

Dynamic Principal Component Analysis in Time Domain

Sangmyung Ha

September 18, 2026

Theorem 1 (Spectral Isometry). *Let $\{X_t\}_{t \in \mathbb{Z}}$ be a zero-mean, covariance-stationary $\mathbb{C}^n$ -valued process. Represent the frequency circle by $\mathbb{T} = (-\pi, \pi]$, let F be the matrix-valued spectral measure of X_t , and let Z_X be the associated orthogonal-increment random measure, so that*

$$X_t = \int_{\mathbb{T}} e^{-it\lambda} dZ_X(\lambda), \quad \mathbb{E}[Z_X(I)Z_X(J)^*] = F(I \cap J)$$

for intervals $I, J \subseteq \mathbb{T}$, where $*$ denotes conjugate transpose. For a positive integer m , define $\mathcal{L}^2(F)$ as the completion of $\mathbb{C}^{m \times n}$ -valued step functions under the seminorm

$$\|\varphi\|_{\mathcal{L}^2(F)}^2 = \int_{\mathbb{T}} \text{tr}(\varphi(\lambda) dF(\lambda) \varphi(\lambda)^*),$$

after identifying functions whose difference has zero seminorm. Its inner product, conjugate-linear in the first argument, is

$$\langle \varphi, \psi \rangle_F = \int_{\mathbb{T}} \text{tr}(\psi(\lambda) dF(\lambda) \varphi(\lambda)^*).$$

Equip $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{C}^m)$ with the inner product $\langle Y, W \rangle = \mathbb{E}[Y^* W]$. Then the stochastic integral defines a linear isometry

$$\tau : \mathcal{L}^2(F) \longrightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{C}^m), \quad \tau(\varphi) = \int_{\mathbb{T}} \varphi(\lambda) dZ_X(\lambda).$$

In particular,

$$\mathbb{E}[\tau(\varphi)\tau(\psi)^*] = \int_{\mathbb{T}} \varphi(\lambda) dF(\lambda) \psi(\lambda)^*, \tag{1}$$

$$\mathbb{E}[\tau(\varphi)^*\tau(\psi)] = \int_{\mathbb{T}} \text{tr}(\psi(\lambda) dF(\lambda) \varphi(\lambda)^*). \tag{2}$$

The mapping is an isometric embedding; it need not be onto the entire space of square-integrable random vectors.

Proof. First consider two step functions. By passing to a common refinement of their partitions, write

$$\varphi(\lambda) = \sum_{j=1}^r C_j \mathbf{1}_{I_j}(\lambda), \quad \psi(\lambda) = \sum_{j=1}^r D_j \mathbf{1}_{I_j}(\lambda),$$

where the intervals $I_j = (\lambda_{j-1}, \lambda_j]$ partition $\mathbb{T}$ and $C_j, D_j \in \mathbb{C}^{m \times n}$. Define

$$\tau(\varphi) = \sum_{j=1}^r C_j Z_X(I_j), \quad \tau(\psi) = \sum_{j=1}^r D_j Z_X(I_j).$$

Orthogonality of increments on disjoint intervals gives

$$\begin{aligned} \mathbb{E}[\tau(\varphi)\tau(\psi)^*] &= \sum_{j=1}^r \sum_{\ell=1}^r C_j \mathbb{E}[Z_X(I_j)Z_X(I_\ell)^*] D_\ell^* \\ &= \sum_{j=1}^r C_j F(I_j) D_j^* \\ &= \int_{\mathbb{T}} \varphi(\lambda) dF(\lambda) \psi(\lambda)^*. \end{aligned}$$

Taking the trace after interchanging φ and ψ establishes (2). In particular,

$$\mathbb{E}\|\tau(\varphi)\|^2 = \|\varphi\|_{\mathcal{L}^2(F)}^2.$$

Functions whose difference has zero seminorm have the same stochastic integral. A Cauchy sequence of integrands produces a Cauchy sequence of random vectors in mean square. Since the target Hilbert space is complete, the mapping extends uniquely to $\mathcal{L}^2(F)$. The covariance and inner-product identities extend by the Cauchy–Schwarz inequality and mean-square convergence. $\square$

1 Dynamic PCA in the Frequency Domain

We consider the optimal-reconstruction formulation of dynamic principal components described by Brillinger (1981). Let $\{X_t\}_{t \in \mathbb{Z}}$ be a zero-mean, covariance-stationary $\mathbb{C}^n$ -valued process with autocovariance matrices

$$\Gamma_X(h) = \mathbb{E}[X_{t+h}X_t^*], \quad h \in \mathbb{Z}.$$

Throughout this section, assume

$$\sum_{h \in \mathbb{Z}} \|\Gamma_X(h)\|_{\text{op}} < \infty. \quad (3)$$

We use the spectral-density normalization

$$\begin{aligned} f_X(\lambda) &= \sum_{h \in \mathbb{Z}} \Gamma_X(h) e^{ih\lambda}, \quad dF_X(\lambda) = \frac{1}{2\pi} f_X(\lambda) d\lambda, \\ \Gamma_X(h) &= \frac{1}{2\pi} \int_{\mathbb{T}} f_X(\lambda) e^{-ih\lambda} d\lambda. \end{aligned} \quad (4)$$

The matrix $f_X(\lambda)$ is Hermitian and positive semidefinite. Assumption (3) also implies that it is continuous and bounded:

$$M_X := \sup_{\lambda \in \mathbb{T}} \|f_X(\lambda)\|_{\text{op}} < \infty.$$

All frequency-domain identities below are understood to hold Lebesgue-almost everywhere unless stated otherwise. The real-valued case is discussed after the reconstruction theorem.

For $1 \leq q < n$, an analysis filter produces a q -dimensional process

$$\xi_t = b(L)X_t = \sum_{k \in \mathbb{Z}} b_k X_{t-k}, \quad b_k \in \mathbb{C}^{q \times n},$$

and a reconstruction filter produces

$$\tilde{X}_t = c(L)\xi_t = \sum_{k \in \mathbb{Z}} c_k \xi_{t-k}, \quad c_k \in \mathbb{C}^{n \times q}.$$

Here $LX_t = X_{t-1}$, and infinite sums, when used as time-domain representations, are interpreted as mean-square limits of symmetric partial sums. With the convention in (4), the corresponding responses are

$$b(e^{i\lambda}) = \sum_{k \in \mathbb{Z}} b_k e^{ik\lambda}, \quad c(e^{i\lambda}) = \sum_{k \in \mathbb{Z}} c_k e^{ik\lambda}.$$

For square-summable coefficients, the Fourier series converge in $L^2(d\lambda)$. Spectral integrability determines whether the associated filtered processes have finite variance.

Admissible frequency-domain filters. To formulate the optimization without interchanging infinite sums, let $\mathcal{A}_q$ be the class of measurable response pairs (b, c) , of dimensions $q \times n$ and $n \times q$, satisfying

$$\begin{aligned} \int_{\mathbb{T}} \text{tr} [b(e^{i\lambda}) f_X(\lambda) b(e^{i\lambda})^*] d\lambda &< \infty, \\ \int_{\mathbb{T}} \text{tr} [c(e^{i\lambda}) b(e^{i\lambda}) f_X(\lambda) b(e^{i\lambda})^* c(e^{i\lambda})^*] d\lambda &< \infty. \end{aligned}$$

These responses define finite-variance processes by

$$\begin{aligned} \xi_t &= \int_{\mathbb{T}} e^{-it\lambda} b(e^{i\lambda}) dZ_X(\lambda), \\ \tilde{X}_t &= \int_{\mathbb{T}} e^{-it\lambda} c(e^{i\lambda}) b(e^{i\lambda}) dZ_X(\lambda). \end{aligned}$$

The notation $b(L)$ and $c(L)$ refers to these spectral filters. We will show that the minimizing responses have square-summable Fourier coefficients that realize both filters as mean-square convergent time-domain sums. No causality or finite-length restriction is imposed in $\mathcal{A}_q$.

The responses are chosen jointly to minimize

$$\mathbb{E} \|X_t - \tilde{X}_t\|^2 = \mathbb{E} \|X_t - c(L)b(L)X_t\|^2.$$

Theorem 2 (Optimal Dynamic Reconstruction). *Under (3), let*

$$\mu_1(\lambda) \geq \cdots \geq \mu_n(\lambda) \geq 0$$

be the eigenvalues of $f_X(\lambda)$. Choose associated orthonormal eigenvectors $u_1(\lambda), \dots, u_n(\lambda)$ measurably in λ , and write

$$U_q(\lambda) = (u_1(\lambda) \quad \cdots \quad u_q(\lambda)), \quad P_q(\lambda) = U_q(\lambda) U_q(\lambda)^*.$$

A jointly optimal pair in $\mathcal{A}_q$ is

$$b(e^{i\lambda}) = U_q(\lambda)^*, \quad c(e^{i\lambda}) = U_q(\lambda). \quad (5)$$

Its reconstruction response is $c(e^{i\lambda})b(e^{i\lambda}) = P_q(\lambda)$, and the minimum error is

$$\min_{(b,c) \in \mathcal{A}_q} \mathbb{E} \|X_t - \tilde{X}_t\|^2 = \frac{1}{2\pi} \int_{\mathbb{T}} \sum_{j=q+1}^n \mu_j(\lambda) d\lambda. \quad (6)$$

The Fourier coefficients

$$\begin{aligned} b_k &= \frac{1}{2\pi} \int_{\mathbb{T}} b(e^{i\lambda}) e^{-ik\lambda} d\lambda, \\ c_k &= \frac{1}{2\pi} \int_{\mathbb{T}} c(e^{i\lambda}) e^{-ik\lambda} d\lambda \end{aligned} \quad (7)$$

define the optimal time-domain filters, with both filtering sums converging in mean square. Distinct eigenvalues are not required for the existence or optimality of this construction.

Proof. For any admissible pair, let $A(e^{i\lambda}) = c(e^{i\lambda})b(e^{i\lambda})$, which has rank at most q at each frequency. Theorem 1 gives

$$\begin{aligned} \mathbb{E} \|X_t - \tilde{X}_t\|^2 &= \frac{1}{2\pi} \int_{\mathbb{T}} \text{tr}[(I_n - A)f_X(I_n - A)^*] d\lambda \\ &= \frac{1}{2\pi} \int_{\mathbb{T}} \|f_X^{1/2} - Af_X^{1/2}\|_F^2 d\lambda, \end{aligned}$$

where the frequency arguments are suppressed inside the integrals. Since $Af_X^{1/2}$ has rank at most q , the best rank- q approximation inequality implies

$$\|f_X(\lambda)^{1/2} - A(e^{i\lambda})f_X(\lambda)^{1/2}\|_F^2 \geq \sum_{j=q+1}^n \mu_j(\lambda).$$

The choice $A(e^{i\lambda}) = P_q(\lambda)$ attains this lower bound because

$$\|(I_n - P_q(\lambda))f_X(\lambda)^{1/2}\|_F^2 = \sum_{j=q+1}^n \mu_j(\lambda).$$

A measurable orthonormal eigenvector selection exists for a measurable finite-dimensional Hermitian matrix function. The resulting responses in (5) are bounded, with operator norms equal to one, and therefore belong to $\mathcal{A}_q$. Integrating the pointwise lower bound and using this admissible attaining pair proves (6).

It remains to justify the time-domain representations. The optimal responses belong to ordinary $L^2(d\lambda)$, so their Fourier partial sums $b^{(r)}(e^{i\lambda}) = \sum_{|k| \leq r} b_k e^{ik\lambda}$ and $c^{(r)}(e^{i\lambda}) = \sum_{|k| \leq r} c_k e^{ik\lambda}$ converge in $L^2(d\lambda)$. Parseval's identity gives

$$\sum_{k \in \mathbb{Z}} \|b_k\|_F^2 = \sum_{k \in \mathbb{Z}} \|c_k\|_F^2 = q.$$

By spectral isometry and boundedness of f_X ,

$$\begin{aligned} \mathbb{E} \left\| \sum_{|k| \leq r} b_k X_{t-k} - \xi_t \right\|^2 &\leq \frac{M_X}{2\pi} \int_{\mathbb{T}} \|b^{(r)}(e^{i\lambda}) - b(e^{i\lambda})\|_F^2 d\lambda \\ &\longrightarrow 0. \end{aligned}$$

Moreover, the spectral density of the optimal score process is

$$f_\xi(\lambda) = b(e^{i\lambda}) f_X(\lambda) b(e^{i\lambda})^* = \text{diag}(\mu_1(\lambda), \dots, \mu_q(\lambda)),$$

so $\sup_\lambda \|f_\xi(\lambda)\|_{\text{op}} \leq M_X$. Applying the same argument to the reconstruction filter gives

$$\begin{aligned} \mathbb{E} \left\| \sum_{|k| \leq r} c_k \xi_{t-k} - \tilde{X}_t \right\|^2 &\leq \frac{M_X}{2\pi} \int_{\mathbb{T}} \|c^{(r)}(e^{i\lambda}) - c(e^{i\lambda})\|_F^2 d\lambda \\ &\longrightarrow 0. \end{aligned}$$

Thus, both optimal spectral filters are realized by the asserted time-domain sums. $\square$

The adjoint filter and real-valued coefficients. For the normalized optimal pair, $c(e^{i\lambda}) = b(e^{i\lambda})^*$ implies

$$\begin{aligned} c_k &= \frac{1}{2\pi} \int_{\mathbb{T}} b(e^{i\lambda})^* e^{-ik\lambda} d\lambda \\ &= \left(\frac{1}{2\pi} \int_{\mathbb{T}} b(e^{i\lambda}) e^{ik\lambda} d\lambda \right)^* = b_{-k}^*. \end{aligned} \tag{8}$$

This identity does not require X_t or the coefficients to be real-valued. If $X_t \in \mathbb{R}^n$, then $f_X(-\lambda) = \overline{f_X(\lambda)}$. The eigenvectors can be chosen to satisfy

$$U_q(-\lambda) = \overline{U_q(\lambda)},$$

with frequencies interpreted on the circle. For example, choose them measurably on $(0, \pi)$ and extend by conjugation; real eigenvectors can be chosen at 0 and π . This selection yields $b(e^{-i\lambda}) = \overline{b(e^{i\lambda})}$ and real Fourier coefficients. In that case, (8) becomes $c_k = b'_{-k}$, where $'$ denotes ordinary transpose. An arbitrary eigenvector selection need not have the required conjugate symmetry.

Nonuniqueness. The theorem specifies an optimal pair, not a unique pair. Eigenvectors have phase indeterminacy, and repeated eigenvalues permit rotations within their eigenspaces. The projection $P_q(\lambda)$ is uniquely determined wherever $\mu_q(\lambda) > \mu_{q+1}(\lambda)$, but a tie at the cutoff can make the retained subspace nonunique. Moreover, replacing b by Qb and c by cQ^* , for a measurable unitary $q \times q$ response Q , preserves the reconstruction and total score variance. Such a change of coordinates need not preserve diagonality of the score spectral density. The all-lead-and-lag orthogonality below refers specifically to the eigenvector-based coordinates in (5).

Properties of the dynamic principal components. For the eigenvector-based score process $\xi_t = b(L)X_t$,

$$f_\xi(\lambda) = \text{diag}(\mu_1(\lambda), \dots, \mu_q(\lambda)).$$

Thus, distinct components have zero cross-spectral density almost everywhere. Their coherence is zero wherever both marginal spectral densities are positive; coherence is undefined where its denominator vanishes. Fourier inversion gives

$$\Gamma_\xi(h) = \frac{1}{2\pi} \text{diag} \left(\int_{\mathbb{T}} \mu_1(\lambda) e^{-ih\lambda} d\lambda, \dots, \int_{\mathbb{T}} \mu_q(\lambda) e^{-ih\lambda} d\lambda \right).$$

Consequently, for $i \neq j$,

$$\mathbb{E}[\xi_{i,t+h} \overline{\xi_{j,t}}] = 0, \quad h \in \mathbb{Z}.$$

The component processes are mutually uncorrelated at all leads and lags, a stronger property than the contemporaneous uncorrelatedness guaranteed by static PCA. Individual components may nevertheless be serially correlated; they are not generally white noise.

For any response satisfying $b(e^{i\lambda})b(e^{i\lambda})^* = I_q$, the eigenvalue variational inequality gives

$$\text{tr}[b(e^{i\lambda})f_X(\lambda)b(e^{i\lambda})^*] \leq \sum_{j=1}^q \mu_j(\lambda). \quad (9)$$

The eigenvector-based response attains equality. Therefore,

$$\max_{bb^*=I_q} \text{tr}(\Gamma_\xi(0)) = \frac{1}{2\pi} \int_{\mathbb{T}} \sum_{j=1}^q \mu_j(\lambda) d\lambda.$$

The individual eigenvalues have a sequential variance-maximization interpretation, with each eigenvector constrained to be orthogonal to those previously selected. Under the joint normalization alone, the relevant objective is the sum of the retained variances. No maximization of the nonzero-lag autocovariances $\Gamma_\xi(h)$ is imposed.

Claim (Variance Maximization and Reconstruction). *Let $\mathcal{B}_q$ be the set of square-summable coefficient sequences $\{b_k\}_{k \in \mathbb{Z}}$ satisfying the static and dynamic constraints*

$$\sum_{k \in \mathbb{Z}} b_k b_k^* = I_q, \quad (10)$$

$$\sum_{p \in \mathbb{Z}} b_p b_{p+j}^* = 0, \quad j = 1, 2, \dots \quad (11)$$

These constraints are equivalent to $b(e^{i\lambda})b(e^{i\lambda})^* = I_q$ almost everywhere. For every $\{b_k\} \in \mathcal{B}_q$, choose the adjoint reconstruction filter $c_k = b_{-k}^*$. Then

$$\mathbb{E}\|X_t - \tilde{X}_t\|^2 = \text{tr}(\Gamma_X(0)) - \text{tr}(\Gamma_\xi(0)). \quad (12)$$

Moreover,

$$\min_{(b,c) \in \mathcal{A}_q} \mathbb{E}\|X_t - \tilde{X}_t\|^2 = \text{tr}(\Gamma_X(0)) - \max_{\{b_k\} \in \mathcal{B}_q} \text{tr}(\Gamma_\xi(0)). \quad (13)$$

In particular, there exists a jointly optimal normalized pair with $c_k = b_{-k}^*$.

Proof. First establish the equivalence of the frequency-domain and time-domain constraints. For every integer j , define

$$R_j := \sum_{p \in \mathbb{Z}} b_p b_{p+j}^*.$$

Each sum converges absolutely in Frobenius norm because

$$\sum_{p \in \mathbb{Z}} \|b_p b_{p+j}^*\|_F \leq \left(\sum_{p \in \mathbb{Z}} \|b_p\|_F^2 \right)^{1/2} \left(\sum_{p \in \mathbb{Z}} \|b_{p+j}\|_F^2 \right)^{1/2} < \infty.$$

Let $b^{(r)}$ be the symmetric Fourier partial sums of b . Since $b^{(r)} \rightarrow b$ in L^2 , the products converge in L^1 :

$$\|b^{(r)} b^{(r)*} - b b^*\|_{L^1} \leq (\|b^{(r)}\|_{L^2} + \|b\|_{L^2}) \|b^{(r)} - b\|_{L^2} \rightarrow 0,$$

where these norms use the Frobenius norm and normalized Lebesgue measure. Computing Fourier coefficients for the finite sums and taking limits yields

$$R_j = \frac{1}{2\pi} \int_{\mathbb{T}} b(e^{i\lambda}) b(e^{i\lambda})^* e^{ij\lambda} d\lambda. \quad (14)$$

By uniqueness of Fourier coefficients for L^1 functions, $b b^* = I_q$ almost everywhere if and only if $R_0 = I_q$ and $R_j = 0$ for every nonzero integer j . Since $R_{-j} = R_j^*$, the positive-lag restrictions suffice. This proves the equivalence with (10)–(11).

For a normalized response, let $P(\lambda) = b(e^{i\lambda})^* b(e^{i\lambda})$. Then $P^* = P$ and $P^2 = P$. The response b is bounded, so $f_\xi = b f_X b^*$ is bounded as well. The square-summable Fourier coefficients of b and its adjoint therefore define mean-square filters by the same argument used in Theorem 2. With $c = b^*$, the reconstruction response is P , and

$$\begin{aligned} \text{tr}[(I_n - P) f_X (I_n - P)] &= \text{tr}(f_X) - 2 \text{tr}(P f_X) + \text{tr}(P f_X P) \\ &= \text{tr}(f_X) - \text{tr}(P f_X) \\ &= \text{tr}(f_X) - \text{tr}(b f_X b^*). \end{aligned}$$

Integrating this identity and applying spectral isometry proves (12).

Finally, (9) bounds the retained variance for every normalized response, and the response in (5) attains that bound. Its Fourier coefficients belong to $\mathcal{B}_q$ by (14). Comparing the resulting variance with the minimum reconstruction error in (6) proves (13). This establishes equivalence at the joint optimum; it does not assert that the adjoint decoder is conditionally optimal for every normalized analysis filter. $\square$

Optimal reconstruction for a fixed analysis filter. The last distinction is important. At a fixed frequency, suppress the frequency arguments and set $S = b f_X b^*$. Minimizing

$$J(c) = \text{tr}[(I_n - c b) f_X (I_n - c b)^*]$$

over c gives the normal equation

$$c S = f_X b^*.$$

When S is nonsingular, the unique minimizer is

$$c_{\text{opt}} = f_X b^* (b f_X b^*)^{-1}.$$

For singular S , replacing the inverse by the Moore–Penrose pseudoinverse gives a pointwise minimizer. Under $bb^* = I_q$, the choice $c = b^*$ satisfies the normal equation if and only if

$$(I_n - b^*b)f_X b^* = 0. \quad (15)$$

This means that the subspace spanned by the columns of b^* is invariant under f_X . It holds for the eigenvector-based solution, but does not follow from $bb^* = I_q$ alone. Equation (12), in contrast, holds for every normalized analysis filter paired with its adjoint, irrespective of conditional optimality. Also, the pointwise formula for c_{opt} does not guarantee a decoder of a prescribed finite length or a causal decoder.

Interpretation of the coefficient constraints. The restrictions (10)–(11) describe orthogonality of shifted filter coefficient sequences. They do not require $\Gamma_\xi(h) = 0$ for $h \neq 0$. They normalize the frequency response of the analysis filter, not the temporal dependence of the input or the scores. For real coefficients, every conjugate transpose in these restrictions can be replaced by an ordinary transpose.

2 Dynamic PCA in the Time Domain

Consider finite-length, two-sided analysis and reconstruction filters,

$$\xi_t = \sum_{k=-m}^m b_k X_{t-k}, \quad \tilde{X}_t = \sum_{k=-m}^m c_k \xi_{t-k},$$

where $m \geq 0$, $b_k \in \mathbb{C}^{q \times n}$, and $c_k \in \mathbb{C}^{n \times q}$. Extend both coefficient sequences by zero outside $[-m, m]$. We study a population optimization problem over these finite sequences. The following identity needs only zero-mean covariance stationarity of X_t ; absolute summability of its autocovariances is not required.

Claim (Finite-Filter Reconstruction Identity). *Suppose that*

$$c_k = b_{-k}^*, \quad k = -m, \dots, m, \quad (16)$$

$$\sum_{k=-m}^m b_k b_k^* = I_q, \quad (17)$$

$$\sum_{p=-m}^{m-j} b_p b_{p+j}^* = 0, \quad j = 1, \dots, 2m. \quad (18)$$

Then

$$\mathbb{E} \|X_t - \tilde{X}_t\|^2 = \text{tr}(\Gamma_X(0)) - \text{tr}(\Gamma_\xi(0)). \quad (19)$$

Consequently, maximizing $\text{tr}(\Gamma_\xi(0))$ over analysis coefficients satisfying (17)–(18) is equivalent to minimizing reconstruction error over analysis and reconstruction coefficients satisfying (16)–(18).

Proof. The score autocovariances are

$$\Gamma_\xi(h) = \sum_{k=-m}^m \sum_{\ell=-m}^m b_k \Gamma_X(h + \ell - k) b_\ell^*. \quad (20)$$

Under (16), write

$$\tilde{X}_t = \sum_{r=-m}^m b_r^* \xi_{t+r}.$$

Joint covariance stationarity of X_t and ξ_t gives

$$\begin{aligned} \mathbb{E}[\tilde{X}_t^* X_t] &= \sum_{r=-m}^m \mathbb{E}[\xi_{t+r}^* b_r X_t] \\ &= \sum_{r=-m}^m \mathbb{E}[\xi_t^* b_r X_{t-r}] \\ &= \mathbb{E} \left[\xi_t^* \sum_{r=-m}^m b_r X_{t-r} \right] \\ &= \mathbb{E}[\xi_t^* \xi_t] = \text{tr}(\Gamma_\xi(0)). \end{aligned}$$

This quantity is real, so its complex conjugate $\mathbb{E}[X_t^* \tilde{X}_t]$ has the same value.

Next, define $R_h = \sum_{s \in \mathbb{Z}} b_s b_{s+h}^*$ using the zero extension. Conditions (17)–(18) give $R_0 = I_q$ and $R_h = 0$ for $h \neq 0$, where negative lags follow from $R_{-h} = R_h^*$. Therefore,

$$\begin{aligned} \mathbb{E}\|\tilde{X}_t\|^2 &= \sum_{r=-m}^m \sum_{s=-m}^m \text{tr}[b_r^* \Gamma_\xi(r-s) b_s] \\ &= \sum_{h=-2m}^{2m} \text{tr} \left[\left(\sum_{s \in \mathbb{Z}} b_s b_{s+h}^* \right) \Gamma_\xi(h) \right] \\ &= \sum_{h=-2m}^{2m} \text{tr}[R_h \Gamma_\xi(h)] \\ &= \text{tr}(\Gamma_\xi(0)). \end{aligned}$$

Expanding the reconstruction error now yields

$$\begin{aligned} \mathbb{E}\|X_t - \tilde{X}_t\|^2 &= \mathbb{E}\|X_t\|^2 - 2 \text{Re} \mathbb{E}[\tilde{X}_t^* X_t] + \mathbb{E}\|\tilde{X}_t\|^2 \\ &= \text{tr}(\Gamma_X(0)) - \text{tr}(\Gamma_\xi(0)), \end{aligned}$$

which proves the result. $\square$

The finite-dimensional variance problem. By (20), the objective is

$$\text{tr}(\Gamma_\xi(0)) = \sum_{k=-m}^m \sum_{\ell=-m}^m \text{tr}[b_k \Gamma_X(\ell-k) b_\ell^*]. \quad (21)$$

Equivalently, define

$$\mathcal{X}_t = \begin{pmatrix} X_{t+m} \\ \vdots \\ X_{t-m} \end{pmatrix}, \quad B = (b_{-m} \quad \cdots \quad b_m), \quad S_m = \mathbb{E}[\mathcal{X}_t \mathcal{X}_t^*].$$

Then $\xi_t = B\mathcal{X}_t$ and

$$\text{tr}(\Gamma_\xi(0)) = \text{tr}(BS_m B^*).$$

The static constraint is $BB^* = I_q$, but the dynamic constraints in (18) must also be imposed. Ordinary PCA of the stacked vector $\mathcal{X}_t$ imposes only $BB^* = I_q$ and therefore does not generally solve this constrained problem. The feasible set is nonempty, closed, and bounded, so the continuous population objective attains its maximum. For example, a feasible sequence is obtained by taking b_0 to have orthonormal rows and setting all other coefficients to zero.

Finite constraints are not preserved by truncation. Solving the finite constrained problem is different from truncating an infinite optimal filter. If $\{b_k\}_{k \in \mathbb{Z}}$ satisfies (10), then its truncation satisfies

$$\sum_{k=-m}^m b_k b_k^* = I_q - \sum_{|k| > m} b_k b_k^*,$$

which generally differs from I_q . The truncated shifted-product sums need not vanish either. Thus, a simple truncation does not generally satisfy (17)–(18), and the exact identity (19) need not hold for the truncated pair. Under the spectral assumptions of the preceding section, the maximum retained variance in the finite constrained problem cannot exceed the unrestricted value in (9), integrated over frequency. Equality, or convergence to the unrestricted value as m increases, requires a separate approximation argument.

Scope of the finite-filter equivalence. Claim 2 concerns reconstruction under all three restrictions (16)–(18). It does not identify the optimum over arbitrary finite analysis and reconstruction filters. The pointwise optimization in Theorem 2 cannot simply be repeated with finite-length restrictions: finite Fourier polynomials link the choices across frequencies, and the adjoint relationship need not be without loss within that restricted class.

For example, take $q = 1$, let z_t be scalar white noise with variance one and let

$$X_t = \begin{pmatrix} z_t \\ z_t + z_{t-1} \end{pmatrix}.$$

The finite filters $\xi_t = X_{1t} = z_t$ and $\tilde{X}_t = (\xi_t, \xi_t + \xi_{t-1})'$ reconstruct X_t exactly. However, with

$$h(\lambda) = \begin{pmatrix} 1 \\ 1 + e^{i\lambda} \end{pmatrix}, \quad f_X(\lambda) = h(\lambda)h(\lambda)^*,$$

exact reconstruction by a normalized rank-one adjoint pair requires the orthogonal projection

$$P(\lambda) = \frac{h(\lambda)h(\lambda)^*}{h(\lambda)^*h(\lambda)}, \quad P_{11}(\lambda) = \frac{1}{3 + 2 \cos \lambda}.$$

The last function is not a finite trigonometric polynomial, whereas $b(e^{i\lambda})^* b(e^{i\lambda})$ is a finite trigonometric polynomial for every finite analysis filter. Hence no finite-length normalized adjoint pair gives this exact reconstruction, even though unrestricted finite analysis and reconstruction filters do.

Population identities and sample implementation. The preceding results involve population covariances and expectations. A sample implementation may replace S_m by a sample covariance of complete stacked observations, while retaining the coefficient constraints. Statistical consistency requires additional assumptions and is not established by the population reconstruction identity alone. Boundary treatment also matters: composing two filters supported on $[-m, m]$ can use observations from $t - 2m$ through $t + 2m$. With observations $X_1, \dots, X_T$ and no padding, complete reconstructions are available only for $2m + 1 \leq t \leq T - 2m$. Arbitrary endpoint truncation does not preserve the population identity as an exact finite-sample equality.