcal{H}_{\mathrm{B}} = \{f \in H^1(I) : f(0) = 0\}$ with $\langle f, g \rangle_{\mathcal{H}_{\mathrm{B}}} = \int_I f' g'$ (Berlinet & Thomas-Agnan, 2004). Given support functions $\mathcal{S}$ on an input space $\mathcal{X}$, each $u : \mathcal{X} \to \mathbb{R}$, and a probability measure μ on $\mathcal{S}$, the Brownian integral kernel $k[\mathcal{S}, \mu](\mathbf{x}, \mathbf{x}') = \int_{\mathcal{S}} k_{\mathrm{B}}(u(\mathbf{x}), u(\mathbf{x}'))\, \mathrm{d}\mu(u)$ has an RKHS defining the next layer; iterating from linear projections yields a hierarchy of Brownian RKHSs. Variation Brownian Kernel Ladders (VBKLs) instead use an atomic representation (Mohammadigohari, 2026): composing support functions with Brownian profiles yields a dictionary $\mathcal{U}_L = \{g \circ u : u \in \mathcal{U}_{L-1},\ g \in \mathcal{H}_{\mathrm{B}},\ \|g\|_{\mathcal{H}_{\mathrm{B}}} \le 1\}$ of depth-L atoms, with variation space $\mathcal{V}^{(L)} = \{F = \int_{\mathcal{U}_L} u\, \mathrm{d}\mu(u) : \|\mu\|_{\mathrm{TV}} < \infty\}$ and complexity $\widehat{\mathcal{C}}^{(L)}_{\mathrm{var}}(F) = \inf\{\|\mu\|_{\mathrm{TV}} : F = \int_{\mathcal{U}_L} u\, \mathrm{d}\mu(u)\}$. In both, profile functions lie in infinite-dimensional Brownian RKHSs, so implementations must discretize them. The VBKL construction already employs anchored piecewise-linear profiles and their increment-energy constraint; the present paper makes the corresponding finite RKHS and coordinate metrics explicit for a controlled optimization comparison.

3 FINITE BROWNIAN RKHS

We collect the classical finite-element and RKHS machinery in the paper's two-sided anchored normalization (Strang & Fix, 1973; Brenner & Scott, 2008), ensuring that later optimizer comparisons use the same function space, norm, and approximation operator.

Definition 1 (Finite Brownian profile space)**.** Let $I = [-A, A]$ for some $A > 0$, let $G \in \mathbb{N}$, and let $t_i = -A + ih$ for $i = 0, \dots, G$ be the uniform grid with mesh size $h = \frac{2A}{G}$, so that $-A = t_0 < t_1 < \cdots < t_G = A$. For each $i = 0, \dots, G$, let $\phi_i \in C(I)$ be the hat function that is affine on each subinterval $[t_j, t_{j+1}]$ and satisfies $\phi_i(t_j) = \delta_{ij}$ for every $j = 0, \dots, G$ (Brenner & Scott, 2008). Define the finite-element reconstruction operator $\mathcal{R} : \mathbb{R}^{G+1} \to C(I)$ by $\mathcal{R}(\mathbf{v}) = \sum_{i=0}^{G} v_i \phi_i$, where $\mathbf{v} = (v_0, \dots, v_G)^\top \in \mathbb{R}^{G+1}$. The *finite Brownian profile space* is $\mathcal{H}_h = \mathcal{R}(\mathbb{R}^{G+1}) = \operatorname{span}(\phi_0, \dots, \phi_G)$. Since $\phi_i(t_j) = \delta_{ij}$, the family $\{\phi_i\}_{i=0}^{G}$ is linearly independent, so every $f \in \mathcal{H}_h$ has a unique coefficient vector $\mathbf{v} \in \mathbb{R}^{G+1}$ with $f = \mathcal{R}(\mathbf{v})$.

Definition 2 (Brownian energy form)**.** On $\mathcal{H}_h$ we define the symmetric bilinear form $\langle f, g \rangle_{B,h} = \int_{-A}^{A} f'(t) g'(t)\, dt$, with derivatives taken elementwise on each $[t_i, t_{i+1}]$ and hence defined almost everywhere, and the Brownian seminorm $\|f\|_{B,h} = \sqrt{\langle f, f \rangle_{B,h}}$.

Proposition 1 (Finite Brownian energy representation)**.** *Let $\mathbf{D} \in \mathbb{R}^{G \times (G+1)}$ be the first-difference matrix, $(\mathbf{D}\mathbf{v})_i = v_{i+1} - v_i$ for $i = 0, \dots, G-1$, and let $\mathbf{K} = (K_{rs})_{r,s=0}^{G}$ with $K_{rs} = \langle \phi_r, \phi_s \rangle_{B,h}$ be the stiffness matrix. For $f = \mathcal{R}(\mathbf{v}) \in \mathcal{H}_h$,*

$$\|f\|_{B,h}^2 = \mathbf{v}^\top \mathbf{K} \mathbf{v} = \frac{1}{h} \|\mathbf{D}\mathbf{v}\|_2^2 = \frac{1}{h} \sum_{i=0}^{G-1} (v_{i+1} - v_i)^2 .$$

Since both matrices are symmetric, this is equivalent to the matrix identity $\mathbf{K} = \frac{1}{h} \mathbf{D}^\top \mathbf{D}$.

This recovers the classical stiffness-matrix form of the Dirichlet energy (Strang & Fix, 1973; Brenner & Scott, 2008); we record it because it identifies the regularizer that finite realizations implement.

Corollary 1 (Implemented Brownian regularizer)**.** *For $\mathbf{v} \in \mathbb{R}^{G+1}$ let $\Omega_h(\mathbf{v}) := \sum_{i=0}^{G-1} (v_{i+1} - v_i)^2$. Then $\Omega_h(\mathbf{v}) = h \|\mathcal{R}(\mathbf{v})\|_{B,h}^2$: the implemented regularizer is the Brownian energy times the fixed mesh factor $h > 0$.*

3.1 ANCHORED FINITE BROWNIAN RKHS

The Brownian energy is constant-invariant and hence only a seminorm on $\mathcal{H}_h$. As in the Cameron–Martin space (Berlinet & Thomas-Agnan, 2004), uniqueness is recovered by fixing the value at a reference point.

Assumption 1. $G = 2m$ with $m \in \mathbb{N}$, and $i_0 := m$, so $t_{i_0} = 0$. In force for the remainder of the paper.

Definition 3 (Anchored finite Brownian profile space)**.** $\mathcal{H}_h^0 = \{f \in \mathcal{H}_h : f(0) = 0\}$; equivalently $f = \mathcal{R}(\mathbf{v}) \in \mathcal{H}_h^0$ iff $v_{i_0} = 0$, and $\dim \mathcal{H}_h^0 = G$.

Since the anchor removes one coordinate, we work with a free parameter in $\mathbb{R}^G$. The ordering below traverses each half-grid from its exterior endpoint toward the anchor, which is what makes the two diagonal blocks of Theorem 3 the same matrix.

Definition 4 (Reduced anchored coordinates)**.** For an anchored nodal vector $\mathbf{v} \in \mathbb{R}^{G+1}$ with $v_{i_0} = 0$, the *reduced anchored vector* is $\widetilde{\mathbf{v}} := (v_0, \dots, v_{m-1}, v_{2m}, v_{2m-1}, \dots, v_{m+1})^\top \in \mathbb{R}^G$, and the *anchoring reconstruction matrix* is $\mathbf{R} := (\mathbf{e}_0, \dots, \mathbf{e}_{m-1}, \mathbf{e}_{2m}, \mathbf{e}_{2m-1}, \dots, \mathbf{e}_{m+1}) \in \mathbb{R}^{(G+1) \times G}$, so that $\mathbf{R}\widetilde{\mathbf{v}} = \mathbf{v}$. We write $\mathbf{K}_0 := \mathbf{R}^\top \mathbf{K} \mathbf{R}$ for the *anchored stiffness matrix* and $\mathbf{D}_0 := \mathbf{D}\mathbf{R} \in \mathbb{R}^{G \times G}$.

The anchored energy is positive definite, which is what makes the anchored space an inner-product space rather than merely a seminormed one.

Lemma 3.1 (Positive definiteness)**.** $\ker \mathbf{K} = \operatorname{span}(\mathbf{1})$ *and* $\mathbf{K}_0 \succ 0$. *Consequently $\langle \cdot, \cdot \rangle_{B,h}$ is an inner product on $\mathcal{H}_h^0$.*

We can now record the Brownian-specific kernel formula and its covariance interpretation. The value of the result here is its exact normalization and its use in the subsequent controlled comparison, rather than finite-dimensional RKHS existence itself.

Theorem 1 (Finite Brownian RKHS and its reproducing kernel)**.** *The space $(\mathcal{H}_h^0, \langle\cdot,\cdot\rangle_{B,h})$ is a G-dimensional RKHS. Let $\mathcal{J} := \{0,\dots,G\} \setminus \{i_0\}$ and let $\pi : \mathcal{J} \to [G]$ give the position of a node in the reduced ordering of Definition 4. With $\widetilde{\phi}(t) := (\phi_{\pi^{-1}(1)}(t), \dots, \phi_{\pi^{-1}(G)}(t))^\top$, its reproducing kernel satisfies:*

(i) $k_h(s,t) = \widetilde{\phi}(s)^\top \mathbf{K}_0^{-1} \widetilde{\phi}(t)$ *for all* $s,t \in I$.

(ii) $(\mathbf{K}_0^{-1})_{\pi(r),\pi(q)} = k_{\mathrm{B}}(t_r,t_q)$ *for* $r,q \in \mathcal{J}$*, and* $k_h(\cdot,t_r) = k_{\mathrm{B}}(\cdot,t_r)$ *for every grid node. More generally,* k_h *is the tensor-product piecewise-bilinear interpolant of* k_{B} *on the grid, and therefore agrees with* k_{B} *whenever at least one argument is a node.*

Corollary 2 (Conforming Brownian interpolation space)**.**

$$\mathcal{H}_h^0 = \operatorname{span}\{k_{\mathrm{B}}(\cdot,t_r) : r \in \mathcal{J}\} \subset \mathcal{H}_{\mathrm{B}}, \qquad \|g\|_{B,h} = \|g\|_{\mathcal{H}_{\mathrm{B}}} \quad (g \in \mathcal{H}_h^0).$$

Thus the inclusion is isometric, and the $\mathcal{H}_{\mathrm{B}}$-orthogonal projection $\Pi_h : \mathcal{H}_{\mathrm{B}} \to \mathcal{H}_h^0$ is exactly nodal interpolation.

Corollary 3 (Exact residual kernel and sharp approximation)**.** *Let $r_h := k_{\mathrm{B}} - k_h$ and let $T_i = [a_i,b_i] = [t_i,t_{i+1}]$. Then*

$$r_h(s,t) = \begin{cases} \dfrac{(\min\{s,t\} - a_i)\,(b_i - \max\{s,t\})}{h}, & s,t \in T_i, \\ 0, & s,t \text{ lie in different grid elements}. \end{cases} \tag{1}$$

Consequently, for $t \in T_i$ the power function is

$$p_h(t)^2 := r_h(t,t) = \frac{(t-a_i)(b_i-t)}{h}, \qquad |f(t) - \Pi_h f(t)| \le p_h(t)\|f\|_{\mathcal{H}_{\mathrm{B}}}. \tag{2}$$

For every $f \in \mathcal{H}_{\mathrm{B}}$,

$$\|f - \Pi_h f\|_{\mathcal{H}_{\mathrm{B}}}^2 = \|f\|_{\mathcal{H}_{\mathrm{B}}}^2 - \|\Pi_h f\|_{\mathcal{H}_{\mathrm{B}}}^2,$$

$$\|f - \Pi_h f\|_\infty \le \frac{\sqrt{h}}{2}\|f\|_{\mathcal{H}_{\mathrm{B}}}, \qquad \|f - \Pi_h f\|_{L^2(I)} \le \frac{h}{\pi}\|f\|_{\mathcal{H}_{\mathrm{B}}}.$$

Both constants are sharp over the unit ball. These quantities depend only on the subspace $\mathcal{H}_h^0$, so they are identical under every coordinate realization of that space.

The residual has the Brownian-bridge covariance form, and the bounds are standard power-function and sharp interval-inequality consequences (Rasmussen & Williams, 2006; Fasshauer, 2007; Brenner & Scott, 2008); their role is to certify identical approximation across coordinates.

4 Equivalent Coordinate Representations

The finite Brownian RKHS admits nodal, increment, and spectral realizations. Their invertible linear relation is elementary; the substantive control is that they realize exactly the same space, Brownian norm, and approximation operator. Consequently, any mapped optimization difference must arise from the optimizer or the coordinate metric rather than representational capacity.

4.1 Nodal and Increment Coordinate Representations

We first consider the nodal finite-element realization and its associated increment representation.

Definition 5 (Nodal coordinates)**.** Every function $f \in \mathcal{H}_h^0$ admits the unique nodal representation $f = \mathcal{R}(\mathbf{R}\widetilde{\mathbf{v}})$ with reduced anchored vector $\widetilde{\mathbf{v}} \in \mathbb{R}^G$, and $\|f\|_{B,h}^2 = \widetilde{\mathbf{v}}^\top \mathbf{K}_0 \widetilde{\mathbf{v}} = \mathbf{v}^\top \mathbf{K} \mathbf{v}$ for $\mathbf{v} = \mathbf{R}\widetilde{\mathbf{v}}$. Throughout, $\widetilde{\mathbf{v}} \in \mathbb{R}^G$ is the free parameter; the anchor is not a trainable coordinate.

The nodal realization is the parameterization induced directly by the finite-element discretization: each trainable parameter is the value of the profile at a grid point.

Definition 6 (Increment coordinates)**.** The increment coordinates associated with the reduced nodal vector $\widetilde{\mathbf{v}}$ are defined by $\mathbf{w} = \mathbf{D}_0\widetilde{\mathbf{v}}$, or equivalently, $w_i = v_{i+1} - v_i$ for $i = 0, \dots, G-1$ with $\mathbf{v} = \mathbf{R}\widetilde{\mathbf{v}}$.

The increment realization stores local differences rather than grid values. Because the profile is anchored, $\mathbf{D}_0$ is invertible (Theorem 2ii) and the increments determine f uniquely. In these coordinates the Brownian energy is simply $\|f\|_{B,h}^2 = \frac{1}{h}\|\mathbf{w}\|_2^2$.

4.2 Spectral Coordinate Representation

A complementary realization is obtained by diagonalizing the anchored Brownian energy, yielding an orthogonal spectral parameterization of the finite Brownian RKHS.

Definition 7 (Spectral coordinates)**.** By Lemma 3.1, $\mathbf{K}_0 \succ 0$. Let $\mathbf{K}_0 = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top$ be a fixed orthogonal eigendecomposition, with $\mathbf{Q} \in \mathbb{R}^{G\times G}$ orthogonal and $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_G)$ having strictly positive diagonal. For the reduced anchored nodal vector $\widetilde{\mathbf{v}} \in \mathbb{R}^G$, define $\mathbf{c} = \mathbf{Q}^\top\widetilde{\mathbf{v}}$, or equivalently, $\widetilde{\mathbf{v}} = \mathbf{Q}\mathbf{c}$.

These coordinates diagonalize the Brownian norm, so each spectral coefficient contributes independently to the total energy. The reconstruction and norm identities are proved in Appendix B.4; Section 5 gives the explicit DCT-VIII block-structured choice of $\mathbf{Q}$.

The three coordinate systems differ in how they store a Brownian profile: the nodal realization stores profile values, the increment realization stores local differences, and the spectral realization stores coefficients in an orthogonal Brownian basis. The following theorem shows that they are nevertheless equivalent at the level of the function space they represent.

Theorem 2 (Equivalence of implementations)**.** *Let $f \in \mathcal{H}_h^0$ be arbitrary, let $\widetilde{\mathbf{v}} \in \mathbb{R}^G$ denote its reduced anchored nodal coefficient vector, and set $\mathbf{v} = \mathbf{R}\widetilde{\mathbf{v}} \in \mathbb{R}^{G+1}$ and $\mathbf{D}_0 := \mathbf{D}\mathbf{R} \in \mathbb{R}^{G\times G}$. Define the increment and spectral coordinates by $\mathbf{w} = \mathbf{D}_0\widetilde{\mathbf{v}}$ and $\mathbf{c} = \mathbf{Q}^\top\widetilde{\mathbf{v}}$, respectively. Then the following statements hold.*

(i) The nodal, increment, and spectral representations determine exactly the same function $f \in \mathcal{H}_h^0$.

(ii) The maps $\widetilde{\mathbf{v}} \mapsto \mathbf{w}$ and $\widetilde{\mathbf{v}} \mapsto \mathbf{c}$ are invertible linear transformations. Consequently, each coordinate system uniquely determines the other two.

(iii) The Brownian norm admits the equivalent representations

$$\|f\|_{B,h}^2 = \mathbf{v}^\top\mathbf{K}\mathbf{v} = \frac{1}{h}\|\mathbf{w}\|_2^2 = \mathbf{c}^\top\mathbf{\Lambda}\mathbf{c}.$$

(iv) The three realizations therefore represent the same hypothesis class and, by Corollary 3, share the same approximation properties.

(v) The nodal-to-spectral transformation is an isometry for the Euclidean metric on the reduced anchored coordinates, whereas the nodal-to-increment transformation satisfies $\mathbf{D}_0^\top\mathbf{D}_0 = h\mathbf{K}_0$. Consequently, for $G = 2m \geq 4$, the increment coordinate map is nonorthogonal, whereas the spectral coordinate map is orthogonal.

The restriction $G \geq 4$ excludes only $G = 2$, where $h\mathbf{K}_0 = \mathbf{I}_2$ and the increment map is itself orthogonal. Theorem 2 gives both the function-space equivalence and the exact relation between the induced Euclidean metrics: the spectral map is orthogonal, the increment map induces the metric $h\mathbf{K}_0$. Section 6 shows that this one matrix identity determines which optimizers see the three realizations as the same problem.

5 Explicit Spectral Representation

Using classical DCT nomenclature (Martucci, 1994; Strang, 1999; Masera et al., 2017), we give all eigenpairs of the fixed mixed-boundary stiffness matrix in closed block DCT-VIII form; "complete" does not mean a new transform family.

5.1 Block Decomposition

After anchoring, the stiffness matrix decouples into two independent tridiagonal blocks. The reason is that the anchor node $t_{i_0} = 0$ is the unique node coupling the two half-grids; removing it severs them. In the reduced ordering of Definition 4, which traverses both half-grids from the exterior endpoint inward, the two blocks are literally the same matrix.

Theorem 3 (Complete DCT-VIII eigendecomposition)**.** *Let $\mathbf{T}_m = (\tau_{rs})_{r,s=0}^{m-1} \in \mathbb{R}^{m\times m}$ be the tridiagonal matrix with $\tau_{00} = 1$, $\tau_{rr} = 2$ for $1 \le r \le m-1$, $\tau_{rs} = -1$ for $|r-s| = 1$, and $\tau_{rs} = 0$ otherwise. Then $\mathbf{K}_0 = \frac{1}{h}\,\mathrm{blkdiag}(\mathbf{T}_m, \mathbf{T}_m)$ (Lemma D3). For each $k \in [m]$, define $\theta_k = \frac{(2k-1)\pi}{2m+1}$ and $\nu_k = 2 - 2\cos(\theta_k) = 4\sin^2\left(\frac{\theta_k}{2}\right)$, and let the vector $\mathbf{q}_k \in \mathbb{R}^m$ be defined componentwise by $q_k(j) = \frac{2}{\sqrt{2m+1}}\cos\left(\left(j+\frac{1}{2}\right)\theta_k\right)$ for $j = 0, \dots, m-1$.*

Then the following statements hold.

(i) The vectors $\mathbf{q}_1, \dots, \mathbf{q}_m$ form an orthonormal basis of $\mathbb{R}^m$.

(ii) With $\mathbf{Q}_m = [\mathbf{q}_1, \dots, \mathbf{q}_m]$, the matrix $\mathbf{T}_m$ admits the eigendecomposition $\mathbf{T}_m = \mathbf{Q}_m\,\mathrm{diag}(\nu_1, \dots, \nu_m)\mathbf{Q}_m^\top$.

(iii) The anchored stiffness matrix admits the orthogonal eigendecomposition $\mathbf{K}_0 = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top$, where $\mathbf{Q} = \mathrm{blkdiag}(\mathbf{Q}_m, \mathbf{Q}_m)$ and $\mathbf{\Lambda} = \frac{1}{h}\,\mathrm{diag}(\nu_1, \dots, \nu_m, \nu_1, \dots, \nu_m)$.

(iv) The eigenvalues of $\mathbf{K}_0$ are $\lambda_k = \frac{4}{h}\sin^2\left(\frac{(2k-1)\pi}{2(2m+1)}\right)$ for $k \in [m]$, each of multiplicity two.

The eigendecomposition induces the basis $\psi_i := \mathcal{R}(\mathbf{R}\mathbf{Q}\mathbf{e}_i)$ of $\mathcal{H}_h^0$, satisfying $\langle \psi_i, \psi_j\rangle_{B,h} = \lambda_i \delta_{ij}$ (Lemma D6). Hence the finite RKHS kernel has the spectral basis expansion

$$k_h(s,t) = \sum_{i=1}^{G} \lambda_i^{-1}\psi_i(s)\psi_i(t).$$

This is a finite RKHS expansion; no input measure or integral-operator eigenproblem is required.

Remark 1. Every eigenvalue has multiplicity two, so the eigenbasis is not canonical inside each left–right eigenspace. We fix $\mathbf{Q} = \mathrm{blkdiag}(\mathbf{Q}_m, \mathbf{Q}_m)$, the explicit block-structured choice matching the reduced ordering. Any such orthogonal choice gives the same gradient descent trajectory after mapping, but coordinatewise adaptive methods may distinguish the choices.

Corollary 4 (Exact stiffness and coordinate conditioning)**.** *Under Assumption 1,*

$$\kappa_2(\mathbf{K}_0) = \frac{\cos^2\left(\frac{\pi}{2m+1}\right)}{\sin^2\left(\frac{\pi}{4m+2}\right)} = \frac{4}{\pi^2}(2m+1)^2 + O(1) = \Theta(G^2), \tag{3}$$

$$\sigma_{\min}(\mathbf{D}_0) = 2\sin\left(\frac{\pi}{4m+2}\right), \quad \sigma_{\max}(\mathbf{D}_0) = 2\cos\left(\frac{\pi}{2m+1}\right), \tag{4}$$

$$\kappa_2(\mathbf{D}_0) = \frac{\cos\left(\frac{\pi}{2m+1}\right)}{\sin\left(\frac{\pi}{4m+2}\right)} = \sqrt{\kappa_2(\mathbf{K}_0)} = \Theta(G). \tag{5}$$

The condition numbers are independent of A. Moreover, $h\lambda_{\max}(\mathbf{K}_0) \to 4$ and $\lambda_{\min}(\mathbf{K}_0) \sim \pi^2/[h(2m+1)^2]$ as $m \to \infty$.

The first identity gives the classical $O(h^{-2})$ stiffness conditioning; the second and third quantify the anisotropy of the actual nodal-to-increment coordinate map.

6 Coordinates and Optimizers

We begin with the standard chain-rule reparameterization identity (Gunasekar et al., 2018; Amid & Warmuth, 2020; Kristiadi et al., 2023). Let $L : \mathbb{R}^G \to \mathbb{R}$ be differentiable in reduced nodal coordinates and, for an invertible matrix $\mathbf{A}$, define $L_{\mathbf{A}}(\mathbf{z}) := L(\mathbf{A}^{-1}\mathbf{z})$.

Proposition 2 (Linear reparameterization is exact preconditioning). *If* $\mathbf{z}_{k+1} = \mathbf{z}_k - \eta_k \nabla L_{\mathbf{A}}(\mathbf{z}_k)$ *and* $\widetilde{\mathbf{v}}_k = \mathbf{A}^{-1}\mathbf{z}_k$, *then*

$$\widetilde{\mathbf{v}}_{k+1} = \widetilde{\mathbf{v}}_k - \eta_k (\mathbf{A}^\top \mathbf{A})^{-1} \nabla L(\widetilde{\mathbf{v}}_k). \tag{6}$$

The same identity holds for stochastic gradients computed from the same minibatches and transformed by the chain rule.

Proposition 3 (Coordinate–optimizer correspondence). *For corresponding initializations, identical scalar step sizes, identical minibatches, and exact arithmetic:*

(i) Since $\mathbf{c} = \mathbf{Q}^\top \widetilde{\mathbf{v}}$ *and* $\mathbf{Q}$ *is orthogonal, nodal and spectral gradient descent produce identical mapped iterates and hence identical functions at every step.*

(ii) Since $\mathbf{w} = \mathbf{D}_0 \widetilde{\mathbf{v}}$ *and* $\mathbf{D}_0^\top \mathbf{D}_0 = h\mathbf{K}_0$, *increment gradient descent maps to*

$$\widetilde{\mathbf{v}}_{k+1} = \widetilde{\mathbf{v}}_k - \frac{\eta_k}{h} \mathbf{K}_0^{-1} \nabla L(\widetilde{\mathbf{v}}_k). \tag{7}$$

Thus it is an explicit Euler step generated by the constant Brownian/Sobolev metric with Gram matrix $\mathbf{K}_0$, *with time step* η_k/h. *We use "Riemannian-gradient step" only in this discrete, constant-metric sense; the statement is not equality with the continuous gradient flow.*

Corollary 5 (Blockwise extension to recursive architectures). *Let an arbitrary differentiable composed objective depend on profile blocks* $\widetilde{\mathbf{v}}^{(1)}, \dots, \widetilde{\mathbf{v}}^{(B)}$ *and possibly other parameters, and apply the block transformation* $\mathbf{A} = \mathrm{blkdiag}(\mathbf{A}_1, \dots, \mathbf{A}_B, \mathbf{I})$. *Then Proposition 2 holds with block preconditioner* $\mathrm{blkdiag}((\mathbf{A}_1^\top \mathbf{A}_1)^{-1}, \dots, (\mathbf{A}_B^\top \mathbf{A}_B)^{-1}, \mathbf{I})$. *Hence simultaneous spectral changes leave the full GD/SGD trajectory invariant, while simultaneous increment coordinates implement blockwise Brownian Riemannian-gradient steps with the corresponding per-block factors* $1/h_b$. *No quadratic or layer-separability assumption is required.*

This is the block-diagonal application of Proposition 2; cross-block coupling does not invalidate the identity. Here "mesh-independent" means independent of G for fixed A, $\rho > 0$, sample domain, and normalization, not uniform in A or ρ.

The previous identities describe trajectories. The next result gives the corresponding quantitative conditioning statement for the regularized regression problem generated by the finite RKHS.

Proposition 4 (Mesh-independent increment conditioning). *For samples* $x_1, \dots, x_n \in I$, *targets* $y_1, \dots, y_n \in \mathbb{R}$, *and* $\rho > 0$, *consider*

$$L(\widetilde{\mathbf{v}}) = \frac{1}{2n} \sum_{i=1}^{n} \left(\widetilde{\boldsymbol{\phi}}(x_i)^\top \widetilde{\mathbf{v}} - y_i \right)^2 + \frac{\rho}{2} \widetilde{\mathbf{v}}^\top \mathbf{K}_0 \widetilde{\mathbf{v}}.$$

Nodal and spectral Hessians are orthogonally similar. In increment coordinates $\mathbf{w} = \mathbf{D}_0 \widetilde{\mathbf{v}}$, *the Hessian* $\mathbf{H}_{\mathrm{inc}}$ *satisfies*

$$\frac{\rho}{h} \mathbf{I}_G \preceq \mathbf{H}_{\mathrm{inc}} \preceq \frac{\rho + A}{h} \mathbf{I}_G, \qquad \kappa_2(\mathbf{H}_{\mathrm{inc}}) \le 1 + \frac{A}{\rho}, \tag{8}$$

independently of G. *For the pure Brownian quadratic, the increment condition number is exactly* 1, *whereas the nodal and spectral condition numbers equal* $\kappa_2(\mathbf{K}_0) = \Theta(G^2)$.

Standard Adam requires a separate statement because momentum prevents its update from being represented as a diagonal matrix times only the current gradient.

Proposition 5 (Exact orthogonal equivariance group of standard Adam). *Consider standard Adam with scalar* $\beta_1, \beta_2 \in [0, 1)$, *scalar* $\varepsilon > 0$, *zero initial moment states, bias correction, a positive first step size, and no weight decay, clipping, or other modification. Among orthogonal coordinate changes, Adam is equivariant for every gradient sequence and corresponding initialization if and only if the coordinate change is a signed permutation. For* $m \ge 2$, *the DCT-VIII matrix* $\mathbf{Q}$ *is not a signed permutation; therefore there exists even a linear objective for which corresponding nodal and spectral Adam runs differ at the first step.*

This maximality statement strengthens basis dependence to an if-and-only-if universal group under the exact convention; it does not cover AdamW, clipping, nonzero moments, or matrix-valued adaptivity.

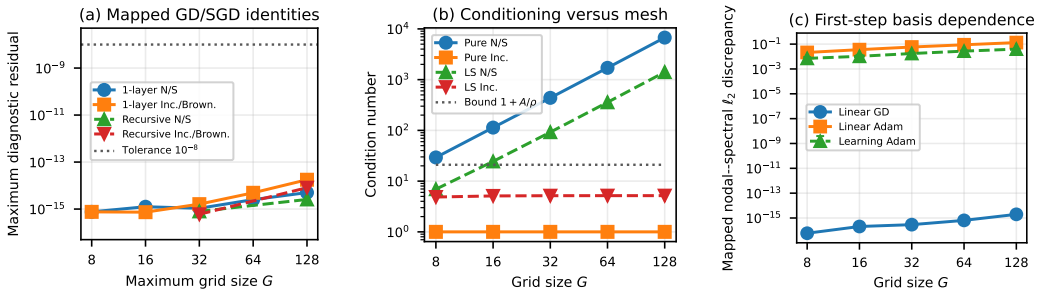

Figure 1: Direct numerical verification of the coordinate–optimizer theory. (a) Each marker is the maximum over the recorded parameter, function, objective, gradient, and preconditioner diagnostics and over full-batch GD and shared-minibatch SGD. The recursive markers correspond to grid triples $(8, 16, 32)$ and $(32, 64, 128)$, plotted at their maximum G; dashed segments guide the eye. All measured residuals remain far below the predeclared 10^{-8} tolerance. (b) For the pure Brownian quadratic, nodal/spectral conditioning grows with the mesh whereas increment conditioning is exactly one. Least-squares markers are five-seed means; increment conditioning remains independent of G under the stated fixed quantities and below $1 + A/\rho = 21$. (c) Ordinary GD remains equivariant, whereas standard Adam has a nonzero mapped nodal–spectral first-step discrepancy. Learning-objective markers are five-seed means and bars span the minimum-to-maximum range. The ordinate is a coordinate-mapping discrepancy, not a predictive loss or error. These results establish permitted basis dependence, not a universal performance ordering.

Thus GD gives exact nodal–spectral agreement; increment GD gives exact Brownian preconditioning; and Adam is permitted, but not forced, to separate the orthogonally related realizations. A separation theorem does not impose a universal ordering, and independently sampled initializations do not test trajectory equivalence. These predictions require mapped initial states, the same data order and scalar step sequence, and the same intrinsic Brownian penalty.

7 Experiments

We separate direct theorem-verification tests from predictive comparisons. The verification suite comprises E0, the one-layer optimization tests E1A–E1C, and the recursive extension E2. All five tests use frozen configurations, float64 arithmetic, one common represented initialization, exact coordinate maps, the same scalar step sequences, and, for SGD, identical minibatches; no identity-test arm is tuned separately. Secondary EuroSAT and spatially separated Salinas studies compare different model classes and are reported only in the appendix.

Finite and mapped identities. Across $G \in \{8, 16, 32, 64, 128\}$, the matrix, DCT-VIII spectrum, Brownian-Gram inverse, reconstruction, and intrinsic-energy residuals remain at floating-point precision. Mapped nodal–spectral full-batch GD and shared-minibatch SGD agree in parameters, functions, objectives, and gradients. Increment GD agrees with the correctly scaled Brownian-preconditioned nodal update. The same conclusions hold for a recursively coupled three-profile model with all non-profile parameters trained and shared across arms.

Conditioning and standard Adam. For the pure Brownian quadratic, nodal/spectral conditioning grows from 29.3 at $G = 8$ to 6740.7 at $G = 128$, whereas increment conditioning is exactly one. For Brownian-regularized least squares with $\rho = 0.1$, the largest measured increment condition number is 5.64, below the stated G-independent bound $1 + A/\rho = 21$, while nodal/spectral conditioning and fixed-step iteration counts increase with G. A functional standard-Adam implementation matches `torch.optim.Adam`, but mapped nodal and spectral first steps differ on both the deterministic linear counterexample and a nontrivial regularized learning objective. This demonstrates permitted basis dependence, not a universal performance ordering.

Table 1: Float64 verification of the exact finite-Brownian and mapped-optimizer identities. The last four rows report maxima over the listed diagnostics; these are identity checks, not predictive-performance comparisons.

Check	Scope	Maximum residual	Type
$\mathbf{D}_0^\top \mathbf{D}_0 = h\mathbf{K}_0$	5 grids	0	absolute
$\mathbf{K}_0 = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top$	5 grids	9.99×10^{-15}	relative
Closed-form spectrum	5 grids	1.41×10^{-13}	relative
$\mathbf{K}_0^{-1}$ versus Brownian Gram	5 grids	0	absolute
Coordinate reconstruction	125 random trials	3.12×10^{-14}	absolute
Intrinsic Brownian energy	125 random trials	3.99×10^{-15}	relative
One-layer Nodal–Spectral GD/SGD	50 mapped runs	5.02×10^{-15}	diagnostic envelope
One-layer Increment–Brownian GD/SGD	50 mapped runs	1.77×10^{-14}	diagnostic envelope
Recursive Nodal–Spectral GD/SGD	20 mapped runs	2.57×10^{-15}	diagnostic envelope
Recursive Increment–Brownian GD/SGD	20 mapped runs	8.10×10^{-15}	diagnostic envelope

Table 2: Conditioning, fixed-step convergence, and standard-Adam basis dependence across meshes. "N/S" denotes identical nodal/spectral quantities and "Inc." the increment realization. Iteration counts use the analytic optimal scalar step and relative objective-gap tolerance 10^{-8}; least squares uses $\rho = 0.1$, with increment bound $1 + A/\rho = 21$. The last columns report mapped nodal–spectral first-step discrepancies.

G	$\kappa(\mathbf{K}_0)$	Pure iters N/S / Inc.	LS κ N/S / Inc.	LS iters N/S / Inc.	Adam Δ_1 linear	Adam Δ_1 learning
8	29.3	123.2 / 1.0	6.94 / 4.86	31.2 / 21.6	0.02	7.09×10^{-3}
16	113.5	467.0 / 1.0	24.45 / 5.11	109.6 / 22.6	0.04	0.01
32	437.7	1795.2 / 1.0	92.15 / 5.16	413.0 / 23.4	0.06	0.02
64	1.7×10^3	7004.2 / 1.0	357.64 / 5.17	1601.8 / 23.4	0.09	0.03
128	6.7×10^3	27629.0 / 1.0	1408.71 / 5.17	6310.2 / 23.4	0.13	0.04

External validity. Appendix G reports independently audited EuroSAT and spatially separated Salinas BKL-versus-MLP robustness studies. They provide secondary predictive evidence, not tests of coordinate equivariance.

8 Conclusion

Equivalent coordinates of one finite function space need not define equivalent optimizers. Classical finite-element, interpolation, covariance, and DCT identities provide our fixed Brownian comparison model. Within it, nodal and spectral GD/SGD coincide; increment descent is a $1/h$-scaled constant-metric Brownian/Sobolev Euler step, also blockwise; the least-squares condition number is at most $1 + A/\rho$ independently of G under the stated fixed quantities; and signed permutations are the maximal universal orthogonal equivariances of standard Adam. Float64 tests numerically verify these predictions.

Limitations. Closed forms assume a one-dimensional uniform grid and fixed anchor. The least-squares bound is not uniform in A or ρ or a nonlinear convergence rate; the Adam theorem covers only its stated update and gives no coordinate ranking. EuroSAT and Salinas compare model classes rather than coordinates, and eigenvalue multiplicity makes the displayed DCT-VIII basis noncanonical within each two-dimensional eigenspace.

References

Shun-ichi Amari. Natural gradient works efficiently in learning. *Neural Computation*, 10(2):251–276, 1998. doi: 10.1162/089976698300017746.

Ehsan Amid and Manfred K. Warmuth. Reparameterizing mirror descent as gradient descent. In *Advances in Neural Information Processing Systems*, volume 33, pp. 8430–8439, 2020.

Nachman Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68(3):337–404, 1950.

Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

Alain Berlinet and Christine Thomas-Agnan. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Kluwer Academic Publishers, 2004.

Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 3 edition, 2008.

Gregory E. Fasshauer. *Meshfree Approximation Methods with MATLAB*. World Scientific, 2007.

Suriya Gunasekar, Jason D. Lee, Daniel Soudry, and Nathan Srebro. Characterizing implicit bias in terms of optimization geometry. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 1832–1841, 2018.

Arthur Jacot, Franck Gabriel, and Clément Hongler. Neural tangent kernel: Convergence and generalization in neural networks. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015.

Agustinus Kristiadi, Felix Dangel, and Philipp Hennig. The geometry of neural nets' parameter spaces under reparametrization. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Jaehoon Lee, Lechao Xiao, Samuel S. Schoenholz, Yasaman Bahri, Roman Novak, Jascha Sohl-Dickstein, and Jeffrey Pennington. Wide neural networks of any depth evolve as linear models under gradient descent. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Zhiyuan Li, Tianhao Wang, Jason D. Lee, and Sanjeev Arora. Implicit bias of gradient descent on reparametrized models: On equivalence to mirror descent. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Selena Ling, Nicholas Sharp, and Alec Jacobson. VectorAdam for rotation equivariant geometry optimization, 2022.

James Martens. New insights and perspectives on the natural gradient method. *Journal of Machine Learning Research*, 21(146):1–76, 2020.

Stephen A. Martucci. Symmetric convolution and the discrete sine and cosine transforms. *IEEE Transactions on Signal Processing*, 42(5):1038–1051, 1994.

Maurizio Masera, Maurizio Martina, and Guido Masera. Odd type DCT/DST for video coding: Relationships and low-complexity implementations. In *2017 IEEE International Workshop on Signal Processing Systems (SiPS)*, pp. 1–6, 2017. doi: 10.1109/SiPS.2017.8110009.

Mahdi Mohammadigohari. Variation brownian kernel ladders, 2026. URL https://arxiv.org/abs/2608.13882.

Mahdi Mohammadigohari, Giuseppe Di Fatta, Giuseppe Nicosia, and Panos M. Pardalos. Brownian kernel ladders, 2026. URL https://arxiv.org/abs/2606.15812.

Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Advances in Neural Information Processing Systems*, volume 20, 2007.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

Bernhard Schölkopf and Alexander J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2002.

Ingo Steinwart and Andreas Christmann. *Support Vector Machines*. Springer, 2008.

Gilbert Strang. The discrete cosine transform. *SIAM Review*, 41(1):135–147, 1999. doi: 10.1137/S0036144598336745.

Gilbert Strang and George J. Fix. *An Analysis of the Finite Element Method*. Prentice-Hall, 1973.

Grace Wahba. *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, 1990.

Holger Wendland. *Scattered Data Approximation*. Cambridge University Press, 2005.

Tianyue H. Zhang, Lucas Maes, Alan Milligan, Alexia Jolicoeur-Martineau, Ioannis Mitliagkas, Damien Scieur, Simon Lacoste-Julien, and Charles Guille-Escuret. Understanding Adam requires better rotation-dependent assumptions. In *Advances in Neural Information Processing Systems*, 2025.

AI USE STATEMENT

Generative AI tools were used to assist with the writing and revision of mathematical proofs, feedback on experimental design, editing of research code, language editing, LaTeX formatting, and literature search. All AI-assisted material incorporated into the paper was reviewed and checked for correctness. The authors take responsibility for the final content of this work, including its mathematical claims, proofs, code, experiments, citations, and other artifacts.

A ADDITIONAL NOTATION

Table 3: Summary of the main theoretical results.

Result	Content
Proposition 1	Classical stiffness identity in the Brownian normalization
Theorem 1	Brownian-specific kernel and covariance–precision formula
Corollaries 2 and 3	Classical interpolation consequences with exact constants
Theorem 2	Exact equivalence of the three realizations
Theorem 3 and Corollary 4	Closed-form block DCT-VIII specialization
Propositions 2 and 3	Standard reparameterization and Brownian metric consequence
Corollary 5	Exact blockwise recursive extension
Proposition 4	G-independent conditioning for fixed A and ρ
Proposition 5	Maximal universal orthogonal group of standard Adam

The reduced anchored ordering, the anchoring reconstruction matrix $\mathbf{R}$, the anchored stiffness matrix $\mathbf{K}_0 = \mathbf{R}^\top \mathbf{K}\mathbf{R}$, and $\mathbf{D}_0 = \mathbf{D}\mathbf{R}$ are now defined in the main text (Definition 4), together with the standing assumption $G = 2m$, $i_0 = m$ (Assumption 1); they are not repeated here. For reference, the reduced ordering is

$$\widetilde{\mathbf{v}} := (v_0, \dots, v_{m-1}, v_{2m}, v_{2m-1}, \dots, v_{m+1})^\top \in \mathbb{R}^{2m}, \qquad \mathbf{R}\widetilde{\mathbf{v}} = \mathbf{v}, \tag{9}$$

so that both half-grids run from their exterior endpoint toward the anchor, and

$$\mathbf{R} := (\mathbf{e}_0, \dots, \mathbf{e}_{m-1}, \mathbf{e}_{2m}, \mathbf{e}_{2m-1}, \dots, \mathbf{e}_{m+1}) \in \mathbb{R}^{(2m+1)\times 2m}. \tag{10}$$

That the Brownian energy has a one-dimensional constant nullspace, so that anchoring yields $\mathbf{K}_0 \succ 0$, is established by Lemmas 3.1 and D2; it is proved rather than assumed.

The first-difference operator $\mathbf{D} : \mathbb{R}^{G+1} \to \mathbb{R}^G$ is defined by $(\mathbf{D}\mathbf{v})_i = v_{i+1} - v_i$ for $i = 0, \dots, G-1$, and the finite-element stiffness matrix $\mathbf{K} = (K_{rs})_{r,s=0}^{G}$ entrywise by $K_{rs} = \int_{-A}^{A} \phi_r'(t)\phi_s'(t)\,\mathrm{d}t$.

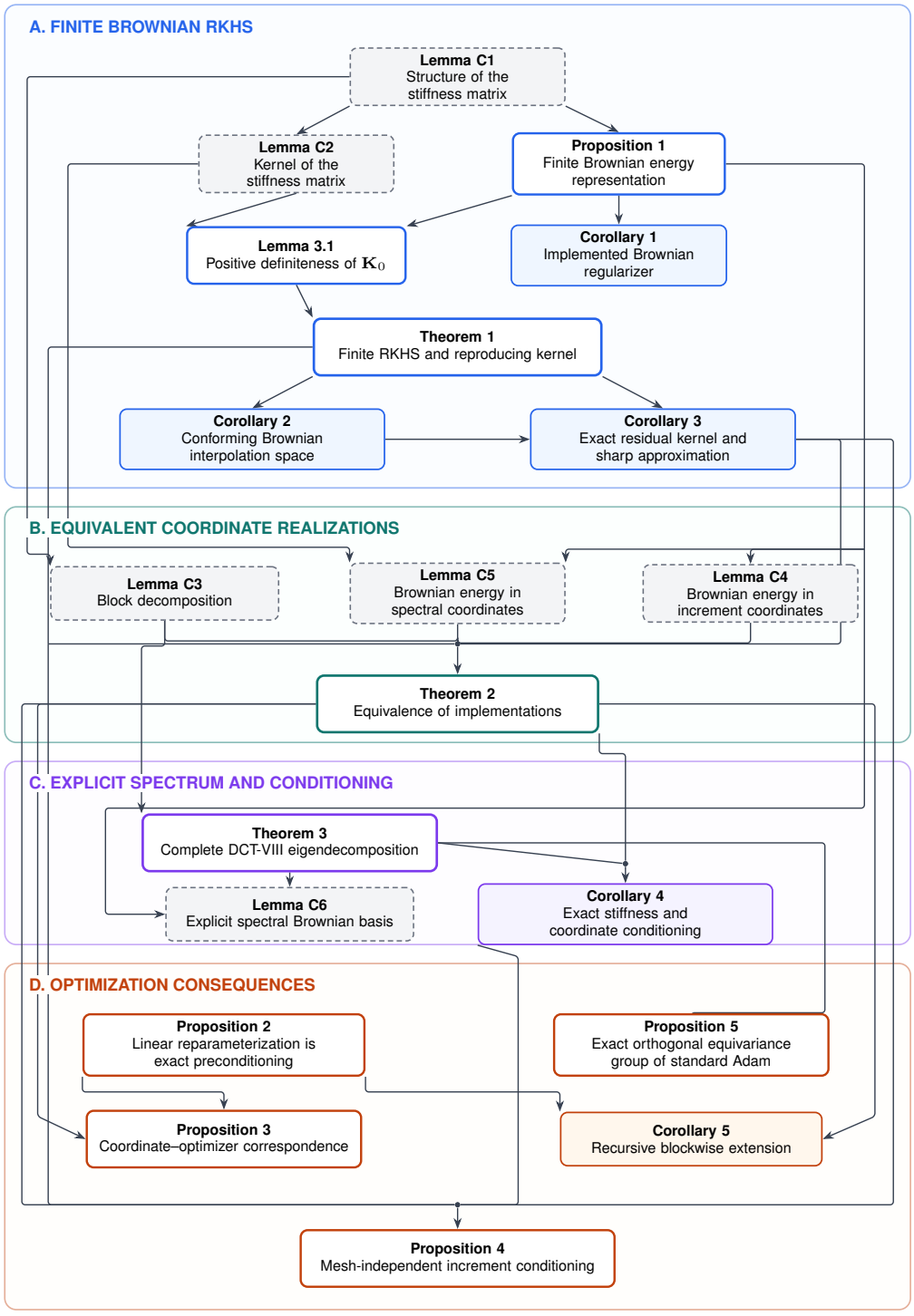

Figure 2: Proof-dependency structure of the theoretical results. The displayed artwork is the supplied dependency graph without redrawing or re-typesetting; panel headings and result nodes are clickable and jump to the corresponding section or statement.

B PROOFS

This section proves the foundational finite Brownian and coordinate results. We first establish the finite Brownian energy and stiffness representation in Proposition 1 and its implemented-regularizer consequence in Corollary 1. We then prove the G-dimensional Hilbert-space and reproducing-kernel structure asserted in Theorem 1; the explicit kernel formula and Brownian Gram identification in items i–ii are isolated in Section C. Next, we prove the exact equivalence of the nodal, increment, and spectral realizations in Theorem 2, including the coordinate maps, Brownian-energy identities, and induced metric relations. Finally, we derive the complete DCT-VIII eigendecomposition in Theorem 3. The auxiliary stiffness, nullspace, block-decomposition, and coordinate-energy results used throughout are collected in Section D.

B.1 PROOF OF PROPOSITION 1

Since $A > 0$ and the grid is strictly increasing, its mesh size satisfies $h = \frac{2A}{G} > 0$. Because $f = \mathcal{R}(\mathbf{v})$, the definition of the reconstruction operator gives, for every $t \in I$,

$$f(t) \overset{(a)}{=} \mathcal{R}(\mathbf{v})(t) \overset{(b)}{=} \sum_{i=0}^{G} v_i \phi_i(t). \tag{11}$$

Here (a) uses the assumed identity $f = \mathcal{R}(\mathbf{v})$, and (b) is the definition of $\mathcal{R}$. Each basis function ϕ_i is continuous and piecewise affine on the finite grid, and is therefore absolutely continuous on I and differentiable away from the grid points. Consequently, f is absolutely continuous and, for almost every $t \in I$,

$$f'(t) \overset{(a)}{=} \left(\sum_{i=0}^{G} v_i \phi_i(t) \right)' \overset{(b)}{=} \sum_{i=0}^{G} v_i \phi_i'(t). \tag{12}$$

Here (a) differentiates (11) at a point at which all basis derivatives exist, and (b) uses the linearity of differentiation and the fact that each coefficient v_i is independent of t. The possible failure of differentiability at the finitely many grid points does not affect any of the integrals below.

We first identify the stiffness-matrix representation of the Brownian energy. By Definition 2,

$$\begin{aligned}
\|f\|_{B,h}^2 &\overset{(a)}{=} \langle f, f \rangle_{B,h} \overset{(b)}{=} \int_{-A}^{A} f'(t) f'(t)\, \mathrm{d}t \\
&\overset{(c)}{=} \int_{-A}^{A} \left(\sum_{r=0}^{G} v_r \phi_r'(t) \right) \left(\sum_{s=0}^{G} v_s \phi_s'(t) \right) \mathrm{d}t \\
&\overset{(d)}{=} \int_{-A}^{A} \sum_{r=0}^{G} \sum_{s=0}^{G} v_r v_s \phi_r'(t) \phi_s'(t)\, \mathrm{d}t \\
&\overset{(e)}{=} \sum_{r=0}^{G} \sum_{s=0}^{G} v_r v_s \int_{-A}^{A} \phi_r'(t) \phi_s'(t)\, \mathrm{d}t \\
&\overset{(f)}{=} \sum_{r=0}^{G} \sum_{s=0}^{G} v_r \langle \phi_r, \phi_s \rangle_{B,h} v_s \overset{(g)}{=} \sum_{r=0}^{G} \sum_{s=0}^{G} v_r K_{rs} v_s \\
&\overset{(h)}{=} \mathbf{v}^\top \mathbf{K} \mathbf{v}.
\end{aligned} \tag{13}$$

Here (a) is the definition of the Brownian seminorm associated with the symmetric bilinear energy form, (b) is the definition of that energy form, (c) substitutes (12), (d) expands the product of the two finite sums, (e) interchanges the integral with finite sums by linearity, (f) applies the definition of $\langle \phi_r, \phi_s \rangle_{B,h}$, (g) uses the definition $K_{rs} = \langle \phi_r, \phi_s \rangle_{B,h}$, and (h) is the componentwise expansion of the quadratic form $\mathbf{v}^\top \mathbf{K} \mathbf{v}$.

By Lemma D1,

$$\mathbf{K} \overset{(a)}{=} \frac{1}{h} \mathbf{D}^\top \mathbf{D}. \tag{14}$$

Here (a) invokes the stiffness-matrix factorization proved in Lemma D1. In particular, the dimensions $\mathbf{D} \in \mathbb{R}^{G\times(G+1)}$ and $\mathbf{v} \in \mathbb{R}^{G+1}$ ensure that $\mathbf{Dv} \in \mathbb{R}^G$ and that every product below is well defined. Using (14),

$$\mathbf{v}^\top \mathbf{K}\mathbf{v} \overset{(a)}{=} \mathbf{v}^\top \left(\frac{1}{h}\mathbf{D}^\top\mathbf{D}\right)\mathbf{v} \overset{(b)}{=} \frac{1}{h}\mathbf{v}^\top\mathbf{D}^\top\mathbf{D}\mathbf{v}$$
$$\overset{(c)}{=} \frac{1}{h}\left(\mathbf{D}\mathbf{v}\right)^\top\left(\mathbf{D}\mathbf{v}\right) \overset{(d)}{=} \frac{1}{h}\left\|\mathbf{D}\mathbf{v}\right\|_2^2. \tag{15}$$

Here (a) substitutes (14), (b) moves the scalar $\frac{1}{h}$ outside the quadratic form, (c) uses the transpose identity $(\mathbf{Dv})^\top = \mathbf{v}^\top\mathbf{D}^\top$ and associativity of matrix multiplication, and (d) applies $\mathbf{w}^\top\mathbf{w} = \|\mathbf{w}\|_2^2$ with $\mathbf{w} = \mathbf{Dv}$.

By the definition of the Euclidean norm and the componentwise definition of $\mathbf{D}$,

$$\left\|\mathbf{D}\mathbf{v}\right\|_2^2 \overset{(a)}{=} \sum_{i=0}^{G-1}\left(\left(\mathbf{D}\mathbf{v}\right)_i\right)^2 \overset{(b)}{=} \sum_{i=0}^{G-1}\left(v_{i+1}-v_i\right)^2. \tag{16}$$

Here (a) expands the squared Euclidean norm of the vector $\mathbf{Dv} \in \mathbb{R}^G$, and (b) uses $(\mathbf{Dv})_i = v_{i+1} - v_i$ for every $i = 0, \ldots, G-1$. Combining (13), (15), and (16) yields

$$\left\|f\right\|_{B,h}^2 \overset{(a)}{=} \mathbf{v}^\top\mathbf{K}\mathbf{v} \overset{(b)}{=} \frac{1}{h}\left\|\mathbf{D}\mathbf{v}\right\|_2^2 \overset{(c)}{=} \frac{1}{h}\sum_{i=0}^{G-1}\left(v_{i+1}-v_i\right)^2.$$

Here (a) is (13), (b) is (15), and (c) substitutes (16). Finally, (14) is exactly the equivalent matrix identity asserted in the proposition. Therefore all claimed representations follow.

□

B.2 Proof of Corollary 1

Fix an arbitrary vector $\mathbf{v} = (v_0, \ldots, v_G)^\top \in \mathbb{R}^{G+1}$ and define

$$f := \mathcal{R}\left(\mathbf{v}\right). \tag{17}$$

By Definition 1, $\mathcal{R}$ maps $\mathbb{R}^{G+1}$ into $\mathcal{H}_h$. Consequently,

$$f \overset{(a)}{=} \mathcal{R}\left(\mathbf{v}\right) \overset{(b)}{\in} \mathcal{H}_h. \tag{18}$$

Here (a) is the definition (17), and (b) follows from $\mathcal{R}(\mathbb{R}^{G+1}) = \mathcal{H}_h$ and $\mathbf{v} \in \mathbb{R}^{G+1}$. Therefore Proposition 1 applies to f and its nodal vector $\mathbf{v}$.

By the definition of the implemented finite Brownian regularizer,

$$\Omega_h\left(\mathbf{v}\right) := \sum_{i=0}^{G-1}\left(v_{i+1}-v_i\right)^2.$$

On the other hand, Proposition 1, applied using (18), gives

$$\left\|\mathcal{R}\left(\mathbf{v}\right)\right\|_{B,h}^2 \overset{(a)}{=} \frac{1}{h}\sum_{i=0}^{G-1}\left(v_{i+1}-v_i\right)^2. \tag{19}$$

Here (a) is the final identity in Proposition 1, applied to the function $f = \mathcal{R}(\mathbf{v})$ established in (18).

Since $h = 2A/G > 0$, multiplying (19) by h gives

$$\Omega_h\left(\mathbf{v}\right) = h\left\|\mathcal{R}\left(\mathbf{v}\right)\right\|_{B,h}^2.$$

As $\mathbf{v} \in \mathbb{R}^{G+1}$ was arbitrary and h depends only on the fixed grid, the implemented regularizer is the discrete Brownian energy scaled by the fixed positive factor h.

□

B.3 Proof of Theorem 1: RKHS existence and dimension

We first identify the linear structure and dimension of the anchored space. Set

$$\mathcal{I}_0 := \{0, \dots, G\} \setminus \{i_0\}. \tag{20}$$

Since $t_{i_0} = 0$ and the nodal basis satisfies $\phi_j(t_r) = \delta_{jr}$, we have, for every $j = 0, \dots, G$,

$$\phi_j(0) \overset{(a)}{=} \phi_j(t_{i_0}) \overset{(b)}{=} \delta_{ji_0}. \tag{21}$$

Here (a) uses $t_{i_0} = 0$, and (b) is the nodal interpolation property of the finite-element basis. Let $f \in \mathcal{H}_h$ be arbitrary, and let $\mathbf{v} = (v_0, \dots, v_G)^\top$ be its unique nodal vector, so that $f = \mathcal{R}(\mathbf{v})$. Evaluating this representation at the anchor and using (21) gives

$$f(0) \overset{(a)}{=} \mathcal{R}(\mathbf{v})(0) \overset{(b)}{=} \sum_{j=0}^{G} v_j \phi_j(0) \overset{(c)}{=} \sum_{j=0}^{G} v_j \delta_{ji_0} \overset{(d)}{=} v_{i_0}. \tag{22}$$

Here (a) uses $f = \mathcal{R}(\mathbf{v})$; (b) is the definition of the reconstruction operator; (c) substitutes (21); and (d) uses the defining property of the Kronecker delta, which removes every term except the term with index $j = i_0$. Consequently,

$$f \in \mathcal{H}_h^0 \overset{(a)}{\iff} f(0) = 0 \overset{(b)}{\iff} v_{i_0} = 0 \overset{(c)}{\iff} f = \sum_{j \in \mathcal{I}_0} v_j \phi_j. \tag{23}$$

Here (a) is the definition of $\mathcal{H}_h^0$; (b) uses (22); and (c) uses the unique nodal representation and the fact that the coefficient of ϕ_{i_0} is zero exactly when $v_{i_0} = 0$. It follows that

$$\mathcal{H}_h^0 \overset{(a)}{=} \operatorname{span}\{\phi_j : j \in \mathcal{I}_0\}. \tag{24}$$

Here (a) is precisely the set equality expressed by (23). In particular, $\mathcal{H}_h^0$ is a vector subspace of $\mathcal{H}_h$.

We next verify that the spanning family in (24) is linearly independent. Suppose that real coefficients $(a_j)_{j \in \mathcal{I}_0}$ satisfy

$$\sum_{j \in \mathcal{I}_0} a_j \phi_j = 0. \tag{25}$$

Fix an arbitrary $r \in \mathcal{I}_0$ and evaluate (25) at the grid point t_r. Then

$$0 \overset{(a)}{=} \sum_{j \in \mathcal{I}_0} a_j \phi_j(t_r) \overset{(b)}{=} \sum_{j \in \mathcal{I}_0} a_j \delta_{jr} \overset{(c)}{=} a_r.$$

Here (a) evaluates the zero-function identity (25) at t_r; (b) uses $\phi_j(t_r) = \delta_{jr}$; and (c) again uses the defining property of the Kronecker delta. Since $r \in \mathcal{I}_0$ was arbitrary, every coefficient a_r vanishes. Therefore $\{\phi_j : j \in \mathcal{I}_0\}$ is a basis of $\mathcal{H}_h^0$. Because $\mathcal{I}_0$ is obtained by removing one index from the $G + 1$ indices $0, \dots, G$,

$$\dim\left(\mathcal{H}_h^0\right) \overset{(a)}{=} |\mathcal{I}_0| \overset{(b)}{=} (G+1) - 1 \overset{(c)}{=} G. \tag{26}$$

Here (a) uses the basis just established; (b) uses the definition (20); and (c) simplifies the integer expression. Thus the anchored space is finite-dimensional.

We now prove that the Brownian energy form is an inner product on this space. Every function in $\mathcal{H}_h^0 \subset \mathcal{H}_h$ is continuous and piecewise affine on a finite partition of I. Hence it is absolutely continuous on I, its derivative exists almost everywhere, and its derivative is piecewise constant and therefore belongs to $L^2(I)$. For arbitrary $f, g \in \mathcal{H}_h^0$, the Cauchy–Schwarz inequality gives

$$\int_{-A}^{A} |f'(t)g'(t)| \, \mathrm{d}t \overset{(a)}{\leq} \left(\int_{-A}^{A} |f'(t)|^2 \, \mathrm{d}t\right)^{1/2} \left(\int_{-A}^{A} |g'(t)|^2 \, \mathrm{d}t\right)^{1/2} \overset{(b)}{<} \infty.$$

Here (a) is the Cauchy–Schwarz inequality on $L^2(I)$, and (b) follows from $f', g' \in L^2(I)$. Thus the integral defining $\langle f, g \rangle_{B,h}$ is finite and the form is well defined.

Let $f_1, f_2, g \in \mathcal{H}_h^0$ and $\alpha, \beta \in \mathbb{R}$. Because $\mathcal{H}_h^0$ is a vector space, the function $\alpha f_1 + \beta f_2$ also belongs to $\mathcal{H}_h^0$. Using linearity of the weak derivative and of the integral,

$$\begin{aligned}\langle \alpha f_1 + \beta f_2, g\rangle_{B,h} &\overset{(a)}{=} \int_{-A}^{A} (\alpha f_1 + \beta f_2)'(t) g'(t)\,\mathrm{d}t \\ &\overset{(b)}{=} \int_{-A}^{A} (\alpha f_1'(t) + \beta f_2'(t))\, g'(t)\,\mathrm{d}t \\ &\overset{(c)}{=} \alpha \int_{-A}^{A} f_1'(t) g'(t)\,\mathrm{d}t + \beta \int_{-A}^{A} f_2'(t) g'(t)\,\mathrm{d}t \\ &\overset{(d)}{=} \alpha \langle f_1, g\rangle_{B,h} + \beta \langle f_2, g\rangle_{B,h}. \end{aligned} \tag{27}$$

Here (a) is the definition of the Brownian energy form; (b) uses $(\alpha f_1 + \beta f_2)' = \alpha f_1' + \beta f_2'$ almost everywhere; (c) distributes the product and uses linearity of the integral; and (d) applies the definition of the Brownian energy form to the two resulting integrals. For arbitrary $f, g \in \mathcal{H}_h^0$, commutativity of multiplication in $\mathbb{R}$ yields

$$\langle f, g\rangle_{B,h} \overset{(a)}{=} \int_{-A}^{A} f'(t) g'(t)\,\mathrm{d}t \overset{(b)}{=} \int_{-A}^{A} g'(t) f'(t)\,\mathrm{d}t \overset{(c)}{=} \langle g, f\rangle_{B,h}. \tag{28}$$

Here (a) and (c) are the definition of the Brownian energy form, and (b) uses $f'(t)g'(t) = g'(t)f'(t)$ for almost every $t \in I$. Linearity in the second argument now follows explicitly from (27) and (28): for $f, g_1, g_2 \in \mathcal{H}_h^0$ and $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned}\langle f, \alpha g_1 + \beta g_2\rangle_{B,h} &\overset{(a)}{=} \langle \alpha g_1 + \beta g_2, f\rangle_{B,h} \\ &\overset{(b)}{=} \alpha \langle g_1, f\rangle_{B,h} + \beta \langle g_2, f\rangle_{B,h} \\ &\overset{(c)}{=} \alpha \langle f, g_1\rangle_{B,h} + \beta \langle f, g_2\rangle_{B,h}. \end{aligned}$$

Here (a) uses symmetry; (b) applies (27); and (c) uses symmetry separately in each term. Therefore the form is bilinear and symmetric. Moreover, for every $f \in \mathcal{H}_h^0$,

$$\langle f, f\rangle_{B,h} \overset{(a)}{=} \int_{-A}^{A} f'(t) f'(t)\,\mathrm{d}t \overset{(b)}{=} \int_{-A}^{A} |f'(t)|^2\,\mathrm{d}t \overset{(c)}{\geq} 0. \tag{29}$$

Here (a) is the definition of the energy form; (b) uses that the functions are real valued; and (c) follows because $|f'(t)|^2 \geq 0$ almost everywhere.

It remains to prove definiteness. Fix $f \in \mathcal{H}_h^0$ and let $\mathbf{v} \in \mathbb{R}^{G+1}$ be its unique nodal vector, so that

$$f \overset{(a)}{=} \mathcal{R}(\mathbf{v}), \qquad v_{i_0} \overset{(b)}{=} 0. \tag{30}$$

Here (a) follows from $f \in \mathcal{H}_h$ and Definition 1, and (b) follows from $f \in \mathcal{H}_h^0$ and (22). Assume that

$$\langle f, f\rangle_{B,h} = 0. \tag{31}$$

By the definition of the Brownian seminorm and Proposition 1,

$$0 \overset{(a)}{=} \langle f, f\rangle_{B,h} \overset{(b)}{=} \|f\|_{B,h}^2 \overset{(c)}{=} \mathbf{v}^\top \mathbf{K} \mathbf{v} \overset{(d)}{=} \frac{1}{h} \|\mathbf{D}\mathbf{v}\|_2^2. \tag{32}$$

Here (a) is the assumption (31); (b) is the definition of the Brownian seminorm; and (c)–(d) are the stiffness and first-difference representations in Proposition 1. The mesh size is strictly positive:

$$h \overset{(a)}{=} \frac{2A}{G} \overset{(b)}{>} 0. \tag{33}$$

Here (a) is the mesh-size definition, and (b) follows because $A > 0$ and the strictly increasing grid contains $G \geq 1$ subintervals. Multiplying the first and last quantities in (32) by h and using (33), we obtain

$$\|\mathbf{D}\mathbf{v}\|_2^2 \overset{(a)}{=} h \left(\frac{1}{h} \|\mathbf{D}\mathbf{v}\|_2^2\right) \overset{(b)}{=} h \cdot 0 \overset{(c)}{=} 0. \tag{34}$$

Here (a) uses $h(1/h) = 1$; (b) uses (32); and (c) uses $h \cdot 0 = 0$. For each $i = 0, \dots, G-1$,

$$0 \overset{(a)}{\leq} ((\mathbf{D}\mathbf{v})_i)^2 \overset{(b)}{\leq} \sum_{r=0}^{G-1} ((\mathbf{D}\mathbf{v})_r)^2 \overset{(c)}{=} \|\mathbf{D}\mathbf{v}\|_2^2 \overset{(d)}{=} 0. \tag{35}$$

Here (a) uses nonnegativity of a square; (b) uses that the sum on the right contains the nonnegative term on the left; (c) is the componentwise definition of the Euclidean norm; and (d) uses (34). Therefore every component $(\mathbf{D}\mathbf{v})_i$ is zero and hence

$$\mathbf{D}\mathbf{v} \overset{(a)}{=} \mathbf{0}. \tag{36}$$

Here (a) follows from (35) for all component indices. Using the factorization in Proposition 1,

$$\mathbf{K}\mathbf{v} \overset{(a)}{=} \left(\frac{1}{h}\mathbf{D}^\top\mathbf{D}\right)\mathbf{v} \overset{(b)}{=} \frac{1}{h}\mathbf{D}^\top(\mathbf{D}\mathbf{v}) \overset{(c)}{=} \frac{1}{h}\mathbf{D}^\top\mathbf{0} \overset{(d)}{=} \mathbf{0}. \tag{37}$$

Here (a) uses $\mathbf{K} = \frac{1}{h}\mathbf{D}^\top\mathbf{D}$; (b) uses associativity of matrix multiplication; (c) substitutes (36); and (d) uses that every matrix maps the zero vector to the zero vector. Thus $\mathbf{v} \in \ker(\mathbf{K})$. Lemma D2 now yields

$$\mathbf{v} \overset{(a)}{\in} \ker(\mathbf{K}) \overset{(b)}{=} \operatorname{span}(\mathbf{1}) \overset{(c)}{\Longrightarrow} \mathbf{v} = c\mathbf{1} \quad \text{for some } c \in \mathbb{R}.$$

Here (a) follows from (37); (b) is Lemma D2; and (c) is the definition of the one-dimensional span of $\mathbf{1} = (1, \dots, 1)^\top \in \mathbb{R}^{G+1}$. The anchoring constraint in (30) then gives

$$0 \overset{(a)}{=} v_{i_0} \overset{(b)}{=} (c\mathbf{1})_{i_0} \overset{(c)}{=} c.$$

Here (a) is the anchoring condition; (b) substitutes $\mathbf{v} = c\mathbf{1}$; and (c) uses $\mathbf{1}_{i_0} = 1$. Therefore $c = 0$ and $\mathbf{v} = \mathbf{0}$. Substituting this vector into the reconstruction formula gives

$$f \overset{(a)}{=} \mathcal{R}(\mathbf{v}) \overset{(b)}{=} \mathcal{R}(\mathbf{0}) \overset{(c)}{=} \sum_{j=0}^{G} 0\phi_j \overset{(d)}{=} 0.$$

Here (a) is (30); (b) uses $\mathbf{v} = \mathbf{0}$; (c) is the definition of the reconstruction operator; and (d) simplifies the zero linear combination. Together with nonnegativity in (29), this proves that the Brownian energy form is positive definite on $\mathcal{H}_h^0$. Hence $\langle\cdot,\cdot\rangle_{B,h}$ is an inner product on $\mathcal{H}_h^0$, and $\|\cdot\|_{B,h}$ is its induced norm.

Completeness and boundedness of evaluation are automatic in finite dimensions, but the two quantitative bounds behind them are used later and are recorded here.

Let $f = \mathcal{R}(\mathbf{v}) \in \mathcal{H}_h^0$, so $v_{i_0} = 0$. Telescoping from the anchor, $v_j = -\sum_{r=j}^{i_0-1}(v_{r+1} - v_r)$ for $j < i_0$ and $v_j = \sum_{r=i_0}^{j-1}(v_{r+1} - v_r)$ for $j > i_0$; in either case Cauchy–Schwarz over at most G terms gives $v_j^2 \leq G\sum_{i=0}^{G-1}(v_{i+1} - v_i)^2 = Gh\|f\|_{B,h}^2$ by Proposition 1. Summing over the $G+1$ nodes,

$$\|\mathbf{v}\|_2 \leq \sqrt{G(G+1)h}\,\|f\|_{B,h}, \qquad \text{and conversely} \qquad \|f\|_{B,h} \leq \frac{2}{\sqrt{h}}\|\mathbf{v}\|_2, \tag{38}$$

the second bound from $(v_{i+1} - v_i)^2 \leq 2(v_{i+1}^2 + v_i^2)$. Thus $\|\cdot\|_{B,h}$ and the Euclidean norm on the coefficient vector are equivalent, so $(\mathcal{H}_h^0, \langle\cdot,\cdot\rangle_{B,h})$ is a finite-dimensional inner-product space and hence complete: a $\|\cdot\|_{B,h}$-Cauchy sequence has Cauchy coefficient vectors by (38), and the limit vector reconstructs the limit function, which lies in $\mathcal{H}_h^0$ because $v_{i_0} = 0$ is preserved.

For $t \in I$ the evaluation functional $\delta_t(f) = f(t)$ is linear, and writing $f(t) = f(t) - f(0) = \int_0^t f'$ and applying Cauchy–Schwarz on an interval of length at most A,

$$|\delta_t(f)| = \left|\int_0^t f'\right| \leq \sqrt{|t|}\left(\int_I |f'|^2\right)^{1/2} \leq \sqrt{A}\,\|f\|_{B,h}.$$

Hence $\|\delta_t\| \leq \sqrt{A}$, uniformly in h: the bound does not degrade under mesh refinement, which is the quantitative form of the isometric embedding of Corollary 2.

Since $\mathcal{H}_h^0$ is a real Hilbert space and δ_t is a bounded linear functional, the Riesz representation theorem provides a unique function $k_t \in \mathcal{H}_h^0$ such that

$$\delta_t(f) \overset{(a)}{=} \langle f, k_t \rangle_{B,h} \qquad \text{for every } f \in \mathcal{H}_h^0. \tag{39}$$

Here (a) is the Riesz representation theorem applied to the bounded linear functional δ_t. Define

$$k_h^0 : I \times I \longrightarrow \mathbb{R}, \qquad k_h^0(s,t) := k_t(s). \tag{40}$$

For every $f \in \mathcal{H}_h^0$ and $t \in I$,

$$f(t) \overset{(a)}{=} \delta_t(f) \overset{(b)}{=} \langle f, k_t \rangle_{B,h} \overset{(c)}{=} \left\langle f, k_h^0(\cdot, t) \right\rangle_{B,h}.$$

Here (a) is the definition of point evaluation; (b) is (39); and (c) uses $k_h^0(\cdot, t) = k_t$ from (40). Thus k_h^0 reproduces every point evaluation. For completeness, the kernel is symmetric because, for arbitrary $s, t \in I$,

$$k_h^0(s,t) \overset{(a)}{=} k_t(s) \overset{(b)}{=} \langle k_t, k_s \rangle_{B,h} \overset{(c)}{=} \langle k_s, k_t \rangle_{B,h} \overset{(d)}{=} k_s(t) \overset{(e)}{=} k_h^0(t,s).$$

Here (a) and (e) use the kernel definition; (b) applies (39) with the function $f = k_t$ and evaluation point s; (c) uses symmetry of the inner product; and (d) applies (39) with the function $f = k_s$ and evaluation point t. Moreover, for arbitrary $N \geq 1$, points $t_1, \ldots, t_N \in I$, and coefficients $a_1, \ldots, a_N \in \mathbb{R}$,

$$\sum_{r=1}^{N}\sum_{s=1}^{N} a_r a_s k_h^0(t_r, t_s) \overset{(a)}{=} \sum_{r=1}^{N}\sum_{s=1}^{N} a_r a_s \langle k_{t_s}, k_{t_r} \rangle_{B,h} \overset{(b)}{=} \left\langle \sum_{s=1}^{N} a_s k_{t_s}, \sum_{r=1}^{N} a_r k_{t_r} \right\rangle_{B,h}$$
$$\overset{(c)}{=} \left\| \sum_{r=1}^{N} a_r k_{t_r} \right\|_{B,h}^2 \overset{(d)}{\geq} 0.$$

Here (a) uses $k_h^0(t_r, t_s) = k_{t_s}(t_r) = \langle k_{t_s}, k_{t_r} \rangle_{B,h}$, which follows from (40) and (39); (b) uses bilinearity of the inner product and the finiteness of the two sums; (c) observes that the two displayed sums represent the same function after renaming the dummy index; and (d) uses nonnegativity of the squared norm. Therefore k_h^0 is a positive-semidefinite reproducing kernel for $\mathcal{H}_h^0$. Combining this reproducing property with (26) and the completeness proved above shows that

$$\left(\mathcal{H}_h^0, \langle \cdot, \cdot \rangle_{B,h}\right)$$

is a finite-dimensional reproducing kernel Hilbert space, which completes the proof. Items i–ii are established in Section C.2.

□

B.4 Proof of Theorem 2

The increment-energy identity used in part iii is Lemma D4, and the positivity of $\mathbf{\Lambda}$ used in the spectral part is Lemma 3.1 with Lemma D5. The derivation below is self-contained, but those lemmas may be substituted for the corresponding steps.

Let $f \in \mathcal{H}_h^0$ be arbitrary. By Definition 3, G is even. Write $G = 2m$ and $i_0 = m$. Since $f \in \mathcal{H}_h^0 \subseteq \mathcal{H}_h$, Definition 1 yields a unique vector $\mathbf{v} = (v_0, \ldots, v_G)^\top \in \mathbb{R}^{G+1}$ such that

$$f \overset{(a)}{=} \mathcal{R}(\mathbf{v}) \overset{(b)}{=} \sum_{r=0}^{G} v_r \phi_r. \tag{41}$$

Here (a) is the unique finite-element representation of f, and (b) is the definition of the reconstruction operator $\mathcal{R}$. For every grid index $j = 0, \ldots, G$, the nodal property of the basis gives

$$f(t_j) \overset{(a)}{=} \sum_{r=0}^{G} v_r \phi_r(t_j) \overset{(b)}{=} \sum_{r=0}^{G} v_r \delta_{rj} \overset{(c)}{=} v_j. \tag{42}$$

Here (a) evaluates (41) at t_j, (b) uses $\phi_r(t_j) = \delta_{rj}$, and (c) uses the defining property of the Kronecker delta. Since $t_{i_0} = 0$ and $f(0) = 0$, it follows that

$$v_{i_0} \overset{(a)}{=} f(t_{i_0}) \overset{(b)}{=} f(0) \overset{(c)}{=} 0. \tag{43}$$

Here (a) applies (42), (b) uses $t_{i_0} = 0$, and (c) uses $f \in \mathcal{H}_h^0$.

Let $\widetilde{\mathbf{v}} \in \mathbb{R}^G$ be the reduced anchored vector ordered as in (9). By the definition of the anchoring reconstruction matrix in (10),

$$\mathbf{v} \overset{(a)}{=} \mathbf{R}\widetilde{\mathbf{v}}. \tag{44}$$

Equality (a) inserts the zero component at index i_0 and places every remaining reduced coordinate at its corresponding nodal index. Define

$$\mathbf{D}_0 := \mathbf{D}\mathbf{R} \in \mathbb{R}^{G\times G}. \tag{45}$$

Then the increment coordinates satisfy

$$\mathbf{w} \overset{(a)}{=} \mathbf{D}\mathbf{v} \overset{(b)}{=} \mathbf{D}\mathbf{R}\widetilde{\mathbf{v}} \overset{(c)}{=} \mathbf{D}_0\widetilde{\mathbf{v}}.$$

Here (a) is Definition 6, (b) substitutes (44), and (c) uses (45). The spectral coordinates satisfy

$$\mathbf{c} \overset{(a)}{=} \mathbf{Q}^\top\widetilde{\mathbf{v}}.$$

Equality (a) is Definition 7.

We first establish ii, because the inverse coordinate maps will then make the assertion in i explicit. For arbitrary $\mathbf{x}, \mathbf{y} \in \mathbb{R}^G$ and $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned}\mathbf{D}_0\left(\alpha\mathbf{x} + \beta\mathbf{y}\right) &\overset{(a)}{=} \mathbf{D}\mathbf{R}\left(\alpha\mathbf{x} + \beta\mathbf{y}\right)\\ &\overset{(b)}{=} \mathbf{D}\left(\alpha\mathbf{R}\mathbf{x} + \beta\mathbf{R}\mathbf{y}\right)\\ &\overset{(c)}{=} \alpha\mathbf{D}\mathbf{R}\mathbf{x} + \beta\mathbf{D}\mathbf{R}\mathbf{y} \overset{(d)}{=} \alpha\mathbf{D}_0\mathbf{x} + \beta\mathbf{D}_0\mathbf{y}.\end{aligned}$$

Here (a) substitutes $\mathbf{D}_0 = \mathbf{D}\mathbf{R}$, (b) uses the linearity of multiplication by $\mathbf{R}$, (c) uses the linearity of multiplication by $\mathbf{D}$, and (d) again uses the definition of $\mathbf{D}_0$. Thus $\mathbf{x} \mapsto \mathbf{D}_0\mathbf{x}$ is linear.

We next prove injectivity. Let $\mathbf{w} \in \mathbb{R}^G$ satisfy

$$\mathbf{D}_0\mathbf{w} = \mathbf{0}. \tag{46}$$

Set $\mathbf{z} := \mathbf{R}\mathbf{w} \in \mathbb{R}^{G+1}$. Then

$$\mathbf{D}\mathbf{z} \overset{(a)}{=} \mathbf{D}\mathbf{R}\mathbf{w} \overset{(b)}{=} \mathbf{D}_0\mathbf{w} \overset{(c)}{=} \mathbf{0}. \tag{47}$$

Here (a) substitutes $\mathbf{z} = \mathbf{R}\mathbf{w}$, (b) uses $\mathbf{D}_0 = \mathbf{D}\mathbf{R}$, and (c) applies (46). By the componentwise definition of $\mathbf{D}$, for every $i = 0, \dots, G-1$,

$$0 \overset{(a)}{=} (\mathbf{D}\mathbf{z})_i \overset{(b)}{=} z_{i+1} - z_i,$$

where (a) follows from (47), and (b) is the definition of the first-difference operator. Therefore $z_{i+1} = z_i$ for every $i = 0, \dots, G-1$. Repeated application of these equalities gives

$$z_j \overset{(a)}{=} z_0, \qquad j = 0, \dots, G. \tag{48}$$

Equality (a) follows by induction on j: it is immediate for $j = 0$, and if $z_j = z_0$, then $z_{j+1} = z_j = z_0$.

The row of $\mathbf{R}$ corresponding to the anchor index i_0 is zero because none of the columns of $\mathbf{R}$ equals the omitted canonical vector $\mathbf{e}_{i_0}$. Hence

$$z_{i_0} \overset{(a)}{=} (\mathbf{R}\mathbf{w})_{i_0} \overset{(b)}{=} 0. \tag{49}$$

Here (a) uses $\mathbf{z} = \mathbf{R}\mathbf{w}$, and (b) uses the zero anchor row of $\mathbf{R}$. Combining (48) and (49) gives, for every j,

$$z_j \overset{(a)}{=} z_{i_0} \overset{(b)}{=} 0,$$

where (a) uses the constancy of $\mathbf{z}$, and (b) uses the anchored value. Thus $\mathbf{z} = \mathbf{0}$.

The columns of $\mathbf{R}$ are pairwise distinct canonical basis vectors of $\mathbb{R}^{G+1}$. Consequently,

$$\mathbf{R}^\top \mathbf{R} \overset{(a)}{=} \mathbf{I}_G. \tag{50}$$

Indeed, the (r, s) entry of $\mathbf{R}^\top \mathbf{R}$ is the Euclidean inner product of columns r and s of $\mathbf{R}$, which is 1 when $r = s$ and 0 otherwise. It follows that

$$\mathbf{w} \overset{(a)}{=} \mathbf{I}_G \mathbf{w} \overset{(b)}{=} \mathbf{R}^\top \mathbf{R}\mathbf{w} \overset{(c)}{=} \mathbf{R}^\top \mathbf{z} \overset{(d)}{=} \mathbf{0}.$$

Here (a) uses the identity matrix, (b) applies (50), (c) uses $\mathbf{z} = \mathbf{R}\mathbf{w}$, and (d) uses $\mathbf{z} = \mathbf{0}$. Therefore $\ker(\mathbf{D}_0) = \{\mathbf{0}\}$, so the increment map is injective.

We now prove surjectivity directly. Let $\mathbf{d} = (d_0, \ldots, d_{G-1})^\top \in \mathbb{R}^G$ be arbitrary, and define $\mathbf{u}(\mathbf{d}) = (u_0, \ldots, u_G)^\top \in \mathbb{R}^{G+1}$ by

$$u_j := \begin{cases} -\sum_{r=j}^{i_0-1} d_r, & 0 \le j \le i_0, \\ \sum_{r=i_0}^{j-1} d_r, & i_0 \le j \le G. \end{cases} \tag{51}$$

The two formulas agree at $j = i_0$ because both sums are empty, and every empty sum is understood to equal zero. In particular,

$$u_{i_0} \overset{(a)}{=} 0. \tag{52}$$

Equality (a) follows from the empty-sum convention. For $i = 0, \ldots, i_0 - 1$,

$$u_{i+1} - u_i \overset{(a)}{=} -\sum_{r=i+1}^{i_0-1} d_r + \sum_{r=i}^{i_0-1} d_r \overset{(b)}{=} d_i. \tag{53}$$

Here (a) applies the first branch of (51), including its empty-sum value when $i = i_0 - 1$, and (b) cancels the common terms $d_{i+1}, \ldots, d_{i_0-1}$, leaving only d_i. For $i = i_0, \ldots, G - 1$,

$$u_{i+1} - u_i \overset{(a)}{=} \sum_{r=i_0}^{i} d_r - \sum_{r=i_0}^{i-1} d_r \overset{(b)}{=} d_i. \tag{54}$$

Here (a) applies the second branch of (51), including the empty second sum when $i = i_0$, and (b) cancels the common terms $d_{i_0}, \ldots, d_{i-1}$, leaving only d_i. Hence, for every $i = 0, \ldots, G - 1$,

$$(\mathbf{D}\mathbf{u}(\mathbf{d}))_i \overset{(a)}{=} u_{i+1} - u_i \overset{(b)}{=} d_i,$$

where (a) is the definition of $\mathbf{D}$, while (b) applies (53) when $i < i_0$ and (54) when $i \ge i_0$. Therefore

$$\mathbf{D}\mathbf{u}(\mathbf{d}) \overset{(a)}{=} \mathbf{d}. \tag{55}$$

Equality (a) follows because the two vectors have equal components at every index.

Because the columns of $\mathbf{R}$ contain every canonical basis vector except $\mathbf{e}_{i_0}$, one also has

$$\mathbf{R}\mathbf{R}^\top \overset{(a)}{=} \sum_{j \ne i_0} \mathbf{e}_j \mathbf{e}_j^\top \overset{(b)}{=} \mathbf{I}_{G+1} - \mathbf{e}_{i_0} \mathbf{e}_{i_0}^\top. \tag{56}$$

Here (a) expands the product as the sum of the outer products of the columns of $\mathbf{R}$, and (b) removes the single omitted canonical projector from the canonical decomposition of the identity. Define

$$\mathbf{x}(\mathbf{d}) := \mathbf{R}^\top \mathbf{u}(\mathbf{d}) \in \mathbb{R}^G. \tag{57}$$

Then

$$\begin{aligned}\mathbf{R}\mathbf{x}(\mathbf{d}) &\overset{(a)}{=} \mathbf{R}\mathbf{R}^\top \mathbf{u}(\mathbf{d}) \\ &\overset{(b)}{=} \left(\mathbf{I}_{G+1} - \mathbf{e}_{i_0}\mathbf{e}_{i_0}^\top\right)\mathbf{u}(\mathbf{d}) \\ &\overset{(c)}{=} \mathbf{u}(\mathbf{d}) - \mathbf{e}_{i_0} u_{i_0} \overset{(d)}{=} \mathbf{u}(\mathbf{d}).\end{aligned} \tag{58}$$

Here (a) substitutes (57), (b) uses (56), (c) evaluates $\mathbf{e}_{i_0}^\top \mathbf{u}(\mathbf{d}) = u_{i_0}$, and (d) uses (52). Consequently,

$$\mathbf{D}_0\mathbf{x}(\mathbf{d}) \overset{(a)}{=} \mathbf{D}\mathbf{R}\mathbf{x}(\mathbf{d}) \overset{(b)}{=} \mathbf{D}\mathbf{u}(\mathbf{d}) \overset{(c)}{=} \mathbf{d}. \tag{59}$$

Here (a) uses $\mathbf{D}_0 = \mathbf{D}\mathbf{R}$, (b) applies (58), and (c) applies (55). Since $\mathbf{d} \in \mathbb{R}^G$ was arbitrary, $\mathbf{D}_0$ is surjective. Together with injectivity, this proves that $\mathbf{D}_0$ is invertible. Moreover, the preceding construction gives the explicit inverse

$$\mathbf{D}_0^{-1}\mathbf{d} \overset{(a)}{=} \mathbf{R}^\top \mathbf{u}(\mathbf{d}),$$

where (a) follows from uniqueness of the preimage under the injective map $\mathbf{D}_0$ and (59).

We next treat the spectral map. For arbitrary $\mathbf{x}, \mathbf{y} \in \mathbb{R}^G$ and $\alpha, \beta \in \mathbb{R}$,

$$\mathbf{Q}^\top(\alpha\mathbf{x} + \beta\mathbf{y}) \overset{(a)}{=} \alpha\mathbf{Q}^\top\mathbf{x} + \beta\mathbf{Q}^\top\mathbf{y}.$$

Equality (a) is the distributivity and homogeneity of matrix multiplication, so the spectral map is linear. Since $\mathbf{Q}$ is orthogonal,

$$\mathbf{Q}^\top\mathbf{Q} \overset{(a)}{=} \mathbf{I}_G, \qquad \mathbf{Q}\mathbf{Q}^\top \overset{(b)}{=} \mathbf{I}_G. \tag{60}$$

Equalities (a) and (b) are the two defining inverse identities for an orthogonal square matrix. Thus, for every $\mathbf{x}, \mathbf{a} \in \mathbb{R}^G$,

$$\begin{aligned}\mathbf{Q}\left(\mathbf{Q}^\top\mathbf{x}\right) &\overset{(a)}{=} \left(\mathbf{Q}\mathbf{Q}^\top\right)\mathbf{x} \overset{(b)}{=} \mathbf{I}_G\mathbf{x} \overset{(c)}{=} \mathbf{x}, \\ \mathbf{Q}^\top\left(\mathbf{Q}\mathbf{a}\right) &\overset{(a)}{=} \left(\mathbf{Q}^\top\mathbf{Q}\right)\mathbf{a} \overset{(b)}{=} \mathbf{I}_G\mathbf{a} \overset{(c)}{=} \mathbf{a}.\end{aligned}$$

In each line, (a) uses associativity, (b) applies (60), and (c) uses the identity matrix. Therefore $\mathbf{x} \mapsto \mathbf{Q}^\top\mathbf{x}$ is invertible, with inverse $\mathbf{a} \mapsto \mathbf{Q}\mathbf{a}$.

The complete coordinate conversion formulas are now

$$\begin{aligned}\mathbf{w} &\overset{(a)}{=} \mathbf{D}_0\widetilde{\mathbf{v}}, & \mathbf{c} &\overset{(b)}{=} \mathbf{Q}^\top\widetilde{\mathbf{v}}, \\ \widetilde{\mathbf{v}} &\overset{(c)}{=} \mathbf{D}_0^{-1}\mathbf{w}, & \widetilde{\mathbf{v}} &\overset{(d)}{=} \mathbf{Q}\mathbf{c}, \\ \mathbf{c} &\overset{(e)}{=} \mathbf{Q}^\top\mathbf{D}_0^{-1}\mathbf{w}, & \mathbf{w} &\overset{(f)}{=} \mathbf{D}_0\mathbf{Q}\mathbf{c}.\end{aligned}$$

Here (a) and (b) are the coordinate definitions, (c) uses the inverse of $\mathbf{D}_0$, (d) uses the inverse of $\mathbf{Q}^\top$, (e) substitutes (c) into (b), and (f) substitutes (d) into (a). Hence each coordinate system uniquely determines the other two, proving ii.

We now prove i. Define the three reconstruction maps by

$$\Phi_{\mathrm{nod}}(\mathbf{x}) := \mathcal{R}(\mathbf{R}\mathbf{x}), \tag{61}$$

$$\Phi_{\mathrm{inc}}(\mathbf{d}) := \mathcal{R}\left(\mathbf{R}\mathbf{D}_0^{-1}\mathbf{d}\right), \tag{62}$$

$$\Phi_{\mathrm{spec}}(\mathbf{a}) := \mathcal{R}\left(\mathbf{R}\mathbf{Q}\mathbf{a}\right). \tag{63}$$

Each map takes values in $\mathcal{H}_h^0$. Indeed, for every $\mathbf{x} \in \mathbb{R}^G$,

$$\begin{aligned}\Phi_{\mathrm{nod}}(\mathbf{x})(0) &\overset{(a)}{=} \sum_{j=0}^{G}(\mathbf{R}\mathbf{x})_j\phi_j(0) \overset{(b)}{=} \sum_{j=0}^{G}(\mathbf{R}\mathbf{x})_j\delta_{ji_0} \\ &\overset{(c)}{=} (\mathbf{R}\mathbf{x})_{i_0} \overset{(d)}{=} 0.\end{aligned} \tag{64}$$

Here (a) expands $\mathcal{R}(\mathbf{Rx})$, (b) uses $0 = t_{i_0}$ and $\phi_j(t_{i_0}) = \delta_{ji_0}$, (c) applies the Kronecker delta, and (d) uses the zero anchor row of $\mathbf{R}$. Thus $\Phi_{\mathrm{nod}}(\mathbf{x}) \in \mathcal{H}_h^0$. The increment and spectral maps are compositions of Φ_{nod} with vectors in $\mathbb{R}^G$, so they also take values in $\mathcal{H}_h^0$.

For the coordinates associated with the fixed function f,

$$\Phi_{\mathrm{nod}}(\widetilde{\mathbf{v}}) \overset{(a)}{=} \mathcal{R}(\mathbf{R}\widetilde{\mathbf{v}}) \overset{(b)}{=} \mathcal{R}(\mathbf{v}) \overset{(c)}{=} f. \tag{65}$$

Here (a) uses (61), (b) uses (44), and (c) uses (41). Likewise,

$$\Phi_{\mathrm{inc}}(\mathbf{w}) \overset{(a)}{=} \mathcal{R}\left(\mathbf{R}\mathbf{D}_0^{-1}\mathbf{w}\right) \overset{(b)}{=} \mathcal{R}\left(\mathbf{R}\mathbf{D}_0^{-1}\mathbf{D}_0\widetilde{\mathbf{v}}\right) \overset{(c)}{=} \mathcal{R}(\mathbf{R}\mathbf{I}_G\widetilde{\mathbf{v}}) \overset{(d)}{=} \mathcal{R}(\mathbf{R}\widetilde{\mathbf{v}}) \overset{(e)}{=} f.$$

Here (a) uses (62), (b) substitutes $\mathbf{w} = \mathbf{D}_0\widetilde{\mathbf{v}}$, (c) uses $\mathbf{D}_0^{-1}\mathbf{D}_0 = \mathbf{I}_G$, (d) uses the identity matrix, and (e) applies (65). Finally,

$$\Phi_{\mathrm{spec}}(\mathbf{c}) \overset{(a)}{=} \mathcal{R}(\mathbf{R}\mathbf{Q}\mathbf{c}) \overset{(b)}{=} \mathcal{R}\left(\mathbf{R}\mathbf{Q}\mathbf{Q}^\top\widetilde{\mathbf{v}}\right) \overset{(c)}{=} \mathcal{R}(\mathbf{R}\mathbf{I}_G\widetilde{\mathbf{v}}) \overset{(d)}{=} \mathcal{R}(\mathbf{R}\widetilde{\mathbf{v}}) \overset{(e)}{=} f.$$

Here (a) uses (63), (b) substitutes $\mathbf{c} = \mathbf{Q}^\top\widetilde{\mathbf{v}}$, (c) uses $\mathbf{Q}\mathbf{Q}^\top = \mathbf{I}_G$, (d) uses the identity matrix, and (e) applies (65). Therefore the three coordinate representations reconstruct exactly the same function f, proving i.

We next prove iii. By Proposition 1,

$$\begin{aligned}\|f\|_{B,h}^2 &\overset{(a)}{=} \mathbf{v}^\top\mathbf{K}\mathbf{v} \overset{(b)}{=} \frac{1}{h}(\mathbf{D}\mathbf{v})^\top(\mathbf{D}\mathbf{v}) \\ &\overset{(c)}{=} \frac{1}{h}\mathbf{w}^\top\mathbf{w} \overset{(d)}{=} \frac{1}{h}\|\mathbf{w}\|_2^2.\end{aligned} \tag{66}$$

Here (a) is the stiffness representation of the Brownian energy, (b) uses $\mathbf{K} = h^{-1}\mathbf{D}^\top\mathbf{D}$ and associativity, (c) uses $\mathbf{w} = \mathbf{D}\mathbf{v}$, and (d) is the definition of the Euclidean norm.

For the spectral representation, first

$$\|f\|_{B,h}^2 \overset{(a)}{=} (\mathbf{R}\widetilde{\mathbf{v}})^\top\mathbf{K}(\mathbf{R}\widetilde{\mathbf{v}}) \overset{(b)}{=} \widetilde{\mathbf{v}}^\top\mathbf{R}^\top\mathbf{K}\mathbf{R}\widetilde{\mathbf{v}} \overset{(c)}{=} \widetilde{\mathbf{v}}^\top\mathbf{K}_0\widetilde{\mathbf{v}}. \tag{67}$$

Here (a) substitutes $\mathbf{v} = \mathbf{R}\widetilde{\mathbf{v}}$ into the first equality of (66), (b) uses $(\mathbf{R}\widetilde{\mathbf{v}})^\top = \widetilde{\mathbf{v}}^\top\mathbf{R}^\top$ and associativity, and (c) uses $\mathbf{K}_0 = \mathbf{R}^\top\mathbf{K}\mathbf{R}$. By Definition 7,

$$\widetilde{\mathbf{v}} \overset{(a)}{=} \mathbf{Q}\mathbf{c}, \qquad \mathbf{K}_0 \overset{(b)}{=} \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top. \tag{68}$$

Here (a) is the inverse spectral transformation proved above, and (b) is the eigendecomposition of the anchored stiffness matrix, whose existence with strictly positive spectrum is established in Lemma 3.1 and Lemma D5, rather than assumed in Definition 7. Consequently,

$$\|f\|_{B,h}^2 \overset{(a)}{=} (\mathbf{Q}\mathbf{c})^\top\left(\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top\right)(\mathbf{Q}\mathbf{c}) \overset{(b)}{=} \mathbf{c}^\top\mathbf{Q}^\top\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top\mathbf{Q}\mathbf{c} \overset{(c)}{=} \mathbf{c}^\top\mathbf{I}_G\mathbf{\Lambda}\mathbf{I}_G\mathbf{c} \overset{(d)}{=} \mathbf{c}^\top\mathbf{\Lambda}\mathbf{c}. \tag{69}$$

Here (a) substitutes both identities in (68) into (67), (b) uses $(\mathbf{Q}\mathbf{c})^\top = \mathbf{c}^\top\mathbf{Q}^\top$ and associativity, (c) uses $\mathbf{Q}^\top\mathbf{Q} = \mathbf{I}_G$, and (d) uses the identity matrix. Combining (66) and (69) gives

$$\|f\|_{B,h}^2 \overset{(a)}{=} \mathbf{v}^\top\mathbf{K}\mathbf{v} \overset{(b)}{=} \frac{1}{h}\|\mathbf{w}\|_2^2 \overset{(c)}{=} \mathbf{c}^\top\mathbf{\Lambda}\mathbf{c}.$$

Here (a) and (b) are the nodal and increment equalities in (66), while (c) uses (69). This proves iii.

We now establish iv. Define the ranges of the three reconstruction maps by

$$\mathcal{F}_{\mathrm{nod}} := \Phi_{\mathrm{nod}}(\mathbb{R}^G), \qquad \mathcal{F}_{\mathrm{inc}} := \Phi_{\mathrm{inc}}(\mathbb{R}^G), \qquad \mathcal{F}_{\mathrm{spec}} := \Phi_{\mathrm{spec}}(\mathbb{R}^G).$$

Equation (64) shows that $\mathcal{F}_{\mathrm{nod}} \subseteq \mathcal{H}_h^0$. Conversely, let $g \in \mathcal{H}_h^0$ be arbitrary and let $\mathbf{a} \in \mathbb{R}^{G+1}$ be its unique nodal vector. As in (43), $a_{i_0} = 0$. Set $\mathbf{x} := \mathbf{R}^\top\mathbf{a} \in \mathbb{R}^G$. Then

$$\begin{aligned}\mathbf{R}\mathbf{x} &\overset{(a)}{=} \mathbf{R}\mathbf{R}^\top\mathbf{a} \overset{(b)}{=} \left(\mathbf{I}_{G+1} - \mathbf{e}_{i_0}\mathbf{e}_{i_0}^\top\right)\mathbf{a} \\ &\overset{(c)}{=} \mathbf{a} - \mathbf{e}_{i_0}a_{i_0} \overset{(d)}{=} \mathbf{a}.\end{aligned} \tag{70}$$

Here (a) uses $\mathbf{x} = \mathbf{R}^\top \mathbf{a}$, (b) applies (56), (c) evaluates the rank-one projector, and (d) uses $a_{i_0} = 0$. Therefore

$$g \overset{(a)}{=} \mathcal{R}(\mathbf{a}) \overset{(b)}{=} \mathcal{R}(\mathbf{R}\mathbf{x}) \overset{(c)}{=} \Phi_{\mathrm{nod}}(\mathbf{x}),$$

where (a) is the nodal representation of g, (b) applies (70), and (c) uses (61). Hence $g \in \mathcal{F}_{\mathrm{nod}}$, so $\mathcal{H}_h^0 \subseteq \mathcal{F}_{\mathrm{nod}}$. We have proved

$$\mathcal{F}_{\mathrm{nod}} \overset{(a)}{=} \mathcal{H}_h^0. \tag{71}$$

Equality (a) combines the two set inclusions.

Using the invertibility of $\mathbf{D}_0$,

$$\mathcal{F}_{\mathrm{inc}} \overset{(a)}{=} \left\{\Phi_{\mathrm{nod}}\left(\mathbf{D}_0^{-1}\mathbf{d}\right) : \mathbf{d} \in \mathbb{R}^G\right\} \overset{(b)}{=} \left\{\Phi_{\mathrm{nod}}(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^G\right\} \overset{(c)}{=} \mathcal{F}_{\mathrm{nod}}. \tag{72}$$

Here (a) uses (62), (b) uses the bijective change of variable $\mathbf{x} = \mathbf{D}_0^{-1}\mathbf{d}$, and (c) uses the definition of $\mathcal{F}_{\mathrm{nod}}$. Similarly, using the orthogonality and hence surjectivity of $\mathbf{Q}$,

$$\mathcal{F}_{\mathrm{spec}} \overset{(a)}{=} \left\{\Phi_{\mathrm{nod}}\left(\mathbf{Q}\mathbf{a}\right) : \mathbf{a} \in \mathbb{R}^G\right\} \overset{(b)}{=} \left\{\Phi_{\mathrm{nod}}(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^G\right\} \overset{(c)}{=} \mathcal{F}_{\mathrm{nod}}. \tag{73}$$

Here (a) uses (63), (b) uses the bijective change of variable $\mathbf{x} = \mathbf{Q}\mathbf{a}$, and (c) again uses the definition of $\mathcal{F}_{\mathrm{nod}}$. Combining (71)–(73) gives

$$\mathcal{F}_{\mathrm{nod}} \overset{(a)}{=} \mathcal{F}_{\mathrm{inc}} \overset{(b)}{=} \mathcal{F}_{\mathrm{spec}} \overset{(c)}{=} \mathcal{H}_h^0.$$

Here (a) is (72), (b) is (73), and (c) is (71).

For completeness, we also verify that the coordinate realizations pull back the same Brownian inner product, not merely the same norm. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^G$, define

$$\begin{aligned} g &:= \Phi_{\mathrm{nod}}(\mathbf{x}), & h &:= \Phi_{\mathrm{nod}}(\mathbf{y}), \\ \boldsymbol{\alpha} &:= \mathbf{D}_0\mathbf{x}, & \boldsymbol{\beta} &:= \mathbf{D}_0\mathbf{y}, \\ \mathbf{p} &:= \mathbf{Q}^\top\mathbf{x}, & \mathbf{q} &:= \mathbf{Q}^\top\mathbf{y}. \end{aligned}$$

The bilinear stiffness identity gives

$$\langle g, h\rangle_{B,h} \overset{(a)}{=} (\mathbf{R}\mathbf{x})^\top \mathbf{K} (\mathbf{R}\mathbf{y}) \overset{(b)}{=} \mathbf{x}^\top \mathbf{K}_0 \mathbf{y}. \tag{74}$$

Here (a) follows by expanding both reconstructed functions in the hat basis and using $K_{rs} = \langle \phi_r, \phi_s\rangle_{B,h}$, while (b) uses $\mathbf{K}_0 = \mathbf{R}^\top\mathbf{K}\mathbf{R}$. Moreover,

$$\mathbf{x}^\top\mathbf{K}_0\mathbf{y} \overset{(a)}{=} \mathbf{x}^\top\mathbf{R}^\top\mathbf{K}\mathbf{R}\mathbf{y} \overset{(b)}{=} \frac{1}{h}\mathbf{x}^\top\mathbf{R}^\top\mathbf{D}^\top\mathbf{D}\mathbf{R}\mathbf{y} \overset{(c)}{=} \frac{1}{h}(\mathbf{D}_0\mathbf{x})^\top(\mathbf{D}_0\mathbf{y}) \overset{(d)}{=} \frac{1}{h}\boldsymbol{\alpha}^\top\boldsymbol{\beta}. \tag{75}$$

Here (a) uses the definition of $\mathbf{K}_0$, (b) substitutes $\mathbf{K} = h^{-1}\mathbf{D}^\top\mathbf{D}$, (c) uses $\mathbf{D}_0 = \mathbf{D}\mathbf{R}$ and the transpose identity $(\mathbf{D}_0\mathbf{x})^\top = \mathbf{x}^\top\mathbf{D}_0^\top$, and (d) uses the definitions of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Since $\mathbf{x} = \mathbf{Q}\mathbf{p}$ and $\mathbf{y} = \mathbf{Q}\mathbf{q}$,

$$\begin{aligned} \mathbf{x}^\top\mathbf{K}_0\mathbf{y} &\overset{(a)}{=} (\mathbf{Q}\mathbf{p})^\top\left(\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\top\right)(\mathbf{Q}\mathbf{q}) \\ &\overset{(b)}{=} \mathbf{p}^\top\mathbf{Q}^\top\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\top\mathbf{Q}\mathbf{q} \overset{(c)}{=} \mathbf{p}^\top\boldsymbol{\Lambda}\mathbf{q}. \end{aligned} \tag{76}$$

Here (a) substitutes the inverse spectral transformations and the eigendecomposition of $\mathbf{K}_0$, (b) uses the transpose identity and associativity, and (c) applies $\mathbf{Q}^\top\mathbf{Q} = \mathbf{I}_G$. Thus

$$\langle g, h\rangle_{B,h} \overset{(a)}{=} \mathbf{x}^\top\mathbf{K}_0\mathbf{y} \overset{(b)}{=} \frac{1}{h}\boldsymbol{\alpha}^\top\boldsymbol{\beta} \overset{(c)}{=} \mathbf{p}^\top\boldsymbol{\Lambda}\mathbf{q}. \tag{77}$$

Here (a), (b), and (c) apply (74), (75), and (76), respectively. Therefore, when the nodal, increment, and spectral parameter spaces are respectively equipped with the pulled-back inner products displayed in (77), their reconstruction maps are linear isometric isomorphisms onto the common space $\mathcal{H}_h^0$ equipped with its Brownian inner product. By Theorem 1, this common inner-product space is a finite-dimensional RKHS. Hence all three realizations generate the same finite Brownian RKHS and the same hypothesis class, proving iv.

It remains to establish v. Let $\widetilde{\mathbf{v}}_1, \widetilde{\mathbf{v}}_2 \in \mathbb{R}^G$ be arbitrary, and define

$$\mathbf{w} := \widetilde{\mathbf{v}}_1 - \widetilde{\mathbf{v}}_2,$$
$$\mathbf{c}_r := \mathbf{Q}^\top \widetilde{\mathbf{v}}_r, \qquad \mathbf{w}_r := \mathbf{D}_0 \widetilde{\mathbf{v}}_r, \qquad r \in \{1, 2\}.$$

For the spectral coordinates,

$$\|\mathbf{c}_1 - \mathbf{c}_2\|_2^2 \overset{(a)}{=} \left\|\mathbf{Q}^\top (\widetilde{\mathbf{v}}_1 - \widetilde{\mathbf{v}}_2)\right\|_2^2 \overset{(b)}{=} \left(\mathbf{Q}^\top \mathbf{w}\right)^\top \left(\mathbf{Q}^\top \mathbf{w}\right)$$
$$\overset{(c)}{=} \mathbf{w}^\top \mathbf{Q}\mathbf{Q}^\top \mathbf{w} \overset{(d)}{=} \mathbf{w}^\top \mathbf{I}_G \mathbf{w} \overset{(e)}{=} \|\mathbf{w}\|_2^2 . \tag{78}$$

Here (a) uses linearity of $\mathbf{Q}^\top$ and the definitions of $\mathbf{c}_1$ and $\mathbf{c}_2$, (b) is the definition of the squared Euclidean norm, (c) uses $(\mathbf{Q}^\top \mathbf{w})^\top = \mathbf{w}^\top \mathbf{Q}$, (d) uses $\mathbf{Q}\mathbf{Q}^\top = \mathbf{I}_G$, and (e) again uses the definition of the Euclidean norm. Thus the nodal-to-spectral map is a Euclidean isometry.

For the increment map,

$$\mathbf{D}_0^\top \mathbf{D}_0 \overset{(a)}{=} (\mathbf{D}\mathbf{R})^\top (\mathbf{D}\mathbf{R}) \overset{(b)}{=} \mathbf{R}^\top \mathbf{D}^\top \mathbf{D}\mathbf{R} \overset{(c)}{=} \mathbf{R}^\top (h\mathbf{K})\, \mathbf{R} \overset{(d)}{=} h\mathbf{R}^\top \mathbf{K}\mathbf{R} \overset{(e)}{=} h\mathbf{K}_0. \tag{79}$$

Here (a) substitutes $\mathbf{D}_0 = \mathbf{D}\mathbf{R}$, (b) uses $(\mathbf{D}\mathbf{R})^\top = \mathbf{R}^\top \mathbf{D}^\top$ and associativity, (c) rearranges the stiffness identity $\mathbf{K} = h^{-1}\mathbf{D}^\top \mathbf{D}$ as $\mathbf{D}^\top \mathbf{D} = h\mathbf{K}$, (d) extracts the scalar h, and (e) uses $\mathbf{K}_0 = \mathbf{R}^\top \mathbf{K}\mathbf{R}$. Consequently,

$$\|\mathbf{w}_1 - \mathbf{w}_2\|_2^2 \overset{(a)}{=} \|\mathbf{D}_0 (\widetilde{\mathbf{v}}_1 - \widetilde{\mathbf{v}}_2)\|_2^2 \overset{(b)}{=} (\mathbf{D}_0 \mathbf{w})^\top (\mathbf{D}_0 \mathbf{w}) \overset{(c)}{=} \mathbf{w}^\top \mathbf{D}_0^\top \mathbf{D}_0 \mathbf{w} \overset{(d)}{=} h\mathbf{w}^\top \mathbf{K}_0 \mathbf{w}.$$

Here (a) uses linearity of $\mathbf{D}_0$ and the definitions of $\mathbf{w}_1$ and $\mathbf{w}_2$, (b) is the definition of the squared Euclidean norm, (c) uses the transpose identity and associativity, and (d) applies (79).

Finally, Lemma D3 gives

$$h\mathbf{K}_0 \overset{(a)}{=} \begin{pmatrix} \mathbf{T}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_m \end{pmatrix}. \tag{80}$$

Equality (a) is exactly the block decomposition after multiplication by h. If $m \geq 2$, then the entrywise definition of $\mathbf{T}_m$ gives

$$(\mathbf{T}_m)_{0,1} \overset{(a)}{=} -1 \overset{(b)}{\neq} 0 \overset{(c)}{=} (\mathbf{I}_m)_{0,1}.$$

Here (a) uses $|0-1| = 1$, (b) is the elementary inequality $-1 \neq 0$, and (c) uses the zero off-diagonal entries of the identity matrix. Therefore $\mathbf{T}_m \neq \mathbf{I}_m$ and, by (80),

$$\mathbf{D}_0^\top \mathbf{D}_0 \overset{(a)}{\neq} \mathbf{I}_G, \qquad m \geq 2.$$

Inequality (a) follows from (79), (80), and $\mathbf{T}_m \neq \mathbf{I}_m$. Hence the increment map is nonorthogonal for $G = 2m \geq 4$. In contrast, (78) shows that the spectral map is an orthogonal change of coordinates of the reduced nodal coordinates. This proves v and completes the proof.

□

B.5 Proof of Theorem 3

By Lemma D3,

$$\mathbf{K}_0 \overset{(a)}{=} \frac{1}{h} \begin{pmatrix} \mathbf{T}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_m \end{pmatrix}, \tag{81}$$

where $\mathbf{T}_m = (\tau_{rs})_{r,s=0}^{m-1}$ is defined by

$$\tau_{rs} := \begin{cases} 1, & r = s = 0, \\ 2, & 1 \leq r = s \leq m-1, \\ -1, & |r-s| = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{82}$$

Equality (a) is exactly the block decomposition proved in Lemma D3. In particular, $\mathbf{T}_1 = (1)$. It is therefore sufficient first to construct an orthonormal eigendecomposition of $\mathbf{T}_m$ and then to apply it independently to the two diagonal blocks in (81).

For every $k \in [m]$, set

$$\alpha_m := \frac{2}{\sqrt{2m+1}}, \qquad \theta_k := \frac{(2k-1)\,\pi}{2m+1}, \qquad \nu_k := 2 - 2\cos(\theta_k),$$

and define the extended cosine sequence

$$\widehat{q}_k(j) := \alpha_m \cos\left(\left(j + \frac{1}{2}\right)\theta_k\right), \qquad j \in \{-1, 0, \ldots, m\}. \tag{83}$$

By the definition of $\mathbf{q}_k$ in the theorem statement,

$$q_k(j) \overset{(a)}{=} \widehat{q}_k(j), \qquad j = 0, \ldots, m-1. \tag{84}$$

Equality (a) follows because (83) and the theorem statement use the same normalization factor and the same cosine formula at these indices. To identify the transform type explicitly, set $r := k - 1 \in \{0, \ldots, m-1\}$. Then

$$\left(j + \frac{1}{2}\right)\theta_k \overset{(a)}{=} \frac{2\pi\left(j + \frac{1}{2}\right)\left(r + \frac{1}{2}\right)}{2m+1}, \qquad j, r \in \{0, \ldots, m-1\}.$$

Equality (a) uses $r = k - 1$, and hence $2k - 1 = 2\left(r + \frac{1}{2}\right)$. Therefore the entries of $\mathbf{Q}_m$ are the type-VIII discrete-cosine kernel with the normalization factor $\frac{2}{\sqrt{2m+1}}$; the orthonormality of this normalization is proved below rather than assumed.

We first verify the two boundary relations encoded by the extended sequence. At the outer boundary,

$$\widehat{q}_k(-1) \overset{(a)}{=} \alpha_m \cos\left(-\frac{\theta_k}{2}\right) \overset{(b)}{=} \alpha_m \cos\left(\frac{\theta_k}{2}\right) \overset{(c)}{=} \widehat{q}_k(0). \tag{85}$$

Here (a) substitutes $j = -1$ into (83), (b) uses the evenness of the cosine, and (c) substitutes $j = 0$ into the same definition. At the anchored boundary,

$$\left(m + \frac{1}{2}\right)\theta_k \overset{(a)}{=} \frac{2m+1}{2}\frac{(2k-1)\,\pi}{2m+1} \overset{(b)}{=} \frac{(2k-1)\,\pi}{2}, \tag{86}$$

where (a) uses $m + \frac{1}{2} = \frac{2m+1}{2}$ and the definition of θ_k, while (b) cancels the nonzero factor $2m+1$. Consequently,

$$\widehat{q}_k(m) \overset{(a)}{=} \alpha_m \cos\left(\frac{(2k-1)\,\pi}{2}\right) \overset{(b)}{=} 0. \tag{87}$$

Here (a) combines (83) and (86), and (b) uses that the cosine of every odd multiple of $\frac{\pi}{2}$ is zero.

We now prove the eigenvalue relation $\mathbf{T}_m\mathbf{q}_k = \nu_k\mathbf{q}_k$. The entrywise definition (82), together with (85) and (87), gives, for every $j \in \{0, \ldots, m-1\}$,

$$(\mathbf{T}_m\mathbf{q}_k)_j \overset{(a)}{=} -\widehat{q}_k(j-1) + 2\widehat{q}_k(j) - \widehat{q}_k(j+1). \tag{88}$$

To justify (a) without suppressing the endpoint cases, observe the following. If $m = 1$, then $j = 0$, $\mathbf{T}_1 = (1)$, $\widehat{q}_k(-1) = q_k(0)$, and $\widehat{q}_k(1) = 0$, so the right-hand side of (88) equals $q_k(0)$. If $m \geq 2$ and $j = 0$, the first row of $\mathbf{T}_m$ gives $q_k(0) - q_k(1)$, which equals the displayed second difference because $\widehat{q}_k(-1) = q_k(0)$. If $m \geq 3$ and $1 \leq j \leq m-2$, the equality is the ordinary tridiagonal row formula. Finally, if $m \geq 2$ and $j = m-1$, the last row gives $-q_k(m-2) + 2q_k(m-1)$, which equals the displayed second difference because $\widehat{q}_k(m) = 0$. These cases exhaust all admissible pairs (m, j).

For fixed $k \in [m]$ and $j \in \{0, \ldots, m-1\}$, define

$$a_{j,k} := \left(j + \frac{1}{2}\right)\theta_k. \tag{89}$$

Then

$$\widehat{q}_k(j-1) \overset{(a)}{=} \alpha_m \cos\left(a_{j,k} - \theta_k\right), \qquad \widehat{q}_k(j) \overset{(b)}{=} \alpha_m \cos\left(a_{j,k}\right),$$
$$\widehat{q}_k(j+1) \overset{(c)}{=} \alpha_m \cos\left(a_{j,k} + \theta_k\right). \tag{90}$$

Equalities (a)–(c) follow by substituting $j-1$, j, and $j+1$ into (83) and using (89). Therefore,

$$\begin{aligned}\left(\mathbf{T}_m \mathbf{q}_k\right)_j &\overset{(a)}{=} \alpha_m \left[-\cos\left(a_{j,k} - \theta_k\right) + 2\cos\left(a_{j,k}\right) - \cos\left(a_{j,k} + \theta_k\right)\right] \\ &\overset{(b)}{=} \alpha_m \left[2\cos\left(a_{j,k}\right) - 2\cos\left(a_{j,k}\right)\cos\left(\theta_k\right)\right] \overset{(c)}{=} \left(2 - 2\cos\left(\theta_k\right)\right)\alpha_m \cos\left(a_{j,k}\right) \\ &\overset{(d)}{=} \nu_k \widehat{q}_k(j) \overset{(e)}{=} \nu_k q_k(j).\end{aligned} \tag{91}$$

Here (a) combines (88) and (90); (b) uses

$$\cos\left(a_{j,k} - \theta_k\right) + \cos\left(a_{j,k} + \theta_k\right) \overset{(a)}{=} 2\cos\left(a_{j,k}\right)\cos\left(\theta_k\right),$$

which is the elementary cosine addition identity; (c) factors the common quantity $\alpha_m \cos\left(a_{j,k}\right)$; (d) uses the definitions of ν_k, $\widehat{q}_k(j)$, and $a_{j,k}$; and (e) applies (84). Since (91) holds at every component,

$$\mathbf{T}_m \mathbf{q}_k \overset{(a)}{=} \nu_k \mathbf{q}_k, \qquad k \in [m]. \tag{92}$$

Equality (a) follows from equality of the corresponding components in $\mathbb{R}^m$.

We next prove that the eigenvalues $\nu_1, \ldots, \nu_m$ are positive and pairwise distinct. For every $k \in [m]$,

$$1 \overset{(a)}{\leq} 2k - 1 \overset{(b)}{\leq} 2m - 1 \overset{(c)}{<} 2m + 1. \tag{93}$$

Here (a) uses $k \geq 1$ and (b) uses $k \leq m$. Since $2m+1 > 0$, multiplication by $\frac{\pi}{2m+1} > 0$ preserves the inequalities and gives

$$0 \overset{(a)}{<} \theta_k \overset{(b)}{<} \pi. \tag{94}$$

Here (a) follows from the first inequality in (93), and (b) follows from its last strict inequality. Moreover, if $1 \leq k < \ell \leq m$, then

$$2k - 1 \overset{(a)}{<} 2\ell - 1 \quad \Longrightarrow \quad \theta_k \overset{(b)}{<} \theta_\ell. \tag{95}$$

Inequality (a) follows from $k < \ell$, and the implication to inequality (b) follows after multiplication by the positive factor $\frac{\pi}{2m+1}$. Since $\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left(2 - 2\cos\vartheta\right) = 2\sin\vartheta > 0$ on $(0, \pi)$, the map $\vartheta \mapsto 2 - 2\cos\vartheta$ is strictly increasing there. Combining (94), (95) with this monotonicity yields

$$\begin{aligned}\nu_k &\overset{(a)}{>} 0, \qquad k \in [m], \\ \nu_k &\overset{(b)}{<} \nu_\ell, \qquad 1 \leq k < \ell \leq m.\end{aligned} \tag{96}$$

Here (a) uses $\theta_k > 0$ and $2 - 2\cos\left(0\right) = 0$, while (b) uses strict monotonicity and the strict ordering of the angles. In particular, the values $\nu_1, \ldots, \nu_m$ are positive and pairwise distinct.

We now prove pairwise orthogonality directly, without invoking the spectral theorem. The entrywise formula (82) is invariant under interchange of r and s, because the conditions $r = s$, $r = s = 0$, and $|r - s| = 1$ are all symmetric. Thus

$$\tau_{rs} \overset{(a)}{=} \tau_{sr} \quad \Longrightarrow \quad \mathbf{T}_m^\top \overset{(b)}{=} \mathbf{T}_m. \tag{97}$$

Here (a) follows from the cases in (82), and (b) is the entrywise definition of a symmetric matrix. Fix $k, \ell \in [m]$ with $k \neq \ell$. Then

$$\begin{aligned}\nu_k \mathbf{q}_k^\top \mathbf{q}_\ell &\overset{(a)}{=} \left(\nu_k \mathbf{q}_k\right)^\top \mathbf{q}_\ell \overset{(b)}{=} \left(\mathbf{T}_m \mathbf{q}_k\right)^\top \mathbf{q}_\ell \\ &\overset{(c)}{=} \mathbf{q}_k^\top \mathbf{T}_m^\top \mathbf{q}_\ell \overset{(d)}{=} \mathbf{q}_k^\top \mathbf{T}_m \mathbf{q}_\ell \\ &\overset{(e)}{=} \mathbf{q}_k^\top \left(\nu_\ell \mathbf{q}_\ell\right) \overset{(f)}{=} \nu_\ell \mathbf{q}_k^\top \mathbf{q}_\ell.\end{aligned} \tag{98}$$

Here (a) moves the real scalar ν_k inside the transpose; (b) applies the eigenvalue relation (92) for k; (c) uses $(\mathbf{A}\mathbf{x})^\top = \mathbf{x}^\top \mathbf{A}^\top$; (d) uses (97); (e) applies (92) for ℓ; and (f) extracts the scalar ν_ℓ. Subtracting the right-hand side from the left-hand side gives

$$(\nu_k - \nu_\ell)\, \mathbf{q}_k^\top \mathbf{q}_\ell \overset{(a)}{=} 0.$$

Equality (a) is a rearrangement of (98). Since $\nu_k \neq \nu_\ell$ by (96), division by the nonzero scalar $\nu_k - \nu_\ell$ yields

$$\mathbf{q}_k^\top \mathbf{q}_\ell \overset{(a)}{=} 0, \qquad k \neq \ell. \tag{99}$$

Equality (a) uses the elementary fact that a product of a nonzero real number and a real number can vanish only if the second factor vanishes.

It remains to compute the norm of each $\mathbf{q}_k$. By definition,

$$\|\mathbf{q}_k\|_2^2 \overset{(a)}{=} \sum_{j=0}^{m-1} q_k(j)^2 \overset{(b)}{=} \frac{4}{2m+1} \sum_{j=0}^{m-1} \cos^2\left(\left(j+\frac{1}{2}\right)\theta_k\right). \tag{100}$$

Here (a) is the definition of the squared Euclidean norm, and (b) substitutes the component formula and uses $\alpha_m^2 = \frac{4}{2m+1}$. The identity $\cos^2(x) = \frac{1+\cos(2x)}{2}$ gives

$$\begin{aligned}\sum_{j=0}^{m-1} \cos^2\left(\left(j+\frac{1}{2}\right)\theta_k\right) &\overset{(a)}{=} \frac{1}{2} \sum_{j=0}^{m-1} \left[1 + \cos\left((2j+1)\,\theta_k\right)\right] \\ &\overset{(b)}{=} \frac{m}{2} + \frac{1}{2} \sum_{j=0}^{m-1} \cos\left((2j+1)\,\theta_k\right).\end{aligned} \tag{101}$$

Here (a) applies the double-angle identity term by term, and (b) uses $\sum_{j=0}^{m-1} 1 = m$ and separates the two finite sums.

Define

$$S_k := \sum_{j=0}^{m-1} \cos\left((2j+1)\,\theta_k\right).$$

For every $j \in \{0, \ldots, m-1\}$, the sine addition formula gives

$$2\sin(\theta_k)\cos\left((2j+1)\,\theta_k\right) \overset{(a)}{=} \sin\left((2j+2)\,\theta_k\right) - \sin(2j\theta_k). \tag{102}$$

Indeed, (a) is the identity $2\sin(x)\cos(y) = \sin(x+y) + \sin(x-y)$ with $x = \theta_k$ and $y = (2j+1)\,\theta_k$, together with the oddness of the sine. Summing (102) from $j = 0$ to $j = m-1$ yields

$$\begin{aligned}2\sin(\theta_k)\, S_k &\overset{(a)}{=} \sum_{j=0}^{m-1} \left[\sin\left((2j+2)\,\theta_k\right) - \sin(2j\theta_k)\right] \\ &\overset{(b)}{=} \sin(2m\theta_k) - \sin(0) \overset{(c)}{=} \sin(2m\theta_k).\end{aligned} \tag{103}$$

Here (a) uses the definition of S_k and linearity of finite summation; (b) uses telescopic cancellation, since every intermediate term $\sin(2r\theta_k)$ for $r = 1, \ldots, m-1$ appears once with sign $+1$ and once with sign -1; and (c) uses $\sin(0) = 0$. Furthermore,

$$2m\theta_k \overset{(a)}{=} (2m+1)\,\theta_k - \theta_k \overset{(b)}{=} (2k-1)\,\pi - \theta_k, \tag{104}$$

where (a) adds and subtracts θ_k, and (b) uses $(2m+1)\,\theta_k = (2k-1)\,\pi$. Since $2k-1$ is odd,

$$\sin(2m\theta_k) \overset{(a)}{=} \sin\left((2k-1)\,\pi - \theta_k\right) \overset{(b)}{=} \sin(\theta_k). \tag{105}$$

Here (a) applies (104), and (b) uses $\sin\left((2r-1)\,\pi - x\right) = \sin(x)$ for every integer r. Combining (103) and (105) gives

$$2\sin(\theta_k)\, S_k \overset{(a)}{=} \sin(\theta_k).$$

Equality (a) substitutes (105) into (103). By (94), $\sin(\theta_k) > 0$, and hence it is nonzero. Division by $2\sin(\theta_k)$ therefore yields

$$S_k \overset{(a)}{=} \frac{1}{2}. \tag{106}$$

Equality (a) performs the valid division by the nonzero quantity $2\sin(\theta_k)$. Substituting (106) into (101) gives

$$\sum_{j=0}^{m-1} \cos^2\left(\left(j+\frac{1}{2}\right)\theta_k\right) \overset{(a)}{=} \frac{m}{2} + \frac{1}{2}\left(\frac{1}{2}\right) \overset{(b)}{=} \frac{2m+1}{4}. \tag{107}$$

Here (a) uses (106), and (b) puts the two terms over the common denominator 4. Finally,

$$\|\mathbf{q}_k\|_2^2 \overset{(a)}{=} \frac{4}{2m+1}\frac{2m+1}{4} \overset{(b)}{=} 1. \tag{108}$$

Here (a) combines (100) and (107), and (b) cancels the positive factor $2m+1$ and the nonzero factor 4.

Equations (99) and (108) imply

$$\mathbf{q}_k^\top \mathbf{q}_\ell \overset{(a)}{=} \delta_{k\ell}, \qquad k,\ell \in [m]. \tag{109}$$

Here (a) uses the zero cross-products when $k \neq \ell$ and the unit norm when $k = \ell$. To verify explicitly that the family is a basis, suppose that $a_1, \ldots, a_m \in \mathbb{R}$ satisfy

$$\sum_{k=1}^{m} a_k \mathbf{q}_k \overset{(a)}{=} \mathbf{0}. \tag{110}$$

For every fixed $\ell \in [m]$, left multiplication by $\mathbf{q}_\ell^\top$ gives

$$0 \overset{(a)}{=} \mathbf{q}_\ell^\top \mathbf{0} \overset{(b)}{=} \mathbf{q}_\ell^\top \sum_{k=1}^{m} a_k \mathbf{q}_k \overset{(c)}{=} \sum_{k=1}^{m} a_k \mathbf{q}_\ell^\top \mathbf{q}_k$$
$$\overset{(d)}{=} \sum_{k=1}^{m} a_k \delta_{\ell k} \overset{(e)}{=} a_\ell.$$

Here (a) uses $\mathbf{q}_\ell^\top \mathbf{0} = 0$; (b) applies (110); (c) uses linearity of the inner product in a finite sum; (d) uses (109); and (e) uses the defining property of the Kronecker delta. Thus $a_\ell = 0$ for every $\ell \in [m]$, so the family is linearly independent. Since it contains exactly m vectors in the m-dimensional space $\mathbb{R}^m$, it is a basis. This proves i.

We next prove ii. Define

$$\mathbf{Q}_m := [\mathbf{q}_1, \ldots, \mathbf{q}_m], \qquad \mathbf{\Theta} := \operatorname{diag}(\nu_1, \ldots, \nu_m).$$

For $k,\ell \in [m]$, the (k,ℓ) entry of $\mathbf{Q}_m^\top \mathbf{Q}_m$ satisfies

$$\left(\mathbf{Q}_m^\top \mathbf{Q}_m\right)_{k\ell} \overset{(a)}{=} \mathbf{q}_k^\top \mathbf{q}_\ell \overset{(b)}{=} \delta_{k\ell} \overset{(c)}{=} (\mathbf{I}_m)_{k\ell}.$$

Here (a) is the rule for multiplying a matrix by its transpose in terms of its columns; (b) applies (109); and (c) is the entrywise definition of the identity matrix. Equality of all entries gives

$$\mathbf{Q}_m^\top \mathbf{Q}_m \overset{(a)}{=} \mathbf{I}_m. \tag{111}$$

Equality (a) follows because the two $m \times m$ matrices have identical entries. Since the columns of $\mathbf{Q}_m$ form a basis, $\mathbf{Q}_m$ is invertible. Multiplying (111) from the right by $\mathbf{Q}_m^{-1}$ gives

$$\mathbf{Q}_m^\top \overset{(a)}{=} \mathbf{Q}_m^{-1},$$

where (a) uses $\mathbf{Q}_m^\top \mathbf{Q}_m \mathbf{Q}_m^{-1} = \mathbf{I}_m \mathbf{Q}_m^{-1}$. Therefore,

$$\mathbf{Q}_m \mathbf{Q}_m^\top \overset{(a)}{=} \mathbf{Q}_m \mathbf{Q}_m^{-1} \overset{(b)}{=} \mathbf{I}_m. \tag{112}$$

Here (a) substitutes $\mathbf{Q}_m^\top = \mathbf{Q}_m^{-1}$, and (b) uses the defining property of an inverse.

Collecting the eigenvalue identities (92) columnwise gives

$$\begin{aligned}\mathbf{T}_m\mathbf{Q}_m &\overset{(a)}{=} \mathbf{T}_m\left[\mathbf{q}_1,\dots,\mathbf{q}_m\right] \overset{(b)}{=} \left[\mathbf{T}_m\mathbf{q}_1,\dots,\mathbf{T}_m\mathbf{q}_m\right]\\ &\overset{(c)}{=} \left[\nu_1\mathbf{q}_1,\dots,\nu_m\mathbf{q}_m\right] \overset{(d)}{=} \mathbf{Q}_m\mathbf{\Theta}.\end{aligned} \tag{113}$$

Here (a) substitutes the definition of $\mathbf{Q}_m$; (b) applies matrix multiplication column by column; (c) uses (92); and (d) uses the definition of the diagonal matrix $\mathbf{\Theta}$. Consequently,

$$\begin{aligned}\mathbf{T}_m &\overset{(a)}{=} \mathbf{T}_m\mathbf{I}_m \overset{(b)}{=} \mathbf{T}_m\mathbf{Q}_m\mathbf{Q}_m^\top \overset{(c)}{=} \mathbf{Q}_m\mathbf{\Theta}\mathbf{Q}_m^\top\\ &\overset{(d)}{=} \mathbf{Q}_m\,\mathrm{diag}\left(\nu_1,\dots,\nu_m\right)\mathbf{Q}_m^\top.\end{aligned} \tag{114}$$

Here (a) inserts the identity on the right; (b) applies (112); (c) uses (113); and (d) substitutes the definition of $\mathbf{\Theta}$. This proves ii.

We now prove iii. Because $A > 0$, $G = 2m > 0$, and $h = \frac{2A}{G}$,

$$h \overset{(a)}{=} \frac{2A}{2m} \overset{(b)}{=} \frac{A}{m} \overset{(c)}{>} 0. \tag{115}$$

Here (a) substitutes $G = 2m$, (b) cancels the nonzero factor 2, and (c) uses $A > 0$ and $m \geq 1$. Define

$$\mathbf{Q} := \begin{pmatrix}\mathbf{Q}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}_m\end{pmatrix}, \qquad \mathbf{\Lambda} := \frac{1}{h}\begin{pmatrix}\mathbf{\Theta} & \mathbf{0}\\ \mathbf{0} & \mathbf{\Theta}\end{pmatrix}. \tag{116}$$

Using (111),

$$\begin{aligned}\mathbf{Q}^\top\mathbf{Q} &\overset{(a)}{=} \begin{pmatrix}\mathbf{Q}_m^\top & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}_m^\top\end{pmatrix}\begin{pmatrix}\mathbf{Q}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}_m\end{pmatrix}\\ &\overset{(b)}{=} \begin{pmatrix}\mathbf{Q}_m^\top\mathbf{Q}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}_m^\top\mathbf{Q}_m\end{pmatrix} \overset{(c)}{=} \begin{pmatrix}\mathbf{I}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_m\end{pmatrix} \overset{(d)}{=} \mathbf{I}_{2m} \overset{(e)}{=} \mathbf{I}_G.\end{aligned} \tag{117}$$

Here (a) transposes the block-diagonal matrix in (116); (b) performs the block multiplication; (c) uses (111); (d) identifies the resulting block matrix as the identity in dimension $2m$; and (e) uses $G = 2m$. Similarly, using (112),

$$\mathbf{Q}\mathbf{Q}^\top \overset{(a)}{=} \begin{pmatrix}\mathbf{Q}_m\mathbf{Q}_m^\top & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}_m\mathbf{Q}_m^\top\end{pmatrix} \overset{(b)}{=} \mathbf{I}_G. \tag{118}$$

Here (a) performs the analogous block multiplication, and (b) applies (112) and $G = 2m$. Thus $\mathbf{Q}$ is orthogonal.

Substituting (114) into (81) gives

$$\begin{aligned}\mathbf{K}_0 &\overset{(a)}{=} \frac{1}{h}\begin{pmatrix}\mathbf{Q}_m\mathbf{\Theta}\mathbf{Q}_m^\top & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}_m\mathbf{\Theta}\mathbf{Q}_m^\top\end{pmatrix}\\ &\overset{(b)}{=} \begin{pmatrix}\mathbf{Q}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}_m\end{pmatrix}\left[\frac{1}{h}\begin{pmatrix}\mathbf{\Theta} & \mathbf{0}\\ \mathbf{0} & \mathbf{\Theta}\end{pmatrix}\right]\begin{pmatrix}\mathbf{Q}_m^\top & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}_m^\top\end{pmatrix}\\ &\overset{(c)}{=} \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top.\end{aligned} \tag{119}$$

Here (a) substitutes the eigendecomposition of both copies of $\mathbf{T}_m$; (b) factors the block-diagonal product and places the scalar factor $\frac{1}{h}$ in the middle block; and (c) applies (116). This proves iii.

It remains to prove iv. The elementary half-angle identity gives

$$\nu_k \overset{(a)}{=} 2 - 2\cos\left(\theta_k\right) \overset{(b)}{=} 4\sin^2\left(\frac{\theta_k}{2}\right), \tag{120}$$

where (a) is the definition of ν_k, and (b) uses $1-\cos(x) = 2\sin^2\left(\frac{x}{2}\right)$. Thus the distinct eigenvalue values appearing on the diagonal of $\mathbf{\Lambda}$ are

$$\lambda_k \overset{(a)}{=} \frac{\nu_k}{h} \overset{(b)}{=} \frac{4}{h}\sin^2\left(\frac{\theta_k}{2}\right) \overset{(c)}{=} \frac{4}{h}\sin^2\left(\frac{(2k-1)\,\pi}{2\,(2m+1)}\right), \qquad k = 1, \ldots, m.$$

Here (a) reads the diagonal scaling from (116); (b) applies (120); and (c) substitutes the definition of θ_k. Because $h > 0$ by (115) and the ν_k are strictly ordered by (96),

$$\lambda_k \overset{(a)}{>} 0, \qquad k \in [m], \\ \lambda_k \overset{(b)}{<} \lambda_\ell, \qquad 1 \le k < \ell \le m. \tag{121}$$

Here (a) and (b) follow by division of the corresponding inequalities in (96) by the positive scalar h.

By (116),

$$\mathbf{\Lambda} \overset{(a)}{=} \operatorname{diag}(\lambda_1, \ldots, \lambda_m, \lambda_1, \ldots, \lambda_m). \tag{122}$$

Equality (a) uses $\lambda_k = \frac{\nu_k}{h}$ in each of the two diagonal blocks. Since (119) and (117) imply $\mathbf{Q}^{-1} = \mathbf{Q}^\top$, the matrices $\mathbf{K}_0$ and $\mathbf{\Lambda}$ are similar. Hence, for an indeterminate z,

$$\det(z\mathbf{I}_G - \mathbf{K}_0) \overset{(a)}{=} \det\left(z\mathbf{I}_G - \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top\right) \overset{(b)}{=} \det\left(\mathbf{Q}\,(z\mathbf{I}_G - \mathbf{\Lambda})\,\mathbf{Q}^\top\right) \\ \overset{(c)}{=} \det(\mathbf{Q})\det(z\mathbf{I}_G - \mathbf{\Lambda})\det\left(\mathbf{Q}^\top\right) \overset{(d)}{=} \det\left(\mathbf{Q}\mathbf{Q}^\top\right)\det(z\mathbf{I}_G - \mathbf{\Lambda}) \\ \overset{(e)}{=} \det(z\mathbf{I}_G - \mathbf{\Lambda}) \overset{(f)}{=} \prod_{k=1}^{m}(z - \lambda_k)^2. \tag{123}$$

Here (a) applies (119); (b) uses $z\mathbf{I}_G = \mathbf{Q}\,(z\mathbf{I}_G)\,\mathbf{Q}^\top$, which follows from (118); (c) uses multiplicativity of the determinant; (d) uses $\det(\mathbf{Q})\det\left(\mathbf{Q}^\top\right) = \det\left(\mathbf{Q}\mathbf{Q}^\top\right)$; (e) applies (118) and $\det(\mathbf{I}_G) = 1$; and (f) uses the repeated diagonal form (122). By (121), the numbers $\lambda_1, \ldots, \lambda_m$ are distinct. Therefore each factor $z - \lambda_k$ occurs exactly twice in the characteristic polynomial (123), and each λ_k has algebraic multiplicity two. This proves iv and completes the proof.

□

Table 4: Summary of the auxiliary theoretical results. Their logical dependencies are illustrated in Fig. 2. For the main theoretical results, see Table 3.

Result	Content	Page
Lemma D1	Structure of the stiffness matrix	page 34
Lemma D2	Kernel of the stiffness matrix	page 36
Lemma D3	Block decomposition of the anchored stiffness matrix	page 38
Lemma D4	Brownian energy in increment coordinates	page 40
Lemma D5	Brownian energy in spectral coordinates	page 41
Lemma D6	Explicit spectral Brownian basis	page 44

C Proofs of the Kernel, Conditioning, and Optimization Results

This section completes the explicit kernel, approximation, conditioning, and optimization consequences of the foundational results proved above. We first establish the positive definiteness of the anchored stiffness matrix in Lemma 3.1, identify the reproducing kernel and Brownian Gram inverse in Theorem 1, items i–ii, and derive the conforming-subspace and sharp-approximation results in Corollaries 2 and 3. We then obtain the exact stiffness and coordinate condition numbers in Corollary 4. Finally, we prove the linear-reparameterization identity in Proposition 2, the nodal–spectral

and increment–Brownian gradient correspondences in Proposition 3, their recursive blockwise extension in Corollary 5, the mesh-independent increment least-squares bound in Proposition 4, and the signed-permutation characterization of standard Adam equivariance in Proposition 5.

Throughout, $\mathcal{J} = \{0, \dots, G\} \setminus \{i_0\}$ indexes the reduced anchored coordinates of Definition 4, and $\widetilde{\boldsymbol{\phi}}(s) = (\phi_j(s))_{j\in\mathcal{J}}$ is understood in that same reduced ordering. Thus, for $f, g \in \mathcal{H}_h^0$ with reduced coordinates $\widetilde{\mathbf{v}}, \widetilde{\mathbf{u}}$,

$$f(s) = \widetilde{\boldsymbol{\phi}}(s)^\top \widetilde{\mathbf{v}}, \qquad \langle f, g\rangle_{B,h} = \widetilde{\mathbf{v}}^\top \mathbf{K}_0 \widetilde{\mathbf{u}}. \tag{124}$$

C.1 Proof of Lemma 3.1

By Lemma D2, $\ker \mathbf{K} = \operatorname{span}(\mathbf{1})$, and by Proposition 1, $\mathbf{K} \succeq 0$. If $\mathbf{x}^\top \mathbf{K}_0 \mathbf{x} = 0$, then $(\mathbf{R}\mathbf{x})^\top \mathbf{K} (\mathbf{R}\mathbf{x}) = 0$, so $\mathbf{R}\mathbf{x} = c\mathbf{1}$. The anchor entry of $\mathbf{R}\mathbf{x}$ is zero; hence $c = 0$. Since $\mathbf{R}^\top \mathbf{R} = \mathbf{I}_G$, $\mathbf{x} = 0$. Thus $\mathbf{K}_0 \succ 0$, and (124) is an inner product. □

C.2 Proof of Theorem 1, items (i)–(ii)

Finite-dimensional RKHS existence and dimension are proved in Appendix B.3. Fix $t \in I$. If the representer $k_h(\cdot, t)$ has reduced vector $\widetilde{\mathbf{k}}_t$, then for every $\widetilde{\mathbf{v}}$,

$$\widetilde{\mathbf{v}}^\top \mathbf{K}_0 \widetilde{\mathbf{k}}_t = f(t) = \widetilde{\boldsymbol{\phi}}(t)^\top \widetilde{\mathbf{v}}.$$

Therefore $\mathbf{K}_0 \widetilde{\mathbf{k}}_t = \widetilde{\boldsymbol{\phi}}(t)$, and positive definiteness gives $\widetilde{\mathbf{k}}_t = \mathbf{K}_0^{-1} \widetilde{\boldsymbol{\phi}}(t)$, proving item (i).

For $r \in \mathcal{J}$, set $g_r = k_{\mathrm{B}}(\cdot, t_r)$. Its only possible kinks are at 0 and t_r, both grid nodes, so $g_r \in \mathcal{H}_h^0$. If $t_r > 0$, then $g_r' = \mathbb{1}_{(0,t_r)}$ almost everywhere; if $t_r < 0$, then $g_r' = -\mathbb{1}_{(t_r,0)}$. Hence, for every $f \in \mathcal{H}_h^0$,

$$\langle f, g_r\rangle_{B,h} = \begin{cases} \int_0^{t_r} f'(t)\,\mathrm{d}t, & t_r > 0, \\ -\int_{t_r}^0 f'(t)\,\mathrm{d}t, & t_r < 0, \end{cases} = f(t_r) - f(0) = f(t_r).$$

Thus $g_r = k_h(\cdot, t_r)$. Since $\widetilde{\boldsymbol{\phi}}(t_r) = \mathbf{e}_{\pi(r)}$,

$$(\mathbf{K}_0^{-1})_{\pi(r),\pi(q)} = k_h(t_r, t_q) = k_{\mathrm{B}}(t_r, t_q).$$

Substitution into item (i) gives

$$k_h(s,t) = \sum_{r,q\in\mathcal{J}} \phi_r(s) k_{\mathrm{B}}(t_r, t_q) \phi_q(t),$$

which is the tensor-product bilinear interpolant. If one argument is a node, one sum collapses and the preceding representer identity gives exact agreement. □

C.3 Proof of Corollary 2

The functions $g_r = k_{\mathrm{B}}(\cdot, t_r)$ belong to $\mathcal{H}_h^0$, and their Gram matrix is $\big(k_{\mathrm{B}}(t_r, t_q)\big)_{r,q} = \mathbf{K}_0^{-1}$ in reduced order. They are therefore linearly independent, and their number equals $\dim \mathcal{H}_h^0 = G$, so they span the space. Every $g \in \mathcal{H}_h^0$ is continuous, piecewise affine, anchored, and has derivative in $L^2(I)$; hence $g \in \mathcal{H}_{\mathrm{B}}$ and $\|g\|_{\mathcal{H}_{\mathrm{B}}}^2 = \int_I (g')^2 = \|g\|_{B,h}^2$.

Let $I_h f$ be nodal interpolation of $f \in \mathcal{H}_{\mathrm{B}}$. For every $r \in \mathcal{J}$,

$$\langle f - I_h f, g_r\rangle_{\mathcal{H}_{\mathrm{B}}} = (f - I_h f)(t_r) = 0.$$

Since the g_r span $\mathcal{H}_h^0$, the residual is orthogonal to that space, so $I_h = \Pi_h$. □

C.4 Proof of Corollary 3

Fix a cell $T_i = [a, b]$ and write $h = b - a$. No cell crosses the anchor because 0 is a grid node. Suppose first $0 \le a \le s \le t \le b$, and set $\alpha = (s-a)/h$, $\beta = (t-a)/h$. The four Brownian corner values are a, a, a, b, so bilinear interpolation gives

$$k_h(s,t) = (1-\alpha)(1-\beta)a + (1-\alpha)\beta a + \alpha(1-\beta)a + \alpha\beta b = a + \alpha\beta h.$$

Since $k_{\mathrm{B}}(s,t) = s = a + \alpha h$,

$$k_{\mathrm{B}}(s,t) - k_h(s,t) = \alpha(1-\beta)h = \frac{(s-a)(b-t)}{h}.$$

For $a \le s \le t \le b \le 0$, the corner values are $-a, -b, -b, -b$ and the same bilinear calculation gives $k_h(s,t) = -b + (1-\alpha)(1-\beta)h$. Since $k_{\mathrm{B}}(s,t) = -t$, the same residual $(s-a)(b-t)/h$ follows. Symmetry replaces s, t by their minimum and maximum. On two different cells the diagonal kink $s = t$ is absent from the corresponding rectangle and k_{B} is affine in each argument there; hence its bilinear interpolant is exact. This proves (1).

Because Π_h is orthogonal, the error functional $f \mapsto f(t) - \Pi_h f(t)$ has representer $r_h(\cdot, t)$ and squared norm $r_h(t,t)$. Cauchy–Schwarz gives (2); equality holds for a normalized residual section whenever t is not a node. Maximizing $(t-a)(b-t)/h$ on a cell gives $h/4$, proving the sharp $\sqrt{h}/2$ constant.

The Pythagorean identity follows from $f - \Pi_h f \perp \Pi_h f$. For the L^2 bound, set $e = f - \Pi_h f$. On every cell, $e(a) = e(b) = 0$, so the sharp Wirtinger inequality gives

$$\|e\|^2_{L^2(T_i)} \le \frac{h^2}{\pi^2} \|e'\|^2_{L^2(T_i)}.$$

The derivative of $\Pi_h f$ on T_i is the cell average of f', so e' is the residual after the $L^2(T_i)$-orthogonal projection onto constants. Summing over cells,

$$\|e\|^2_{L^2(I)} \le \frac{h^2}{\pi^2} \|e'\|^2_{L^2(I)} \le \frac{h^2}{\pi^2} \|f'\|^2_{L^2(I)}.$$

Sharpness follows by taking a sine arch supported on one cell, $f(t) = \sin(\pi(t-a)/h)$ on that cell and zero elsewhere; it has zero nodal values and attains the Wirtinger constant. □

C.5 Proof of Corollary 4

By Theorem 3, the distinct stiffness eigenvalues are

$$\lambda_k = \frac{4}{h} \sin^2\left(\frac{(2k-1)\pi}{2(2m+1)}\right), \qquad k \in [m].$$

They increase with k. Thus

$$\lambda_{\min} = \frac{4}{h} \sin^2\left(\frac{\pi}{4m+2}\right),$$
$$\lambda_{\max} = \frac{4}{h} \cos^2\left(\frac{\pi}{2m+1}\right),$$

which gives (3). Since $\mathbf{D}_0^\top \mathbf{D}_0 = h\mathbf{K}_0$, its squared singular values are $h\lambda_k$. Taking square roots yields (4) and (5). Taylor expansion of sine and cosine gives the stated asymptotics. In particular $h\lambda_{\max} \to 4$; writing $\lambda_{\max} \to 4/h$ would be meaningless because h varies with m. □

C.6 Proof of Proposition 2

The chain rule gives $\nabla L_{\mathbf{A}}(\mathbf{z}) = \mathbf{A}^{-\top} \nabla L(\mathbf{A}^{-1}\mathbf{z})$. Applying $\mathbf{A}^{-1}$ to the update gives

$$\begin{aligned}\widetilde{\mathbf{v}}_{k+1} &= \mathbf{A}^{-1}\mathbf{z}_k - \eta_k \mathbf{A}^{-1}\mathbf{A}^{-\top}\nabla L(\widetilde{\mathbf{v}}_k) \\ &= \widetilde{\mathbf{v}}_k - \eta_k (\mathbf{A}^\top \mathbf{A})^{-1} \nabla L(\widetilde{\mathbf{v}}_k).\end{aligned}$$

The stochastic statement is identical with a shared stochastic gradient and its chain-rule transform. □

C.7 Proof of Proposition 3

For $\mathbf{A} = \mathbf{Q}^\top$, orthogonality gives $\mathbf{A}^\top \mathbf{A} = \mathbf{I}_G$, so (6) is exactly the nodal update. Mapped initializations and induction give identical function iterates. For $\mathbf{A} = \mathbf{D}_0$, $\mathbf{A}^\top \mathbf{A} = h\mathbf{K}_0$, and (6) becomes (7). The Riemannian gradient for the constant metric with Gram matrix $\mathbf{K}_0$ is $\mathbf{K}_0^{-1}\nabla L$; hence the discrete update is explicit Euler for that flow with time increment η_k/h. □

C.8 Proof of Corollary 5

Apply Proposition 2 to the block-diagonal matrix $\mathbf{A}$. Its Gram matrix and inverse are block diagonal:

$$(\mathbf{A}^\top\mathbf{A})^{-1} = \mathrm{blkdiag}\left((\mathbf{A}_1^\top\mathbf{A}_1)^{-1}, \ldots, (\mathbf{A}_B^\top\mathbf{A}_B)^{-1}, \mathbf{I}\right).$$

If every $\mathbf{A}_b$ is orthogonal, this is the identity. If every $\mathbf{A}_b = \mathbf{D}_{0,b}$, the bth block is $(h_b\mathbf{K}_{0,b})^{-1}$. The proof never separates the objective by blocks, so arbitrary cross-layer coupling is allowed. □

C.9 Proof of Proposition 4

Let

$$\mathbf{B} := \frac{1}{n}\sum_{i=1}^{n}\widetilde{\boldsymbol{\phi}}(x_i)\widetilde{\boldsymbol{\phi}}(x_i)^\top \succeq 0.$$

The nodal Hessian is $\mathbf{H}_{\mathrm{nod}} = \mathbf{B} + \rho\mathbf{K}_0$; the spectral Hessian is $\mathbf{Q}^\top\mathbf{H}_{\mathrm{nod}}\mathbf{Q}$, so the two have identical spectra. Since $\mathbf{D}_0^\top\mathbf{D}_0 = h\mathbf{K}_0$, the polar decomposition has the form $\mathbf{D}_0 = \sqrt{h}\,\mathbf{U}\mathbf{K}_0^{1/2}$ for an orthogonal $\mathbf{U}$. Therefore

$$\begin{aligned}\mathbf{H}_{\mathrm{inc}} &= \mathbf{D}_0^{-\top}(\mathbf{B} + \rho\mathbf{K}_0)\mathbf{D}_0^{-1} \\ &= \frac{1}{h}\mathbf{U}\left(\mathbf{K}_0^{-1/2}\mathbf{B}\mathbf{K}_0^{-1/2} + \rho\mathbf{I}_G\right)\mathbf{U}^\top. \end{aligned} \tag{125}$$

Set $\mathbf{C} = \mathbf{K}_0^{-1/2}\mathbf{B}\mathbf{K}_0^{-1/2} \succeq 0$. Then

$$\begin{aligned}\lambda_{\max}(\mathbf{C}) \le \mathrm{tr}(\mathbf{C}) &= \frac{1}{n}\sum_{i=1}^{n}\widetilde{\boldsymbol{\phi}}(x_i)^\top\mathbf{K}_0^{-1}\widetilde{\boldsymbol{\phi}}(x_i) \\ &= \frac{1}{n}\sum_{i=1}^{n}k_h(x_i,x_i) \le \frac{1}{n}\sum_{i=1}^{n}k_{\mathrm{B}}(x_i,x_i) = \frac{1}{n}\sum_{i=1}^{n}|x_i| \le A.\end{aligned}$$

The penultimate inequality follows from the nonnegative residual diagonal in (2). Hence every eigenvalue of $\mathbf{C} + \rho\mathbf{I}_G$ lies in $[\rho, \rho + A]$, and (125) proves (8). If $\mathbf{B} = 0$, the increment Hessian is $(\rho/h)\mathbf{I}_G$, while nodal and spectral Hessians have condition number $\kappa_2(\mathbf{K}_0)$. □

C.10 Proof of Proposition 5

Write standard Adam, for gradient $\mathbf{g}_k$, as

$$\begin{aligned}\mathbf{m}_k &= \beta_1\mathbf{m}_{k-1} + (1-\beta_1)\mathbf{g}_k, \\ \mathbf{s}_k &= \beta_2\mathbf{s}_{k-1} + (1-\beta_2)(\mathbf{g}_k \odot \mathbf{g}_k), \\ \widehat{\mathbf{m}}_k &= \frac{\mathbf{m}_k}{1-\beta_1^{k+1}}, \qquad \widehat{\mathbf{s}}_k = \frac{\mathbf{s}_k}{1-\beta_2^{k+1}}, \\ \mathbf{x}_{k+1} &= \mathbf{x}_k - \eta_k\frac{\widehat{\mathbf{m}}_k}{\sqrt{\widehat{\mathbf{s}}_k} + \varepsilon\mathbf{1}},\end{aligned}$$

where the last two operations are coordinatewise. If $\mathbf{S}$ is a signed permutation, gradients and first moments transform by $\mathbf{S}^\top$, while squared gradients and second moments are permuted by $|\mathbf{S}|^\top$. Induction through the recursion shows that the mapped update is exactly the original one, so signed permutations are equivariances.

Conversely, let $\mathbf{U}$ be orthogonal and suppose Adam is equivariant for every gradient sequence. At the first step, bias correction reduces the Adam direction to $\mathbf{F}_\varepsilon(\mathbf{g}) = \mathbf{g}/(|\mathbf{g}| + \varepsilon\mathbf{1})$. Equivariance under $\mathbf{z} = \mathbf{U}^\top\mathbf{x}$ requires

$$\mathbf{U}\mathbf{F}_\varepsilon(\mathbf{U}^\top\mathbf{g}) = \mathbf{F}_\varepsilon(\mathbf{g}) \qquad \text{for every } \mathbf{g}. \tag{126}$$

Take $\mathbf{g} = t\mathbf{U}\mathbf{e}_j$ with $t > 0$ and let $t \to \infty$. The left side tends to $\mathbf{U}\mathbf{e}_j$, while the right side tends coordinatewise to $\mathrm{sgn}(\mathbf{U}\mathbf{e}_j)$. Thus every nonzero coordinate of every column $\mathbf{U}\mathbf{e}_j$ equals ± 1. Since each column has unit norm, it has exactly one nonzero entry; orthogonality then makes $\mathbf{U}$ a signed permutation.

For $m \ge 2$, the first column of each DCT-VIII block has m strictly positive entries, so $\mathbf{Q}$ is not a signed permutation. Failure of (126) for some vector gives a linear objective with that constant gradient; because the first step size is positive, the first mapped Adam iterates differ. □

D INTERNAL LEMMAS

This section proves the six auxiliary linear-algebra and coordinate results used in Sections B and C: the stiffness factorization and nullspace, the anchored block decomposition, the increment and spectral Brownian-energy identities, and the explicit spectral Brownian basis. Their logical roles are displayed in Figure 2, and their statements and locations are summarized in Table 4.

Lemma D1 (Structure of the stiffness matrix)**.** *The finite-element stiffness matrix* $\mathbf{K} = (K_{rs})_{r,s=0}^{G}$, *with*

$$K_{rs} = \int_{-A}^{A} \phi_r'(t)\phi_s'(t)\,\mathrm{d}t, \qquad r, s = 0, \dots, G,$$

satisfies

$$\mathbf{K} = \frac{1}{h}\mathbf{D}^\top\mathbf{D}.$$

Equivalently,

$$\mathbf{K} = \frac{1}{h}\begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & 2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

Consequently, for every $\mathbf{v} \in \mathbb{R}^{G+1}$ *and* $f = \mathcal{R}(\mathbf{v}) \in \mathcal{H}_h$,

$$\|f\|_{B,h}^2 = \frac{1}{h}\|\mathbf{D}\mathbf{v}\|_2^2 = \frac{1}{h}\sum_{i=0}^{G-1}(v_{i+1} - v_i)^2.$$

Proof. Let $\mathbf{v} = (v_0, \dots, v_G)^\top \in \mathbb{R}^{G+1}$ be arbitrary, and set $f = \mathcal{R}(\mathbf{v})$. Fix an interval index $i \in \{0, \dots, G-1\}$. For every $t \in [t_i, t_{i+1}]$, only the two basis functions ϕ_i and ϕ_{i+1} can be nonzero. Moreover, on this interval,

$$\phi_i(t) = \frac{t_{i+1} - t}{h}, \qquad \phi_{i+1}(t) = \frac{t - t_i}{h}.$$

Therefore,

$$f(t) \overset{(a)}{=} \sum_{j=0}^{G} v_j\phi_j(t) \overset{(b)}{=} v_i\phi_i(t) + v_{i+1}\phi_{i+1}(t)$$
$$\overset{(c)}{=} v_i\frac{t_{i+1} - t}{h} + v_{i+1}\frac{t - t_i}{h}.$$

Here (a) is the definition of the finite-element reconstruction operator, (b) uses the local support of the nodal basis functions, and (c) substitutes the explicit affine formulas for ϕ_i and ϕ_{i+1} on $[t_i, t_{i+1}]$. Differentiating the preceding affine expression on the open interval (t_i, t_{i+1}) gives

$$f'(t) \overset{(a)}{=} v_i\left(-\frac{1}{h}\right) + v_{i+1}\left(\frac{1}{h}\right) \overset{(b)}{=} \frac{v_{i+1} - v_i}{h}. \tag{127}$$

Here (a) differentiates the two affine terms, and (b) combines them over the common denominator h. The derivative may fail to exist at the finitely many grid points, but (127) holds almost everywhere on I, which is sufficient for every integral below.

By the definition of the discrete Brownian seminorm,

$$\|f\|_{B,h}^2 \overset{(a)}{=} \int_{-A}^{A}(f'(t))^2\,\mathrm{d}t \overset{(b)}{=} \sum_{i=0}^{G-1}\int_{t_i}^{t_{i+1}}(f'(t))^2\,\mathrm{d}t \overset{(c)}{=} \sum_{i=0}^{G-1}\int_{t_i}^{t_{i+1}}\left(\frac{v_{i+1} - v_i}{h}\right)^2\mathrm{d}t$$
$$\overset{(d)}{=} \sum_{i=0}^{G-1}\left(\frac{v_{i+1} - v_i}{h}\right)^2\int_{t_i}^{t_{i+1}}1\,\mathrm{d}t \overset{(e)}{=} \sum_{i=0}^{G-1}\left(\frac{v_{i+1} - v_i}{h}\right)^2(t_{i+1} - t_i)$$
$$\overset{(f)}{=} \sum_{i=0}^{G-1}\left(\frac{v_{i+1} - v_i}{h}\right)^2 h \overset{(g)}{=} \frac{1}{h}\sum_{i=0}^{G-1}(v_{i+1} - v_i)^2. \tag{128}$$

Here (a) expands the squared seminorm, (b) partitions $[-A, A]$ into its G mesh intervals, (c) substitutes (127), (d) moves the intervalwise constant factor outside each integral, (e) evaluates the integral of the constant function 1, (f) uses the uniform-grid identity $t_{i+1} - t_i = h$, and (g) simplifies $h/h^2 = 1/h$.

By the componentwise definition of the first-difference operator,

$$\|\mathbf{D}\mathbf{v}\|_2^2 \overset{(a)}{=} \sum_{i=0}^{G-1} (\mathbf{D}\mathbf{v})_i^2 \overset{(b)}{=} \sum_{i=0}^{G-1} (v_{i+1} - v_i)^2 . \tag{129}$$

Here (a) is the definition of the squared Euclidean norm on $\mathbb{R}^G$, and (b) uses $(\mathbf{D}\mathbf{v})_i = v_{i+1} - v_i$. Combining (128) and (129) yields

$$\|f\|_{B,h}^2 = \frac{1}{h} \|\mathbf{D}\mathbf{v}\|_2^2 = \frac{1}{h} \mathbf{v}^\top \mathbf{D}^\top \mathbf{D}\mathbf{v}. \tag{130}$$

The last equality follows from $\|\mathbf{w}\|_2^2 = \mathbf{w}^\top \mathbf{w}$ with $\mathbf{w} = \mathbf{D}\mathbf{v}$, together with $(\mathbf{D}\mathbf{v})^\top = \mathbf{v}^\top \mathbf{D}^\top$.

We next express the same energy through the stiffness matrix. By the entrywise definition of $\mathbf{K}$,

$$\begin{aligned}
\mathbf{v}^\top \mathbf{K} \mathbf{v} &\overset{(a)}{=} \sum_{r=0}^{G} \sum_{s=0}^{G} v_r K_{rs} v_s \overset{(b)}{=} \sum_{r=0}^{G} \sum_{s=0}^{G} v_r v_s \int_{-A}^{A} \phi_r'(t) \phi_s'(t) \, \mathrm{d}t \\
&\overset{(c)}{=} \int_{-A}^{A} \sum_{r=0}^{G} \sum_{s=0}^{G} v_r v_s \phi_r'(t) \phi_s'(t) \, \mathrm{d}t \overset{(d)}{=} \int_{-A}^{A} \left(\sum_{r=0}^{G} v_r \phi_r'(t) \right) \left(\sum_{s=0}^{G} v_s \phi_s'(t) \right) \mathrm{d}t \\
&\overset{(e)}{=} \int_{-A}^{A} (f'(t))^2 \, \mathrm{d}t \overset{(f)}{=} \|f\|_{B,h}^2 .
\end{aligned} \tag{131}$$

Here (a) expands the quadratic form componentwise, (b) substitutes the definition of K_{rs}, (c) interchanges the integral with two finite sums, (d) factors the double sum, (e) uses $f' = \sum_{r=0}^{G} v_r \phi_r'$ almost everywhere, and (f) is the definition of the discrete Brownian seminorm.

Combining (130) and (131) gives

$$\mathbf{v}^\top \left(\mathbf{K} - \frac{1}{h} \mathbf{D}^\top \mathbf{D} \right) \mathbf{v} = 0, \qquad \forall \mathbf{v} \in \mathbb{R}^{G+1}. \tag{132}$$

Define

$$\mathbf{M} := \mathbf{K} - \frac{1}{h} \mathbf{D}^\top \mathbf{D}.$$

The matrix $\mathbf{M}$ is symmetric because $K_{rs} = K_{sr}$ and $(\mathbf{D}^\top \mathbf{D})^\top = \mathbf{D}^\top \mathbf{D}$. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{G+1}$ be arbitrary. Then

$$2\mathbf{x}^\top \mathbf{M} \mathbf{y} \overset{(a)}{=} (\mathbf{x} + \mathbf{y})^\top \mathbf{M} (\mathbf{x} + \mathbf{y}) - \mathbf{x}^\top \mathbf{M} \mathbf{x} - \mathbf{y}^\top \mathbf{M} \mathbf{y} \overset{(b)}{=} 0.$$

Here (a) expands the first quadratic form and uses the symmetry identity $\mathbf{y}^\top \mathbf{M} \mathbf{x} = \mathbf{x}^\top \mathbf{M} \mathbf{y}$, while (b) applies (132) to $\mathbf{x} + \mathbf{y}$, $\mathbf{x}$, and $\mathbf{y}$. Thus $\mathbf{x}^\top \mathbf{M} \mathbf{y} = 0$ for all $\mathbf{x}, \mathbf{y}$. Taking $\mathbf{x} = \mathbf{e}_r$ and $\mathbf{y} = \mathbf{e}_s$ gives

$$M_{rs} \overset{(a)}{=} \mathbf{e}_r^\top \mathbf{M} \mathbf{e}_s \overset{(b)}{=} 0, \qquad r, s = 0, \ldots, G.$$

Here (a) extracts the (r, s) entry of $\mathbf{M}$, and (b) uses the preceding bilinear identity. Therefore $\mathbf{M} = \mathbf{0}$, and hence

$$\mathbf{K} = \frac{1}{h} \mathbf{D}^\top \mathbf{D}. \tag{133}$$

It remains to expand $\mathbf{D}^\top \mathbf{D}$. The entries of $\mathbf{D}$ are

$$D_{ir} = \begin{cases} -1, & r = i, \\ 1, & r = i + 1, \\ 0, & \text{otherwise}, \end{cases} \qquad i = 0, \ldots, G-1, \quad r = 0, \ldots, G.$$

Consequently,

$$\left(\mathbf{D}^\top \mathbf{D}\right)_{rs} \overset{(a)}{=} \sum_{i=0}^{G-1} D_{ir} D_{is} \overset{(b)}{=} \begin{cases} 1, & r = s \in \{0, G\}, \\ 2, & r = s \in \{1, \dots, G-1\}, \\ -1, & |r-s| = 1, \\ 0, & |r-s| \geq 2. \end{cases}$$

Here (a) is the entrywise formula for a matrix product. For (b), an endpoint column of $\mathbf{D}$ contains one nonzero entry of magnitude one, an interior column contains two nonzero entries of magnitude one, adjacent columns overlap in exactly one row with product -1, and nonadjacent columns have disjoint supports. Thus

$$\mathbf{D}^\top \mathbf{D} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & 2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

Combining this identity with (133) proves the matrix formula, while (128) and (129) prove the asserted energy identities.

□

Lemma D2 (Kernel of the stiffness matrix)**.** *The stiffness matrix satisfies*

$$\ker\left(\mathbf{K}\right) = \operatorname{span}\left(\mathbf{1}\right),$$

where

$$\mathbf{1} := (1, \dots, 1)^\top \in \mathbb{R}^{G+1}.$$

Proof. By Lemma D1, the stiffness matrix admits the factorization

$$\mathbf{K} = \frac{1}{h} \mathbf{D}^\top \mathbf{D}. \tag{134}$$

We first prove that

$$\ker(\mathbf{K}) = \ker(\mathbf{D}). \tag{135}$$

Let $\mathbf{v} \in \ker(\mathbf{K})$ be arbitrary. By the definition of the kernel,

$$\mathbf{K}\mathbf{v} = \mathbf{0}. \tag{136}$$

Multiplying this identity from the left by $\mathbf{v}^\top$ and using (134) gives

$$\begin{aligned} 0 &\overset{(a)}{=} \mathbf{v}^\top \mathbf{0} \overset{(b)}{=} \mathbf{v}^\top \left(\mathbf{K}\mathbf{v}\right) \overset{(c)}{=} \mathbf{v}^\top \mathbf{K}\mathbf{v} \\ &\overset{(d)}{=} \frac{1}{h} \mathbf{v}^\top \mathbf{D}^\top \mathbf{D}\mathbf{v} \overset{(e)}{=} \frac{1}{h} \left(\mathbf{D}\mathbf{v}\right)^\top \left(\mathbf{D}\mathbf{v}\right) \overset{(f)}{=} \frac{1}{h} \left\|\mathbf{D}\mathbf{v}\right\|_2^2. \end{aligned} \tag{137}$$

Here (a) uses the identity $\mathbf{v}^\top \mathbf{0} = 0$, (b) substitutes (136), (c) uses the associativity of matrix multiplication, (d) substitutes the factorization (134), (e) uses $\mathbf{v}^\top \mathbf{D}^\top = \left(\mathbf{D}\mathbf{v}\right)^\top$, and (f) applies the Euclidean identity $\mathbf{w}^\top \mathbf{w} = \|\mathbf{w}\|_2^2$ with $\mathbf{w} = \mathbf{D}\mathbf{v}$. Since $A > 0$ and $G \geq 1$, the mesh size $h = 2A/G$ satisfies $h > 0$. Multiplying (137) by h therefore yields

$$\left\|\mathbf{D}\mathbf{v}\right\|_2^2 = 0. \tag{138}$$

By the componentwise definition of the Euclidean norm,

$$0 \overset{(a)}{=} \left\|\mathbf{D}\mathbf{v}\right\|_2^2 \overset{(b)}{=} \sum_{i=0}^{G-1} \left(\left(\mathbf{D}\mathbf{v}\right)_i\right)^2. \tag{139}$$

Here (a) is (138), and (b) is the definition of the squared Euclidean norm on $\mathbb{R}^G$. Fix $i \in \{0, \dots, G-1\}$. Since every real square is nonnegative, we have

$$0 \overset{(a)}{\leq} ((\mathbf{D}\mathbf{v})_i)^2 \overset{(b)}{\leq} \sum_{r=0}^{G-1} ((\mathbf{D}\mathbf{v})_r)^2 \overset{(c)}{=} 0. \tag{140}$$

Here (a) uses the nonnegativity of the square of a real number, (b) uses the fact that the sum contains the displayed term and that all of its remaining terms are nonnegative, and (c) applies (139). The two inequalities in (140) force $((\mathbf{D}\mathbf{v})_i)^2 = 0$. A real number has square zero if and only if the number itself is zero, so

$$(\mathbf{D}\mathbf{v})_i = 0.$$

Since i was arbitrary, every component of $\mathbf{D}\mathbf{v}$ is zero. It follows that

$$\mathbf{D}\mathbf{v} = \mathbf{0}.$$

Thus $\mathbf{v} \in \ker(\mathbf{D})$, and consequently

$$\ker(\mathbf{K}) \subseteq \ker(\mathbf{D}). \tag{141}$$

Conversely, let $\mathbf{v} \in \ker(\mathbf{D})$ be arbitrary. Then

$$\mathbf{D}\mathbf{v} = \mathbf{0}. \tag{142}$$

Using (134), we obtain

$$\mathbf{K}\mathbf{v} \overset{(a)}{=} \frac{1}{h}\mathbf{D}^\top\mathbf{D}\mathbf{v} \overset{(b)}{=} \frac{1}{h}\mathbf{D}^\top(\mathbf{D}\mathbf{v}) \overset{(c)}{=} \frac{1}{h}\mathbf{D}^\top\mathbf{0} \overset{(d)}{=} \mathbf{0}.$$

Here (a) substitutes (134), (b) uses the associativity of matrix multiplication, (c) substitutes (142), and (d) uses $\mathbf{D}^\top\mathbf{0} = \mathbf{0}$ and $(1/h)\mathbf{0} = \mathbf{0}$. Therefore $\mathbf{v} \in \ker(\mathbf{K})$, and hence

$$\ker(\mathbf{D}) \subseteq \ker(\mathbf{K}). \tag{143}$$

Combining (141) and (143) proves (135).

It remains to characterize $\ker(\mathbf{D})$. Let $\mathbf{v} = (v_0, \dots, v_G)^\top \in \ker(\mathbf{D})$. Then

$$\mathbf{D}\mathbf{v} = \mathbf{0}. \tag{144}$$

For every $i \in \{0, \dots, G-1\}$, we therefore have

$$0 \overset{(a)}{=} (\mathbf{D}\mathbf{v})_i \overset{(b)}{=} v_{i+1} - v_i. \tag{145}$$

Here (a) takes the ith component of (144), and (b) uses the definition of the first-difference operator. Rearranging (145) gives

$$v_{i+1} = v_i, \qquad i = 0, \dots, G-1. \tag{146}$$

We now show by induction that

$$v_j = v_0, \qquad j = 0, \dots, G. \tag{147}$$

The statement is immediate for $j = 0$. Assume that it holds for some $j \in \{0, \dots, G-1\}$, so that $v_j = v_0$. Applying (146) with $i = j$ gives

$$v_{j+1} \overset{(a)}{=} v_j \overset{(b)}{=} v_0.$$

Here (a) is (146), and (b) is the induction hypothesis. Thus (147) holds for every $j = 0, \dots, G$. Setting $c := v_0$ and recalling that $\mathbf{1} = (1, \dots, 1)^\top \in \mathbb{R}^{G+1}$, we obtain

$$\mathbf{v} \overset{(a)}{=} \begin{pmatrix} v_0 \\ \vdots \\ v_G \end{pmatrix} \overset{(b)}{=} \begin{pmatrix} c \\ \vdots \\ c \end{pmatrix} \overset{(c)}{=} c\mathbf{1}.$$

Here (a) writes $\mathbf{v}$ componentwise, (b) uses (147), and (c) factors out the common scalar c. Therefore every vector in $\ker(\mathbf{D})$ belongs to $\mathrm{span}(\mathbf{1})$, and hence

$$\ker(\mathbf{D}) \subseteq \mathrm{span}(\mathbf{1}). \tag{148}$$

For the reverse inclusion, let $\mathbf{v} \in \mathrm{span}(\mathbf{1})$. By the definition of the linear span, there exists $c \in \mathbb{R}$ such that

$$\mathbf{v} = c\mathbf{1}. \tag{149}$$

For every $i \in \{0, \dots, G-1\}$, the corresponding component of $\mathbf{Dv}$ satisfies

$$(\mathbf{Dv})_i \overset{(a)}{=} v_{i+1} - v_i \overset{(b)}{=} c - c \overset{(c)}{=} 0.$$

Here (a) uses the definition of $\mathbf{D}$, (b) uses (149), and (c) simplifies the scalar difference. Thus every component of $\mathbf{Dv}$ is zero, so

$$\mathbf{Dv} = \mathbf{0}.$$

Therefore $\mathbf{v} \in \ker(\mathbf{D})$, which proves

$$\mathrm{span}(\mathbf{1}) \subseteq \ker(\mathbf{D}). \tag{150}$$

Combining (148) and (150) yields

$$\ker(\mathbf{D}) = \mathrm{span}(\mathbf{1}). \tag{151}$$

Finally, combining (135) and (151) gives

$$\ker(\mathbf{K}) = \ker(\mathbf{D}) = \mathrm{span}(\mathbf{1}),$$

which proves the claim.

□

Lemma D3 (Block decomposition)**.** *Assume that $G = 2m$, so that the anchor is located at $t_{i_0} = 0$ with $i_0 = m$, and use the reduced anchored ordering in (9). Then the anchored stiffness matrix satisfies*

$$\mathbf{K}_0 = \frac{1}{h}\begin{pmatrix} \mathbf{T}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_m \end{pmatrix}, \tag{152}$$

where $\mathbf{T}_m = (\tau_{rs})_{r,s=0}^{m-1} \in \mathbb{R}^{m\times m}$ is defined entrywise by

$$\tau_{rs} := \begin{cases} 1, & r = s = 0, \\ 2, & r = s \in \{1, \dots, m-1\}, \\ -1, & |r-s| = 1, \\ 0, & \textit{otherwise}. \end{cases} \tag{153}$$

Equivalently, for $m \geq 2$,

$$\mathbf{T}_m = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & 2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix},$$

whereas $\mathbf{T}_1 = (1)$.

Proof. Let $\widetilde{\mathbf{v}} \in \mathbb{R}^{2m}$ be arbitrary, and partition it as

$$\widetilde{\mathbf{v}} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}, \qquad \mathbf{a} = (a_0, \dots, a_{m-1})^\top, \qquad \mathbf{b} = (b_0, \dots, b_{m-1})^\top. \tag{154}$$

Set

$$\mathbf{v} := \mathbf{R}\widetilde{\mathbf{v}} \in \mathbb{R}^{2m+1}.$$

By the reduced-coordinate ordering in (9), the components of the three vectors satisfy

$$a_j = \widetilde{v}_j = v_j, \qquad b_j = \widetilde{v}_{m+j} = v_{2m-j}, \qquad j = 0, \dots, m-1, \tag{155}$$

and the inserted anchored component satisfies

$$v_m = 0. \tag{156}$$

Define $\mathbf{E}_m = (E_{rs})_{r,s=0}^{m-1} \in \mathbb{R}^{m\times m}$ entrywise by

$$E_{rs} := \begin{cases} -1, & s = r, \\ 1, & s = r+1 \text{ and } r \in \{0, \dots, m-2\}, \\ 0, & \text{otherwise}, \end{cases} \tag{157}$$

and let $\mathbf{J}_m = (J_{rs})_{r,s=0}^{m-1}$ denote the reversal permutation matrix defined by

$$J_{rs} := \begin{cases} 1, & r+s = m-1, \\ 0, & \text{otherwise}. \end{cases} \tag{158}$$

The definition of $\mathbf{E}_m$ gives, for every $\mathbf{x} = (x_0, \dots, x_{m-1})^\top \in \mathbb{R}^m$,

$$(\mathbf{E}_m\mathbf{x})_r = \begin{cases} x_{r+1} - x_r, & r = 0, \dots, m-2, \\ -x_{m-1}, & r = m-1. \end{cases} \tag{159}$$

We first compute the action of $\mathbf{DR}$ on the left half-grid. For $i = 0, \dots, m-1$,

$$\begin{aligned} (\mathbf{DR}\widetilde{\mathbf{v}})_i &\overset{(a)}{=} (\mathbf{Dv})_i \overset{(b)}{=} v_{i+1} - v_i \\ &\overset{(c)}{=} \begin{cases} a_{i+1} - a_i, & i = 0, \dots, m-2, \\ -a_{m-1}, & i = m-1, \end{cases} \overset{(d)}{=} (\mathbf{E}_m\mathbf{a})_i. \end{aligned} \tag{160}$$

Here (a) uses $\mathbf{v} = \mathbf{R}\widetilde{\mathbf{v}}$, (b) uses the componentwise definition of $\mathbf{D}$, (c) uses (155) and (156), and (d) follows from (159).

We next compute the right-half action. By (158), for every $s = 0, \dots, m-1$,

$$\begin{aligned} (-\mathbf{J}_m\mathbf{E}_m\mathbf{b})_s &\overset{(a)}{=} -(\mathbf{E}_m\mathbf{b})_{m-1-s} \\ &\overset{(b)}{=} \begin{cases} b_{m-1}, & s = 0, \\ b_{m-s-1} - b_{m-s}, & s = 1, \dots, m-1. \end{cases} \end{aligned} \tag{161}$$

Here (a) uses the fact that $\mathbf{J}_m$ reverses the order of the components, and (b) applies (159): when $s = 0$, the index is $m-1$, whereas when $s \geq 1$, the index $m-1-s$ belongs to $\{0, \dots, m-2\}$. On the other hand, for $s = 0, \dots, m-1$,

$$\begin{aligned} (\mathbf{DR}\widetilde{\mathbf{v}})_{m+s} &\overset{(a)}{=} (\mathbf{Dv})_{m+s} \overset{(b)}{=} v_{m+s+1} - v_{m+s} \\ &\overset{(c)}{=} \begin{cases} b_{m-1}, & s = 0, \\ b_{m-s-1} - b_{m-s}, & s = 1, \dots, m-1, \end{cases} \overset{(d)}{=} (-\mathbf{J}_m\mathbf{E}_m\mathbf{b})_s. \end{aligned} \tag{162}$$

Here (a) again uses $\mathbf{v} = \mathbf{R}\widetilde{\mathbf{v}}$, (b) uses the definition of $\mathbf{D}$, (c) uses $v_m = 0$ and the identity $v_{m+r} = b_{m-r}$ for $r = 1, \dots, m$, which follows from (155), and (d) invokes (161).

Combining (160) and (162) yields

$$\begin{aligned} \mathbf{DR}\widetilde{\mathbf{v}} &\overset{(a)}{=} \begin{pmatrix} \mathbf{E}_m\mathbf{a} \\ -\mathbf{J}_m\mathbf{E}_m\mathbf{b} \end{pmatrix} \overset{(b)}{=} \begin{pmatrix} \mathbf{E}_m & \mathbf{0} \\ \mathbf{0} & -\mathbf{J}_m\mathbf{E}_m \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \\ &\overset{(c)}{=} \begin{pmatrix} \mathbf{E}_m & \mathbf{0} \\ \mathbf{0} & -\mathbf{J}_m\mathbf{E}_m \end{pmatrix} \widetilde{\mathbf{v}}. \end{aligned} \tag{163}$$

Here (a) concatenates the first m components from (160) and the last m components from (162), (b) is block-matrix multiplication, and (c) uses (154). Because $\widetilde{\mathbf{v}} \in \mathbb{R}^{2m}$ was arbitrary, (163) implies the matrix identity

$$\mathbf{DR} = \begin{pmatrix} \mathbf{E}_m & \mathbf{0} \\ \mathbf{0} & -\mathbf{J}_m\mathbf{E}_m \end{pmatrix}. \tag{164}$$

The reversal matrix is orthogonal. Indeed, for every $r, s \in \{0, \dots, m-1\}$,

$$(\mathbf{J}_m^\top \mathbf{J}_m)_{rs} \overset{(a)}{=} \sum_{q=0}^{m-1} J_{qr} J_{qs} \overset{(b)}{=} \begin{cases} 1, & r = s, \\ 0, & r \neq s, \end{cases} \overset{(c)}{=} (\mathbf{I}_m)_{rs}.$$

Here (a) is the componentwise formula for a matrix product. For (b), the factor J_{qr} is nonzero only for $q = m-1-r$, whereas J_{qs} is nonzero only for $q = m-1-s$; these two row indices coincide exactly when $r = s$. Finally, (c) is the entrywise definition of the identity matrix. Since the entries agree for every r, s,

$$\mathbf{J}_m^\top \mathbf{J}_m = \mathbf{I}_m. \tag{165}$$

We now apply the stiffness factorization from Lemma D1. The anchored stiffness matrix satisfies

$$\begin{aligned} \mathbf{K}_0 &\overset{(a)}{=} \mathbf{R}^\top \mathbf{K} \mathbf{R} \overset{(b)}{=} \frac{1}{h} \mathbf{R}^\top \mathbf{D}^\top \mathbf{D} \mathbf{R} \overset{(c)}{=} \frac{1}{h} (\mathbf{DR})^\top (\mathbf{DR}) \\ &\overset{(d)}{=} \frac{1}{h} \begin{pmatrix} \mathbf{E}_m^\top & \mathbf{0} \\ \mathbf{0} & -\mathbf{E}_m^\top \mathbf{J}_m^\top \end{pmatrix} \begin{pmatrix} \mathbf{E}_m & \mathbf{0} \\ \mathbf{0} & -\mathbf{J}_m \mathbf{E}_m \end{pmatrix} \\ &\overset{(e)}{=} \frac{1}{h} \begin{pmatrix} \mathbf{E}_m^\top \mathbf{E}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_m^\top \mathbf{J}_m^\top \mathbf{J}_m \mathbf{E}_m \end{pmatrix} \overset{(f)}{=} \frac{1}{h} \begin{pmatrix} \mathbf{E}_m^\top \mathbf{E}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_m^\top \mathbf{E}_m \end{pmatrix}. \end{aligned} \tag{166}$$

Here (a) is the definition of $\mathbf{K}_0$, (b) substitutes $\mathbf{K} = \frac{1}{h}\mathbf{D}^\top \mathbf{D}$ from Lemma D1, (c) uses $(\mathbf{DR})^\top = \mathbf{R}^\top \mathbf{D}^\top$, (d) substitutes (164) and uses $(-\mathbf{J}_m \mathbf{E}_m)^\top = -\mathbf{E}_m^\top \mathbf{J}_m^\top$, (e) performs the block-matrix multiplication, and (f) uses (165).

It remains to identify $\mathbf{E}_m^\top \mathbf{E}_m$. For $r, s \in \{0, \dots, m-1\}$,

$$(\mathbf{E}_m^\top \mathbf{E}_m)_{rs} \overset{(a)}{=} \sum_{q=0}^{m-1} E_{qr} E_{qs} \overset{(b)}{=} \begin{cases} 1, & r = s = 0, \\ 2, & r = s \in \{1, \dots, m-1\}, \\ -1, & |r-s| = 1, \\ 0, & \text{otherwise}, \end{cases} \overset{(c)}{=} \tau_{rs}.$$

Here (a) is the componentwise formula for matrix multiplication. To justify (b), observe from (157) that column zero has the single nonzero entry $E_{00} = -1$, so its squared Euclidean norm is one. Every column $r \in \{1, \dots, m-1\}$ has exactly the two nonzero entries $E_{r-1,r} = 1$ and $E_{rr} = -1$, so its squared Euclidean norm is $1^2 + (-1)^2 = 2$. Two consecutive columns have exactly one common nonzero row: if $s = r+1$, their product in row r is $E_{rr} E_{r,r+1} = (-1)(1) = -1$, and the symmetric case $r = s+1$ is identical. Columns whose indices differ by at least two have no common nonzero row, so their inner product is zero. Finally, (c) uses the definition of τ_{rs} in (153). Since the entries agree for all r, s,

$$\mathbf{E}_m^\top \mathbf{E}_m = \mathbf{T}_m. \tag{167}$$

Substituting (167) into (166) gives

$$\mathbf{K}_0 = \frac{1}{h} \begin{pmatrix} \mathbf{T}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_m \end{pmatrix},$$

which is exactly (152) and proves the lemma.

□

Lemma D4 (Brownian energy in increment coordinates)**.** *Let $f \in \mathcal{H}_h^0$, and let $\widetilde{\mathbf{v}} \in \mathbb{R}^G$ be its reduced anchored nodal coefficient vector. Define the corresponding full anchored nodal vector and increment vector by*

$$\mathbf{v} := \mathbf{R}\widetilde{\mathbf{v}} \in \mathbb{R}^{G+1}, \tag{168}$$

$$\mathbf{w} := \mathbf{D}\mathbf{v} = \mathbf{DR}\widetilde{\mathbf{v}} = (w_0, \dots, w_{G-1})^\top \in \mathbb{R}^G, \tag{169}$$

so that $f = \mathcal{R}(\mathbf{v}) = \mathcal{R}(\mathbf{R}\widetilde{\mathbf{v}})$. *Then*

$$\|f\|_{B,h}^2 = \frac{1}{h}\|\mathbf{w}\|_2^2 = \frac{1}{h}\sum_{i=0}^{G-1} w_i^2. \tag{170}$$

Proof. By (168), $\mathbf{v} \in \mathbb{R}^{G+1}$, and by (169), $\mathbf{w} \in \mathbb{R}^G$. For every $i \in \{0, \dots, G-1\}$, the ith increment satisfies

$$w_i \overset{(a)}{=} (\mathbf{D}\mathbf{v})_i \overset{(b)}{=} v_{i+1} - v_i.$$

Here (a) takes the ith component of the defining identity $\mathbf{w} = \mathbf{D}\mathbf{v}$ in (169), and (b) uses the definition of the first-difference operator $\mathbf{D}$.

Since $f = \mathcal{R}(\mathbf{v})$, Proposition 1 gives both the quadratic-form representation of the Brownian energy and the factorization of the stiffness matrix. Therefore,

$$\begin{aligned}\|f\|_{B,h}^2 &\overset{(a)}{=} \mathbf{v}^\top \mathbf{K}\mathbf{v} \overset{(b)}{=} \mathbf{v}^\top \left(\frac{1}{h}\mathbf{D}^\top\mathbf{D}\right)\mathbf{v} \overset{(c)}{=} \frac{1}{h}\mathbf{v}^\top\mathbf{D}^\top\mathbf{D}\mathbf{v} \overset{(d)}{=} \frac{1}{h}(\mathbf{D}\mathbf{v})^\top(\mathbf{D}\mathbf{v}) \\ &\overset{(e)}{=} \frac{1}{h}\mathbf{w}^\top\mathbf{w} \overset{(f)}{=} \frac{1}{h}\|\mathbf{w}\|_2^2 \overset{(g)}{=} \frac{1}{h}\sum_{i=0}^{G-1} w_i^2.\end{aligned} \tag{171}$$

Here (a) is the Brownian-energy quadratic-form identity established in Proposition 1; (b) substitutes the stiffness factorization $\mathbf{K} = \frac{1}{h}\mathbf{D}^\top\mathbf{D}$ from the same proposition; (c) moves the scalar factor $1/h$ outside the matrix product; (d) uses $(\mathbf{D}\mathbf{v})^\top = \mathbf{v}^\top\mathbf{D}^\top$ and the associativity of matrix multiplication; (e) substitutes $\mathbf{w} = \mathbf{D}\mathbf{v}$ from (169); (f) uses the definition $\|\mathbf{x}\|_2^2 = \mathbf{x}^\top\mathbf{x}$ for vectors in Euclidean space; and (g) expands that norm componentwise using $\mathbf{w} = (w_0, \dots, w_{G-1})^\top$. Thus (171) is exactly (170), which proves the lemma.

□

Lemma D5 (Brownian energy in spectral coordinates)**.** *Let* $f \in \mathcal{H}_h^0$*, and let* $\widetilde{\mathbf{v}} \in \mathbb{R}^G$ *be its reduced anchored nodal coefficient vector, so that*

$$f = \mathcal{R}(\mathbf{R}\widetilde{\mathbf{v}}). \tag{172}$$

Then the anchored stiffness matrix

$$\mathbf{K}_0 = \mathbf{R}^\top\mathbf{K}\mathbf{R}$$

is symmetric positive definite. Consequently, it admits an orthogonal eigendecomposition

$$\mathbf{K}_0 = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top, \qquad \mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_G), \qquad \lambda_i > 0, \quad i = 1, \dots, G.$$

where $\mathbf{Q} \in \mathbb{R}^{G\times G}$ *is orthogonal. Define*

$$\mathbf{c} := \mathbf{Q}^\top\widetilde{\mathbf{v}} = (c_1, \dots, c_G)^\top. \tag{173}$$

Then

$$\|f\|_{B,h}^2 = \mathbf{c}^\top\mathbf{\Lambda}\mathbf{c} = \sum_{i=1}^{G}\lambda_i c_i^2. \tag{174}$$

D.1 Proof of Lemma D5

Proof. Set

$$\mathbf{v} := \mathbf{R}\widetilde{\mathbf{v}} \in \mathbb{R}^{G+1}. \tag{175}$$

By (172) and (175),

$$f = \mathcal{R}(\mathbf{v}).$$

We first verify that $\mathbf{K}_0$ is symmetric positive definite. Write the anchoring reconstruction matrix as

$$\mathbf{R} = [\mathbf{r}_1, \dots, \mathbf{r}_G],$$

where $\mathbf{r}_1, \dots, \mathbf{r}_G \in \mathbb{R}^{G+1}$ are its columns. By the explicit construction in (10), these columns are pairwise distinct canonical basis vectors of $\mathbb{R}^{G+1}$. Therefore, for every $r, s \in [G]$,

$$\left(\mathbf{R}^\top \mathbf{R}\right)_{rs} \overset{(a)}{=} \mathbf{r}_r^\top \mathbf{r}_s \overset{(b)}{=} \delta_{rs} \overset{(c)}{=} \left(\mathbf{I}_G\right)_{rs}. \tag{176}$$

Here (a) is the entrywise rule for the product $\mathbf{R}^\top \mathbf{R}$, (b) uses the orthonormality of pairwise distinct canonical basis vectors, and (c) is the entrywise definition of the identity matrix. Since (176) holds for every $r, s \in [G]$,

$$\mathbf{R}^\top \mathbf{R} = \mathbf{I}_G. \tag{177}$$

By Proposition 1,

$$\mathbf{K} = \frac{1}{h} \mathbf{D}^\top \mathbf{D}.$$

Hence

$$\mathbf{K}^\top \overset{(a)}{=} \left(\frac{1}{h} \mathbf{D}^\top \mathbf{D}\right)^\top \overset{(b)}{=} \frac{1}{h} \mathbf{D}^\top \left(\mathbf{D}^\top\right)^\top \overset{(c)}{=} \frac{1}{h} \mathbf{D}^\top \mathbf{D} \overset{(d)}{=} \mathbf{K}. \tag{178}$$

Here (a) substitutes the stiffness factorization, (b) uses $(\alpha \mathbf{A})^\top = \alpha \mathbf{A}^\top$ for real α and $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$, (c) uses $(\mathbf{D}^\top)^\top = \mathbf{D}$, and (d) substitutes the stiffness factorization again. It follows that

$$\mathbf{K}_0^\top \overset{(a)}{=} \left(\mathbf{R}^\top \mathbf{K} \mathbf{R}\right)^\top \overset{(b)}{=} \mathbf{R}^\top \mathbf{K}^\top \mathbf{R} \overset{(c)}{=} \mathbf{R}^\top \mathbf{K} \mathbf{R} \overset{(d)}{=} \mathbf{K}_0. \tag{179}$$

Here (a) uses the definition of $\mathbf{K}_0$, (b) applies the transpose-of-a-product rule, (c) uses (178), and (d) again uses the definition of $\mathbf{K}_0$. Thus $\mathbf{K}_0$ is symmetric.

To prove positive definiteness, let $\mathbf{z} \in \mathbb{R}^G \setminus \{\mathbf{0}\}$ be arbitrary and define

$$\mathbf{y} := \mathbf{R}\mathbf{z} \in \mathbb{R}^{G+1}. \tag{180}$$

Using (177), we obtain

$$\begin{aligned} \|\mathbf{y}\|_2^2 &\overset{(a)}{=} (\mathbf{R}\mathbf{z})^\top (\mathbf{R}\mathbf{z}) \overset{(b)}{=} \mathbf{z}^\top \mathbf{R}^\top \mathbf{R} \mathbf{z} \\ &\overset{(c)}{=} \mathbf{z}^\top \mathbf{I}_G \mathbf{z} \overset{(d)}{=} \|\mathbf{z}\|_2^2 \overset{(e)}{>} 0. \end{aligned}$$

Here (a) uses $\|\mathbf{y}\|_2^2 = \mathbf{y}^\top \mathbf{y}$ and (180), (b) uses $(\mathbf{R}\mathbf{z})^\top = \mathbf{z}^\top \mathbf{R}^\top$ and associativity, (c) substitutes (177), (d) uses $\mathbf{I}_G \mathbf{z} = \mathbf{z}$ and the definition of the Euclidean norm, and (e) follows from $\mathbf{z} \neq \mathbf{0}$. Consequently,

$$\mathbf{y} \neq \mathbf{0}. \tag{181}$$

Moreover, the construction of $\mathbf{R}$ inserts a zero at the anchor index, and therefore

$$y_{i_0} = 0. \tag{182}$$

The quadratic form of $\mathbf{K}_0$ at $\mathbf{z}$ satisfies

$$\begin{aligned} \mathbf{z}^\top \mathbf{K}_0 \mathbf{z} &\overset{(a)}{=} \mathbf{z}^\top \mathbf{R}^\top \mathbf{K} \mathbf{R} \mathbf{z} \overset{(b)}{=} (\mathbf{R}\mathbf{z})^\top \mathbf{K} (\mathbf{R}\mathbf{z}) \\ &\overset{(c)}{=} \mathbf{y}^\top \mathbf{K} \mathbf{y} \overset{(d)}{=} \frac{1}{h} \|\mathbf{D}\mathbf{y}\|_2^2 \overset{(e)}{\geq} 0. \end{aligned} \tag{183}$$

Here (a) substitutes $\mathbf{K}_0 = \mathbf{R}^\top \mathbf{K} \mathbf{R}$, (b) uses $\mathbf{z}^\top \mathbf{R}^\top = (\mathbf{R}\mathbf{z})^\top$ and associativity, (c) substitutes (180), (d) applies Proposition 1 to $\mathbf{y}$, and (e) uses $h > 0$ and the nonnegativity of a squared Euclidean norm.

Suppose, for contradiction, that

$$\mathbf{z}^\top \mathbf{K}_0 \mathbf{z} = 0. \tag{184}$$

Combining (183) and (184) gives

$$0 \overset{(a)}{=} \frac{1}{h}\,\|\mathbf{D}\mathbf{y}\|_2^2 \overset{(b)}{\Longrightarrow} \|\mathbf{D}\mathbf{y}\|_2^2 = 0 \overset{(c)}{\Longrightarrow} \mathbf{D}\mathbf{y} = \mathbf{0}. \tag{185}$$

Here (a) uses (183) with equality as imposed by (184), (b) multiplies by the positive scalar h, and (c) uses the fact that a vector has zero squared Euclidean norm if and only if it is the zero vector. By Lemma D2, (185) implies that there exists $a \in \mathbb{R}$ such that

$$\mathbf{y} = a\mathbf{1}. \tag{186}$$

Taking the anchor component in (186) and using (182) yields

$$0 \overset{(a)}{=} y_{i_0} \overset{(b)}{=} a,$$

where (a) is (182) and (b) takes the i_0th component of (186). Thus $a = 0$, and (186) gives $\mathbf{y} = \mathbf{0}$, contradicting (181). Therefore

$$\mathbf{z}^\top \mathbf{K}_0 \mathbf{z} > 0 \qquad \text{for every } \mathbf{z} \in \mathbb{R}^G \setminus \{\mathbf{0}\}. \tag{187}$$

Together with (179), this proves that $\mathbf{K}_0$ is symmetric positive definite.

By the spectral theorem for real symmetric matrices, there exist an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{G\times G}$ and a real diagonal matrix $\mathbf{\Lambda}$ such that

$$\mathbf{K}_0 = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top.$$

The positive definiteness in (187) implies that every diagonal entry λ_i of $\mathbf{\Lambda}$ is strictly positive. Indeed, if $\mathbf{q}_i$ denotes the ith column of $\mathbf{Q}$, then $\|\mathbf{q}_i\|_2 = 1$ and $\mathbf{K}_0\mathbf{q}_i = \lambda_i\mathbf{q}_i$, so

$$\lambda_i \overset{(a)}{=} \lambda_i \mathbf{q}_i^\top \mathbf{q}_i \overset{(b)}{=} \mathbf{q}_i^\top \mathbf{K}_0 \mathbf{q}_i \overset{(c)}{>} 0.$$

Here (a) uses $\mathbf{q}_i^\top\mathbf{q}_i = 1$, (b) uses $\mathbf{K}_0\mathbf{q}_i = \lambda_i\mathbf{q}_i$, and (c) applies (187) to the nonzero vector $\mathbf{q}_i$.

We now derive the Brownian energy identity. By (131), (175), and the definition of $\mathbf{K}_0$,

$$\begin{aligned}
\|f\|_{B,h}^2 &\overset{(a)}{=} \mathbf{v}^\top \mathbf{K} \mathbf{v} \overset{(b)}{=} (\mathbf{R}\widetilde{\mathbf{v}})^\top \mathbf{K}\,(\mathbf{R}\widetilde{\mathbf{v}}) \\
&\overset{(c)}{=} \widetilde{\mathbf{v}}^\top \mathbf{R}^\top \mathbf{K}\mathbf{R}\widetilde{\mathbf{v}} \overset{(d)}{=} \widetilde{\mathbf{v}}^\top \mathbf{K}_0 \widetilde{\mathbf{v}} \\
&\overset{(e)}{=} \widetilde{\mathbf{v}}^\top \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top \widetilde{\mathbf{v}} \\
&\overset{(f)}{=} \left(\mathbf{Q}^\top\widetilde{\mathbf{v}}\right)^\top \mathbf{\Lambda}\left(\mathbf{Q}^\top\widetilde{\mathbf{v}}\right) \overset{(g)}{=} \mathbf{c}^\top\mathbf{\Lambda}\mathbf{c}.
\end{aligned} \tag{188}$$

Here (a) applies (131) to $f = \mathcal{R}(\mathbf{v})$, (b) substitutes (175), (c) uses $(\mathbf{R}\widetilde{\mathbf{v}})^\top = \widetilde{\mathbf{v}}^\top\mathbf{R}^\top$ and associativity, (d) uses $\mathbf{K}_0 = \mathbf{R}^\top\mathbf{K}\mathbf{R}$, (e) substitutes the orthogonal eigendecomposition of $\mathbf{K}_0$, (f) uses $\widetilde{\mathbf{v}}^\top\mathbf{Q} = (\mathbf{Q}^\top\widetilde{\mathbf{v}})^\top$, and (g) substitutes (173).

Because $\mathbf{\Lambda}$ is diagonal,

$$\Lambda_{ij} = \lambda_i\delta_{ij}, \qquad i,j \in [G]. \tag{189}$$

Therefore

$$\begin{aligned}
\mathbf{c}^\top\mathbf{\Lambda}\mathbf{c} &\overset{(a)}{=} \sum_{i=1}^{G}\sum_{j=1}^{G} c_i\Lambda_{ij}c_j \overset{(b)}{=} \sum_{i=1}^{G}\sum_{j=1}^{G} c_i\lambda_i\delta_{ij}c_j \\
&\overset{(c)}{=} \sum_{i=1}^{G} c_i\lambda_i c_i \overset{(d)}{=} \sum_{i=1}^{G}\lambda_i c_i^2.
\end{aligned} \tag{190}$$

Here (a) expands the quadratic form componentwise, (b) substitutes (189), (c) uses $\delta_{ij} = 0$ for $i \neq j$ and $\delta_{ii} = 1$ to collapse the inner sum to its $j = i$ term, and (d) rearranges the scalar factors and uses $c_ic_i = c_i^2$. Combining (188) and (190) yields (174), which proves the lemma.

□

Lemma D6 (Explicit spectral Brownian basis)**.** *Assume that $G = 2m$, and let*

$$\mathbf{K}_0 = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top, \qquad \mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_G) \tag{191}$$

be the orthogonal eigendecomposition obtained in Theorem 3. Thus $\mathbf{Q} \in \mathbb{R}^{G\times G}$ is orthogonal and $\lambda_i > 0$ for every $i \in [G]$. Let $\mathbf{e}_i \in \mathbb{R}^G$ denote the ith canonical basis vector. For every $i \in [G]$, define

$$\begin{aligned}\psi_i &:= \mathcal{R}(\mathbf{R}\mathbf{Q}\mathbf{e}_i) \\ &= \sum_{j=0}^{G} (\mathbf{R}\mathbf{Q}\mathbf{e}_i)_j \, \phi_j = \sum_{j=0}^{G} (\mathbf{R}\mathbf{Q})_{ji} \, \phi_j.\end{aligned} \tag{192}$$

Then the following statements hold.

(i) Each function ψ_i belongs to $\mathcal{H}_h^0$, and the family $\{\psi_1, \dots, \psi_G\}$ forms a basis of $\mathcal{H}_h^0$.

(ii) Let $f \in \mathcal{H}_h^0$, and let $\widetilde{\mathbf{v}} \in \mathbb{R}^G$ be its reduced anchored nodal coefficient vector. Then f admits the unique representation

$$f = \sum_{i=1}^{G} c_i \psi_i, \tag{193}$$

where

$$\mathbf{c} := (c_1, \dots, c_G)^\top = \mathbf{Q}^\top \widetilde{\mathbf{v}} \in \mathbb{R}^G \tag{194}$$

is the spectral coefficient vector of f.

(iii) The basis is orthogonal with respect to the Brownian inner product. More precisely, for every $i, j \in [G]$,

$$\langle \psi_i, \psi_j \rangle_{B,h} = \lambda_i \delta_{ij}. \tag{195}$$

Proof. We first verify that every function defined in (192) belongs to the anchored space. Fix $i \in [G]$. Because ψ_i is a finite linear combination of $\phi_0, \dots, \phi_G$, we have $\psi_i \in \mathcal{H}_h$. Moreover, the anchor is the grid point $t_{i_0} = 0$, and the nodal basis satisfies $\phi_j(0) = \delta_{ji_0}$. Therefore,

$$\begin{aligned}\psi_i(0) &\overset{(a)}{=} \sum_{j=0}^{G} (\mathbf{R}\mathbf{Q})_{ji} \, \phi_j(0) \overset{(b)}{=} \sum_{j=0}^{G} (\mathbf{R}\mathbf{Q})_{ji} \, \delta_{ji_0} \\ &\overset{(c)}{=} (\mathbf{R}\mathbf{Q})_{i_0 i} \overset{(d)}{=} \sum_{r=1}^{G} R_{i_0 r} Q_{ri} \overset{(e)}{=} 0.\end{aligned}$$

Here (a) uses the definition of ψ_i in (192); (b) substitutes $\phi_j(0) = \delta_{ji_0}$; (c) uses the defining property of the Kronecker delta to retain only the term $j = i_0$; (d) expands the (i_0, i) entry of the matrix product $\mathbf{R}\mathbf{Q}$; and (e) uses the construction of the anchoring reconstruction matrix $\mathbf{R}$, whose i_0th row is the zero row because the anchor coefficient is fixed at zero. Thus $\psi_i(0) = 0$, and hence $\psi_i \in \mathcal{H}_h^0$. Since $i \in [G]$ was arbitrary, this conclusion holds for every member of the family.

We next prove existence and uniqueness of the spectral expansion. Let $f \in \mathcal{H}_h^0$ be arbitrary. By the reduced anchored nodal representation, there exists a unique vector $\widetilde{\mathbf{v}} \in \mathbb{R}^G$ such that

$$f = \mathcal{R}(\mathbf{R}\widetilde{\mathbf{v}}). \tag{196}$$

Define $\mathbf{c}$ by (194). Because $\mathbf{Q}$ is orthogonal,

$$\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}_G, \qquad \mathbf{Q}\mathbf{Q}^\top = \mathbf{I}_G. \tag{197}$$

Consequently,

$$\mathbf{Q}\mathbf{c} \overset{(a)}{=} \mathbf{Q}\mathbf{Q}^\top \widetilde{\mathbf{v}} \overset{(b)}{=} \mathbf{I}_G \widetilde{\mathbf{v}} \overset{(c)}{=} \widetilde{\mathbf{v}}. \tag{198}$$

Here (a) substitutes the definition $\mathbf{c} = \mathbf{Q}^\top \widetilde{\mathbf{v}}$; (b) uses $\mathbf{Q}\mathbf{Q}^\top = \mathbf{I}_G$ from (197); and (c) uses the defining action of the identity matrix.

Writing $\mathbf{c} = \sum_{i=1}^{G} c_i \mathbf{e}_i$ and using (196) and (198), we obtain

$$\begin{aligned} f &\overset{(a)}{=} \mathcal{R}\left(\mathbf{R}\widetilde{\mathbf{v}}\right) \overset{(b)}{=} \mathcal{R}\left(\mathbf{R}\mathbf{Q}\mathbf{c}\right) \\ &\overset{(c)}{=} \mathcal{R}\left(\mathbf{R}\mathbf{Q}\sum_{i=1}^{G} c_i \mathbf{e}_i\right) \overset{(d)}{=} \mathcal{R}\left(\sum_{i=1}^{G} c_i \mathbf{R}\mathbf{Q}\mathbf{e}_i\right) \\ &\overset{(e)}{=} \sum_{i=1}^{G} c_i \mathcal{R}\left(\mathbf{R}\mathbf{Q}\mathbf{e}_i\right) \overset{(f)}{=} \sum_{i=1}^{G} c_i \psi_i. \end{aligned} \tag{199}$$

Here (a) is (196); (b) substitutes $\widetilde{\mathbf{v}} = \mathbf{Q}\mathbf{c}$ from (198); (c) expands $\mathbf{c}$ in the canonical basis of $\mathbb{R}^G$; (d) uses the linearity of the matrix maps $\mathbf{Q}$ and $\mathbf{R}$; (e) uses the linearity of the finite-element reconstruction operator $\mathcal{R}$; and (f) applies the definition of ψ_i in (192). This proves the existence of the representation (193).

To prove uniqueness, suppose that another vector

$$\mathbf{d} := (d_1, \ldots, d_G)^\top \in \mathbb{R}^G$$

satisfies

$$f = \sum_{i=1}^{G} d_i \psi_i. \tag{200}$$

Repeating the linearity calculation in (199) with $\mathbf{d}$ in place of $\mathbf{c}$ gives

$$\sum_{i=1}^{G} d_i \psi_i = \mathcal{R}\left(\mathbf{R}\mathbf{Q}\mathbf{d}\right). \tag{201}$$

Combining (196), (200), and (201) yields

$$\mathcal{R}\left(\mathbf{R}\widetilde{\mathbf{v}}\right) = \mathcal{R}\left(\mathbf{R}\mathbf{Q}\mathbf{d}\right).$$

The nodal representation in $\mathcal{H}_h$ is unique. Therefore,

$$\mathbf{R}\widetilde{\mathbf{v}} = \mathbf{R}\mathbf{Q}\mathbf{d}. \tag{202}$$

By (177), $\mathbf{R}^\top\mathbf{R} = \mathbf{I}_G$. Left-multiplying (202) by $\mathbf{R}^\top$ gives

$$\begin{aligned} \widetilde{\mathbf{v}} &\overset{(a)}{=} \mathbf{I}_G\widetilde{\mathbf{v}} \overset{(b)}{=} \mathbf{R}^\top\mathbf{R}\widetilde{\mathbf{v}} \overset{(c)}{=} \mathbf{R}^\top\mathbf{R}\mathbf{Q}\mathbf{d} \\ &\overset{(d)}{=} \mathbf{I}_G\mathbf{Q}\mathbf{d} \overset{(e)}{=} \mathbf{Q}\mathbf{d}. \end{aligned} \tag{203}$$

Here (a) uses the action of the identity matrix; (b) substitutes $\mathbf{R}^\top\mathbf{R} = \mathbf{I}_G$; (c) left-multiplies (202) by $\mathbf{R}^\top$; (d) again uses $\mathbf{R}^\top\mathbf{R} = \mathbf{I}_G$; and (e) uses $\mathbf{I}_G\mathbf{Q}\mathbf{d} = \mathbf{Q}\mathbf{d}$. It follows that

$$\mathbf{d} \overset{(a)}{=} \mathbf{I}_G\mathbf{d} \overset{(b)}{=} \mathbf{Q}^\top\mathbf{Q}\mathbf{d} \overset{(c)}{=} \mathbf{Q}^\top\widetilde{\mathbf{v}} \overset{(d)}{=} \mathbf{c}. \tag{204}$$

Here (a) uses the action of the identity matrix; (b) uses $\mathbf{Q}^\top\mathbf{Q} = \mathbf{I}_G$; (c) substitutes (203); and (d) uses the definition of $\mathbf{c}$ in (194). Thus the coefficients in (193) are unique, proving (ii).

The existence of the expansion for every $f \in \mathcal{H}_h^0$ shows that $\{\psi_1, \ldots, \psi_G\}$ spans $\mathcal{H}_h^0$. To verify linear independence, suppose $\sum_{i=1}^{G} d_i\psi_i = 0$. The zero function has reduced anchored nodal coefficient vector $\widetilde{\mathbf{v}} = \mathbf{0}$ and hence, by (194), spectral coefficient vector $\mathbf{Q}^\top\mathbf{0} = \mathbf{0}$. The uniqueness established in (204) therefore gives $\mathbf{d} = \mathbf{0}$. Thus the family is linearly independent. Since it both spans $\mathcal{H}_h^0$ and is linearly independent, it is a basis, proving (i).

It remains to prove the Brownian orthogonality relation. For each $i \in [G]$, set

$$\mathbf{a}_i := \mathbf{R}\mathbf{Q}\mathbf{e}_i = (a_{0i}, \ldots, a_{Gi})^\top \in \mathbb{R}^{G+1}. \tag{205}$$

Then $\psi_i = \mathcal{R}(\mathbf{a}_i)$. Fix arbitrary $i, j \in [G]$. Using the definition of the Brownian inner product and the entrywise definition of the stiffness matrix in Proposition 1, we obtain

$$\begin{aligned}
\langle \psi_i, \psi_j \rangle_{B,h} &\overset{(a)}{=} \int_{-A}^{A} \psi_i'(t)\psi_j'(t)\,\mathrm{d}t \overset{(b)}{=} \int_{-A}^{A} \left(\sum_{r=0}^{G} a_{ri}\phi_r'(t)\right)\left(\sum_{s=0}^{G} a_{sj}\phi_s'(t)\right)\,\mathrm{d}t \\
&\overset{(c)}{=} \sum_{r=0}^{G}\sum_{s=0}^{G} a_{ri}a_{sj}\int_{-A}^{A} \phi_r'(t)\phi_s'(t)\,\mathrm{d}t \overset{(d)}{=} \sum_{r=0}^{G}\sum_{s=0}^{G} a_{ri}K_{rs}a_{sj} \overset{(e)}{=} \mathbf{a}_i^\top \mathbf{K}\mathbf{a}_j \\
&\overset{(f)}{=} (\mathbf{R}\mathbf{Q}\mathbf{e}_i)^\top \mathbf{K}\,(\mathbf{R}\mathbf{Q}\mathbf{e}_j) \overset{(g)}{=} \mathbf{e}_i^\top \mathbf{Q}^\top\mathbf{R}^\top\mathbf{K}\mathbf{R}\mathbf{Q}\mathbf{e}_j \overset{(h)}{=} \mathbf{e}_i^\top \mathbf{Q}^\top\mathbf{K}_0\mathbf{Q}\mathbf{e}_j \\
&\overset{(i)}{=} \mathbf{e}_i^\top \mathbf{Q}^\top\left(\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top\right)\mathbf{Q}\mathbf{e}_j \overset{(j)}{=} \mathbf{e}_i^\top\left(\mathbf{Q}^\top\mathbf{Q}\right)\mathbf{\Lambda}\left(\mathbf{Q}^\top\mathbf{Q}\right)\mathbf{e}_j \overset{(k)}{=} \mathbf{e}_i^\top\mathbf{\Lambda}\mathbf{e}_j \\
&\overset{(l)}{=} \Lambda_{ij} \overset{(m)}{=} \lambda_i\delta_{ij}.
\end{aligned}$$

Here (a) is the definition of the Brownian inner product; (b) uses $\psi_i = \sum_{r=0}^{G} a_{ri}\phi_r$ and $\psi_j = \sum_{s=0}^{G} a_{sj}\phi_s$, and differentiates these finite sums on each open mesh interval; (c) expands the product and uses the linearity of the integral for finite sums; (d) uses $K_{rs} = \int_{-A}^{A} \phi_r'(t)\phi_s'(t)\,\mathrm{d}t$; (e) is the componentwise formula for the bilinear matrix product $\mathbf{a}_i^\top\mathbf{K}\mathbf{a}_j$; (f) substitutes (205); (g) uses $(\mathbf{R}\mathbf{Q}\mathbf{e}_i)^\top = \mathbf{e}_i^\top\mathbf{Q}^\top\mathbf{R}^\top$ and associativity; (h) substitutes $\mathbf{K}_0 = \mathbf{R}^\top\mathbf{K}\mathbf{R}$; (i) substitutes the eigendecomposition (191); (j) uses associativity of matrix multiplication; (k) uses $\mathbf{Q}^\top\mathbf{Q} = \mathbf{I}_G$ and the action of the identity matrix; (l) uses the defining property of canonical basis vectors, $\mathbf{e}_i^\top\mathbf{\Lambda}\mathbf{e}_j = \Lambda_{ij}$; and (m) uses the diagonal structure $\Lambda_{ij} = \lambda_i\delta_{ij}$. This proves (195) and hence (iii). In particular, if $i \neq j$, then $\langle \psi_i, \psi_j \rangle_{B,h} = 0$, whereas if $i = j$, then

$$\|\psi_i\|_{B,h}^2 = \langle \psi_i, \psi_i \rangle_{B,h} = \lambda_i > 0,$$

because $\lambda_i > 0$. This completes the proof.

□

E Implications of the Equivalent Coordinate Representations

Function-space invariance. Theorem 2 establishes that the nodal, increment, and spectral parameterizations are mathematically equivalent realizations of the same finite Brownian RKHS. Consequently, the represented function space, the Brownian norm, and the associated approximation properties are intrinsic to the underlying RKHS and remain invariant under changes of coordinate realization.

Controlled optimization. Because all three realizations represent exactly the same hypothesis class, differences cannot be attributed to representational capacity, approximation error, or intrinsic Brownian regularization. The sharper distinction is optimizer-specific: Proposition 3 forces mapped nodal and spectral GD/SGD trajectories to agree, Proposition 3 identifies increment GD with Brownian preconditioning, and Proposition 5 permits standard Adam to distinguish the DCT-VIII basis. The direct one-layer and recursive tests in Section 7 numerically verify these predictions under mapped initialization and shared update protocols.

Connection to the spectral realization. The equivalence theorem is independent of the particular choice of coordinates. The explicit spectral realization presented in Section 5 constructs a block-structured orthogonal basis of the anchored finite Brownian RKHS through the complete eigendecomposition of the anchored stiffness matrix, providing the spectral implementation used in the optimization study.

F Implications of the Explicit Spectral Representation

Explicit block-structured spectral coordinates. Unlike the nodal and increment parameterizations, the spectral coordinates diagonalize the Brownian energy. The energy determines the

eigenspaces, but the twofold multiplicity of every eigenvalue does not determine a unique eigenbasis inside each left–right eigenspace. The displayed block-diagonal eigendecomposition fixes an explicit orthogonal coordinate system by respecting the reduced-coordinate ordering and the two half-grid blocks.

Diagonalization of the Brownian energy. The explicit spectral basis diagonalizes the Brownian inner product, reducing the Brownian norm to a weighted Euclidean norm of the spectral coefficients. Consequently, each spectral mode contributes independently to the total Brownian energy, providing a natural decomposition of the finite Brownian RKHS into orthogonal Brownian spectral modes.

Implications for optimization. The spectral transformation is orthogonal, so mapped nodal–spectral GD/SGD must agree after coordinates, initialization, scalar steps, and minibatches are mapped consistently. Standard Adam is different: its universal orthogonal equivariance group is only the signed permutations, and the DCT-VIII matrix is not one. The controlled tests in Section 7 numerically verify both the exact GD/SGD identity and the permitted Adam separation.

G Supplementary Experimental Evidence

This appendix separates direct theorem-verification protocols from secondary predictive evidence. The old synthetic tables and legacy coordinate curves are omitted because their raw outputs mixed incompatible sample sizes or lacked the mapped-initialization and optimizer metadata required for a valid equivariance test. The retained EuroSAT and spatially separated Salinas studies were regenerated from frozen pipelines and independently audited from saved predictions and probabilities.

G.1 Theory-Validation Protocols and Reproducibility Audit

Evidence classes. The direct tests E0, E1A–E1C, and E2 numerically verify algebraic identities, mapped optimizer trajectories, conditioning statements, and the theorem-specific standard-Adam counterexample. They are not tuned predictive comparisons. The EuroSAT and spatially separated Salinas studies instead compare different model classes and are retained only as secondary predictive and robustness evidence; they are not used to establish coordinate equivariance.

Table 5: Frozen verification and audit protocols. Identity-test arms share one mapped represented initialization, objective normalization, scalar step sequence, and, for SGD, minibatch order; they are never tuned separately.

Stage	Purpose	Data/runs	Mesh	Budget	Precision
E0	Finite Brownian RKHS, spectrum, and coordinate identities	125 random vectors	$G \in \{8, 16, 32, 64, 128\}$	4097-point dense grid	float64 CPU
E1A	One-layer mapped-trajectory identities for GD and SGD	256 samples, 5 seeds	$G \in \{8, 16, 32, 64, 128\}$	40 GD; 80 SGD; batch 32	float64 CPU
E1B	Conditioning and fixed-step convergence under mesh refinement	256 samples, 5 seeds	$G \in \{8, 16, 32, 64, 128\}$	gap 10^{-8}; analytic steps	float64 CPU
E1C	Basis dependence of standard Adam	256 samples, 5 seeds	$G \in \{8, 16, 32, 64, 128\}$	200 steps; full batch	float64 CPU
E2	Recursive blockwise mapped-trajectory identities	128 samples, 5 seeds	$(8, 16, 32)$ and $(32, 64, 128)$	8 GD; 16 SGD; batch 32	float64 CPU
Real-data audit	Independent metric and statistical reconstruction	390 prediction artifacts	EuroSAT and Salinas	5720 published-value checks	NumPy reconstruction

E0: finite Brownian RKHS identities. We use $A = 2$, $G \in \{8, 16, 32, 64, 128\}$, seeds $\{0, 1, 2, 3, 4\}$, and five random vectors per seed. All matrix, spectrum, Brownian-Gram, reconstruction, and energy diagnostics are evaluated in float64 on CPU; reconstruction is checked on 4097 points.

E1A: one-layer mapped-trajectory identities for GD and SGD. The Brownian-regularized least-squares objective uses $n = 256$, $\rho = 0.1$, and noise standard deviation 0.03. Every coordinate arm begins from one reduced-nodal function and is mapped exactly. Full-batch GD uses 40 updates; SGD uses 80 updates with batch size 32 and an identical minibatch sequence. No coordinate arm is tuned separately.

E1B: conditioning and convergence under mesh refinement. The pure Brownian Hessians are constructed directly in nodal, spectral, and increment coordinates. For least squares, $n = 256$ and $\rho \in \{0.05, 0.1, 0.2, 0.5\}$. The convergence diagnostic uses $\rho = 0.1$, the analytic optimal scalar step

for each SPD quadratic, a relative objective-gap tolerance of 10^{-8}, and at most 200,000 iterations. These steps are analytic, not validation-selected.

E1C: basis dependence of standard Adam. We use scalar $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$, zero initial moment states, bias correction, learning rate 0.003 for the regularized objective, and 0.01 for the deterministic linear counterexample. The experiment uses 200 full-batch updates, no AdamW, weight decay, clipping, or AMSGrad, and no coordinate-specific tuning.

E2: recursive blockwise mapped-trajectory identities. The recursive model contains three trainable Brownian profile blocks with grid sets $(8, 16, 32)$ and $(32, 64, 128)$, trainable cross-layer couplings, and identically initialized non-profile parameters. The objective uses $n = 128$ and $\rho = 0.05$. We run eight full-batch steps at $\eta = 5 \times 10^{-5}$ and sixteen shared-minibatch steps at $\eta = 2 \times 10^{-5}$ with batch size 32. Each increment block is compared with the nodal Brownian-gradient update scaled by its own $1/h_b$.

Independent real-data audit. For the retained EuroSAT and spatially separated Salinas studies, accuracy, multiclass F_1, calibration, and likelihood metrics were reconstructed directly from 390 saved prediction artifacts, including 5,820 per-class records and 88,920 confusion-matrix cells. A separate implementation reproduced 5,720 reported AUC, exact-test, crossed-bootstrap, and variance-component entries across 11 tables with zero mismatch and maximum discrepancy 8.66×10^{-15}.

G.2 EuroSAT Under Test-Time Spectral-Band Dropout

Protocol. Patch MLP and Spatial Path-Atomic BKL receive the same Spatial65 representation of all 13 EuroSAT bands. For each of five paired seeds, both models are trained once on the same clean split. At test time only, one to six bands are zeroed before feature extraction; masks are nested within each seed and shared between models. For a metric M, retention is $M(d)/M(0)$, and the normalized area under the retention curve summarizes the complete degradation path. We additionally record ECE, NLL, per-class F_1, inference time, and parameter count.

Table 6: EuroSAT robustness under nested test-time spectral-band dropout. Values are means over five paired seeds. Retention AUC is normalized over zero to six zeroed bands; higher is better. Maximum ECE is lower-is-better.

Model	Accuracy Retention AUC (%)	Macro-F_1 Retention AUC (%)	Worst Accuracy Retention (%)	Worst Macro-F_1 Retention (%)	Maximum ECE
Patch MLP	37.78	32.16	21.80	13.88	0.774
Spatial Path-Atomic BKL	**46.81**	**42.10**	**23.38**	**17.68**	**0.629**

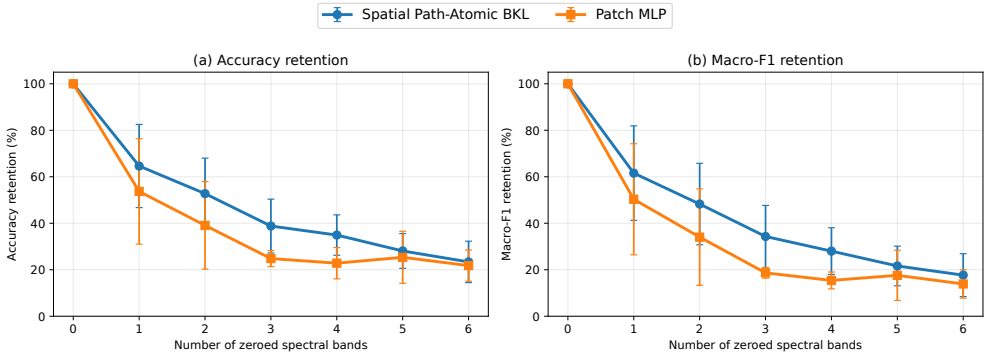

Figure 3: EuroSAT accuracy- and macro-F_1-retention trajectories under nested test-time spectral-band dropout. Lines show paired-seed means and shading one standard deviation.

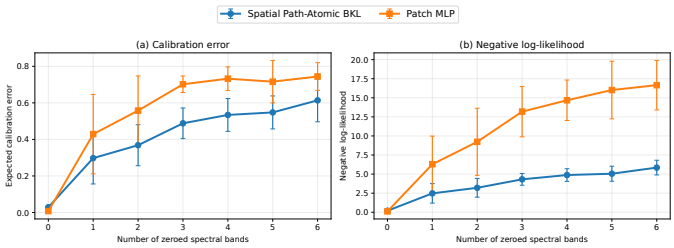

Figure 4: EuroSAT calibration and likelihood degradation under nested spectral-band dropout; lower is better.

Results and scope. Spatial Path-Atomic BKL has normalized accuracy-retention AUC 46.81% versus 37.78% for Patch MLP and macro-F_1-retention AUC 42.10% versus 32.16%. All five pairs favor BKL for both AUCs and for maximum-ECE reduction; each exact two-sided sign-flip test gives $p = 0.0625$, the minimum nonzero value for five pairs. The result concerns the full degradation trajectory rather than uniform superiority at each severity. This image-level multispectral experiment is a BKL-versus-MLP comparison and is not used to establish coordinate equivariance.

Table 7: Exact paired sign-flip analysis over five seeds. Improvements are oriented so that positive values favor Spatial Path-Atomic BKL.

Metric	Mean improvement	Two-sided p	Paired d_z	BKL-favored seeds
Accuracy retention AUC	0.0903	0.0625	1.61	5/5
Macro-F_1 retention AUC	0.0994	0.0625	2.11	5/5
Worst accuracy retention	0.0158	0.6875	0.17	3/5
Worst Macro-F_1 retention	0.0379	0.4375	0.46	3/5
Maximum ECE reduction	0.1449	0.0625	1.53	5/5

G.3 Salinas Evaluation Under Spatial Separation and Missing Spectral Bands

Protocol. We evaluate the mathematically constrained finite VBKL implementation on the 204-band Salinas scene using a spatially separated protocol fixed before the final model evaluations. Labeled pixels are grouped into 24×24 spatial blocks, and a four-pixel cross-split buffer removes 19,152 boundary pixels. The retained split contains 23,998 training, 5,424 validation, and 5,555 test pixels, with all 16 classes represented in every split. Standardization is fitted on the training split only. The standardized spectra are divided by the training 0.99-quantile of their Euclidean norms and are radially projected, when necessary, to the radius-two ball. This projection affected 2 training spectra and no validation or test spectra (0 and 0, respectively).

The constrained VBKL uses 32 paths, depth four under the convention of one normalized linear projection followed by three Brownian profiles, grid size 32, and interval radius $A = 2$. Every effective direction has unit Euclidean norm; every profile is anchored at zero and radially normalized to Brownian energy at most one. Hyperparameters were chosen using validation data only: seed 0 supplied the initial search, seeds 1–3 supplied independent validation confirmation, and the final clean models use seeds 4, 5, 6, 7, 8. The fixed Patch MLP has 15,728 parameters, whereas constrained VBKL has 10,128.

Table 8: Clean Salinas performance (mean $\pm$ standard deviation over five final evaluation seeds).

Model	Acc. (%)	Macro-F1 (%)	Bal. Acc. (%)	Worst F1 (%)	ECE	NLL
Patch MLP	91.87 ± 0.93	94.68 ± 1.14	94.83 ± 1.28	73.93 ± 2.55	0.018 ± 0.002	0.245 ± 0.013
Constrained VBKL	92.18 ± 0.20	95.49 ± 0.40	95.23 ± 0.60	73.83 ± 0.63	0.015 ± 0.004	0.218 ± 0.008

Model	Parameters	Training time (s)	Inference (ms/sample)
Patch MLP	15,728	5.7 ± 0.6	0.001 ± 0.000
Constrained VBKL	10,128	166.2 ± 18.9	0.024 ± 0.000

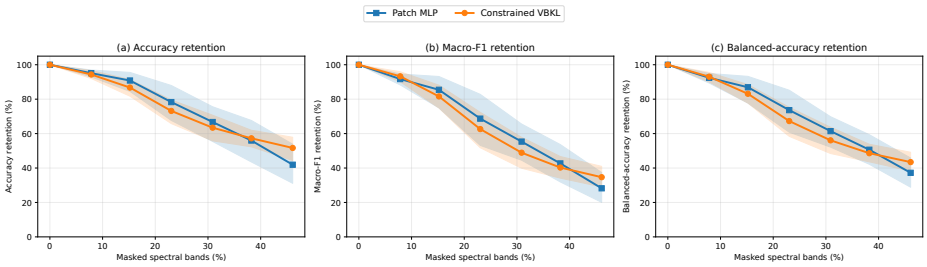

Figure 5: Salinas accuracy, Macro-F1, and balanced-accuracy retention under nested spectral-band masking. Lines are crossed-design means and shaded regions are 95% crossed-bootstrap intervals.

Clean performance. Across the five final seeds, constrained VBKL obtains 92.18% accuracy and 95.49% Macro-F1, compared with 91.87% and 94.68% for Patch MLP. The corresponding balanced accuracies are 95.23% and 94.83%. Constrained VBKL also has lower mean ECE (0.015 versus 0.018) and lower mean NLL. The worst-class F1 values are essentially tied.

Test-time spectral masking. Starting from each clean checkpoint, we mask nested random subsets of 16, 31, 47, 63, 78, and 94 bands. A masked raw coordinate is replaced by its training-set band mean, so it becomes exactly zero after training standardization. We cross five final model seeds with five independent mask seeds, giving 25 cells per corrupted level and 300 model evaluations in total. Confidence intervals resample the model-seed and mask-seed axes separately; exact sign-flip analyses of their marginal means are reported as sensitivity analyses.

Table 9: Aggregate Salinas robustness under nested spectral-band masking. Entries are means with 95% crossed-bootstrap intervals. Retention is higher-is-better; ECE and NLL increases are lower-is-better.

Model	Accuracy-retention AUC (%)	Macro-F1-retention AUC (%)	Balanced-retention AUC (%)
Patch MLP	76.30 [69.82, 82.13]	67.98 [60.57, 74.78]	72.26 [66.21, 77.99]
Constrained VBKL	75.12 [71.53, 78.39]	65.66 [60.29, 70.70]	69.97 [65.71, 74.09]

Model	Level-6 Acc. ret. (%)	Level-6 F1 ret. (%)	ECE-increase AUC	NLL-increase AUC
Patch MLP	41.92 [31.22, 53.95]	28.22 [19.89, 37.42]	0.184 [0.135, 0.235]	1.568 [0.866, 2.367]
Constrained VBKL	51.66 [44.13, 57.85]	34.67 [29.02, 40.85]	0.162 [0.134, 0.195]	1.098 [0.850, 1.378]

Crossover rather than uniform dominance. Across the complete corruption path, Patch MLP has slightly higher mean retention AUC: the BKL-minus-MLP differences are -1.17, -2.32, and -2.29 percentage points for accuracy, Macro-F1, and balanced accuracy, respectively. All corresponding crossed-bootstrap intervals include zero. At the most severe level, however, constrained VBKL retains 51.66% of its clean accuracy and 34.67% of its clean Macro-F1, compared with 41.92% and 28.22% for Patch MLP. The level-six improvements are +9.74 and +6.45 percentage points. All five model-seed marginal means favor constrained VBKL for level-six accuracy retention, yielding an exact two-sided sign-flip value of 0.0625; the crossed-bootstrap interval nevertheless still includes zero because mask realization and model–mask interaction remain substantial.

Calibration and interpretation. Constrained VBKL exhibits a smaller average corruption-induced increase in both ECE and NLL. The crossed mean improvements are +0.021 for ECE-increase AUC and +0.469 for NLL-increase AUC. These effects are directionally favorable but their crossed-bootstrap intervals include zero. Thus this experiment supports a qualified conclusion: Patch MLP has slightly higher average retention AUC over the full masking path, whereas constrained VBKL is more competitive on clean data, has higher mean retention at the most severe masking level, and shows smaller likelihood and calibration deterioration. The effect is heterogeneous across both model initialization and mask realization, so the result does not justify a claim of uniform robustness dominance.

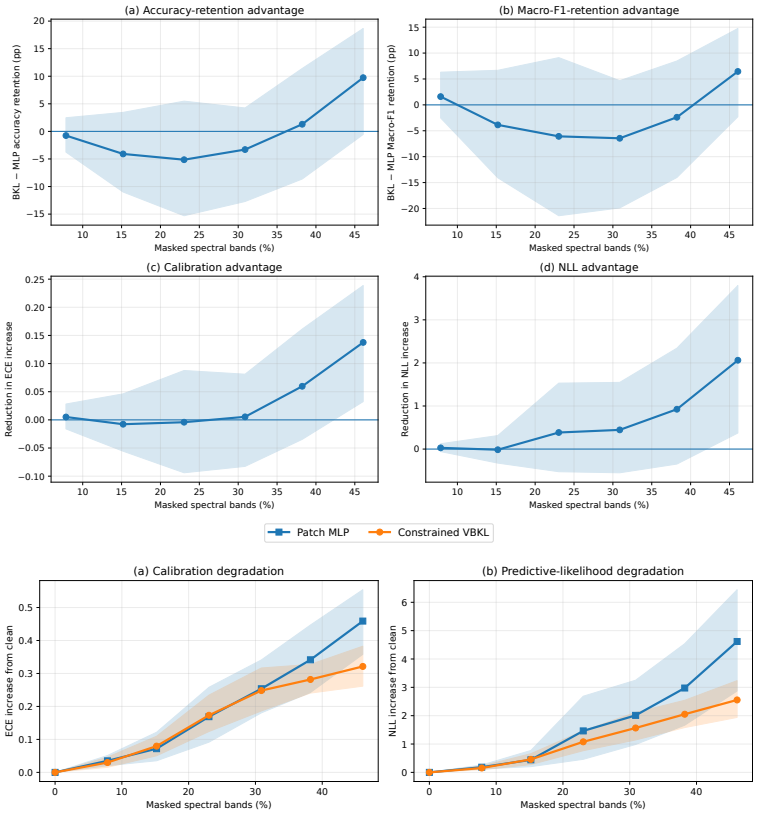

Figure 6: Salinas corruption-path comparisons. The upper panels show BKL-minus-MLP differences, with positive values favoring constrained VBKL; for ECE and NLL, positive values denote a smaller corruption-induced increase. The lower panels show the corresponding calibration and negative-log-likelihood degradation.

Table 10: Crossed Salinas robustness comparisons. Improvements are oriented so that positive values favor constrained VBKL. The confidence interval is obtained by crossed bootstrap. The model-seed and mask-seed columns report BKL-favored/MLP-favored marginal means and exact two-sided sign-flip p-values.

Metric	Favored model	Mean improvement [95% CI]	Cells B/M	Model seeds B/M	p_{model}	Mask seeds B/M	p_{mask}
Accuracy-retention AUC	Patch MLP	$-1.17\,[-6.86,\ +4.33]$	14/11	2/3	0.6875	3/2	0.6250
Macro-F1-retention AUC	Patch MLP	$-2.32\,[-10.28,\ +5.39]$	12/13	2/3	0.6875	2/3	0.4375
Balanced-accuracy-retention AUC	Patch MLP	$-2.29\,[-8.88,\ +4.27]$	12/13	2/3	0.6875	1/4	0.3125
Level-6 accuracy retention	Constrained VBKL	$+9.74\,[-0.54,\ +18.58]$	19/6	5/0	0.0625	3/2	0.2500
Level-6 Macro-F1 retention	Constrained VBKL	$+6.45\,[-2.64,\ +14.72]$	18/7	4/1	0.1875	4/1	0.1875
ECE-increase AUC	Constrained VBKL	$+0.021\,[-0.032,\ +0.074]$	16/9	4/1	0.4375	3/2	0.3125
NLL-increase AUC	Constrained VBKL	$+0.469\,[-0.207,\ +1.194]$	15/10	4/1	0.1250	3/2	0.3125

Table 11: Severity-by-severity Salinas retention. Means are taken over the full 5×5 crossed design at corrupted levels; the clean level uses five model seeds. Differences are BKL minus MLP in percentage points.

Level	Bands	MLP Acc. ret.	BKL Acc. ret.	Δ Acc.	MLP F1 ret.	BKL F1 ret.	Δ F1
0	0	100.00	100.00	+0.00	100.00	100.00	+0.00
1	16	95.09	94.34	-0.75	91.68	93.27	+1.60
2	31	90.83	86.75	-4.09	85.46	81.60	-3.86
3	47	78.30	73.17	-5.13	68.70	62.62	-6.08
4	63	66.72	63.43	-3.29	55.36	48.91	-6.45
5	78	55.97	57.27	+1.29	42.67	40.28	-2.39
6	94	41.92	51.66	+9.74	28.22	34.67	+6.45