Dynamically Optimal Projection onto Slow Spectral Manifolds for Linear Systems

Florian Kogelbauer[1, *] and Ilya Karlin[1, †]

[1]*Department of Mechanical and Process Engineering, ETH Zurich, CH-8092 Zurich, Switzerland*
(Dated: March 25, 2025)

We derive the dynamically optimal projection onto the linear slow manifold from a temporal variational principle. We demonstrate that the projection captures transient dynamics of the overall dissipative system and leads to a considerably improved fit of reduced trajectories compared to full trajectories. We illustrate these optimal model reduction properties on explicit examples, including the linear three-component Grad's moment system.

I. INTRODUCTION

Slow manifolds play a crucial role in model reduction, equilibrium analysis, and data assimilation [15]. Originating from the foundational works of Tikhonov [43] and Fenichel [13] within geometric singular perturbation theory, slow manifolds have become essential tools in fields such as atmospheric sciences [33], fluid dynamics [23], and classical mechanics [34]. The modern framework of slow spectral submanifolds, introduced in [4–6], builds on stable spectral properties near equilibria and quasi-periodic solutions. This theory was further advanced in [20] and has recently found applications in data-driven nonlinear dynamics [7].

The central idea of slow-manifold reduction is based on the separation of time scales, arising from the ordering of negative real parts in the system's linearization. Similar concepts appear in the theory of inertial manifolds for dissipative PDEs [39], which contain the global attractor and are defined by the long-term behavior of the dynamical system [42].

Once the existence of a slow manifold is established and an approximate parametrization is available, an important question arises: how do general initial conditions relate to the reduced dynamics on the slow manifold? While reduced dynamics on the slow manifold can still exhibit transients due to repeated eigenvalues with Jordan blocks, methods such as spectral submanifolds do not capture the transient dynamics associated with trajectories approaching the slow manifold from off-manifold initial conditions.

Even in globally stable systems trajectories converge to a unique minimum but may display prolonged transient behavior. In linear systems, such transients are linked to the non-normality of the governing operator. Transient stable dynamics are pervasive across physics and engineering, with examples from atmospheric science [12], polymer dynamics [24, 25], and thermal convection [26]. In fluid models, long transients have even been proposed as an explanation for turbulence [3, 21], although this remains debated, given the inherently nonlinear nature of turbulence [44]. Transient dynamics are also believed to play a significant role in hypocoercivity estimates [1].

The problem of projecting onto the slow manifold in the presence of transient dynamics has a long history [14]. General model reduction methods often lead to oblique projections that account for non-normality in linear settings [2, 19, 40]. Proper orthogonal decomposition combined with Galerkin approximations is one widely used technique [41] in this context. Near the slow manifold, the correct affine projections become position-dependent and satisfy a nonlinear partial differential equation [37], which can be approximated through series expansions via computer algebra [38]. More recently, projections respecting appropriate fibers for reduced-order modeling have been developed using constrained autoencoders [35].

In this work, we introduce a novel projection onto the slow manifold for linear systems, which we term the *dynamically optimal projection*. This projection optimally captures the transient dynamics of general trajectories in phase space. Derived from a temporal variational principle, the dynamical projection reduces to the standard orthogonal projection in the case of normal operators, ensuring consistency with the classical Riesz projection [22] onto the span of eigenvectors. Unlike existing oblique projections that address non-normality, such as the $\mathcal{H}_2$-optimal projection [19], which compares full and reduced trajectories via their temporal Laplace transforms, our approach presumes knowledge of the slow manifold as a linear combination of eigenvectors associated with slow modes. This combines the rational of the slow spectral submanifold with global phase-space projections for linear stable systems. Additionally, our variational principle is formulated in Hilbert space, relying solely on the spectral properties of the operator generating the semigroup. We also contrast the dynamically optimal projection with other projection methods from statistical physics. One notable example is the thermodynamic projector [17], designed for stable systems with a global Lyapunov function and primarily used in kinetic and chemical dynamics. The thermodynamic projector requires only the gradient of the entropy, imparting a degree of universality, and can define a vector field on general, possibly non-invariant subsets of phase space. In contrast, the dynamically optimal projection requires a more detailed input, particularly the spectral data of the underlying linear operator.

---

* floriank@ethz.ch

† ikarlin@ethz.ch

II. SPECTRAL THEORY AND SLOW MANIFOLDS

Consider an evolution equation of the form

$$\frac{dx}{dt} = Lx, \qquad (1)$$

for a closed linear operator L on the complex Hilbert space H with inner product $\langle .,. \rangle$. For a linear operator $L : H \to H$, we denote its spectrum as $\mathrm{spec}(L)$, its adjoint as $L^\dagger$ and recall that $\mathrm{spec}(L^\dagger) = \mathrm{spec}(L)^*$, see [9]. The commutator of two linear operators L_1 and L_2 is denoted as $[L_1, L_2] = L_1 L_2 - L_2 L_1$.

We assume that the spectrum of L is stable and consists of a continuous part plus a set of isolated eigenvalues of finite multiplicity above the essential spectrum, called the *point spectrum*, i.e.,

$$\mathrm{spec}(L) = \mathrm{spec}_{\mathrm{ess}}(L) \cup \{\lambda_j\}_{1 \le j \le n}, \ \Re\mathrm{spec}(L) < 0, \quad (2)$$

such that $\Re\mathrm{spec}_{\mathrm{ess}}(L) < \Re\{\lambda_j\}_{1 \le j \le n}$, where $n \in \mathbb{N} \cup \{\infty\}$. We assume that the eigenvalues are counted with multiplicity and that they ordered by real part such that $\Re\lambda_i \le \Re\lambda_j$ for $i < j$. Furthermore, we assume that L restricted to the eigenvectors associated to the point spectrum is diagonal, i.e., has a trivial Jordan-block structure. The case of repeated eigenvalues and generalized eigenvectors can be treated analogously and is omitted for notational simplicity.

As mentioned before, operators satisfying (2) appear in the context of fluctuations-dissipation theory as the Fokker–Planck operator of stochastically forced systems [36] and kinetic theory [8]. The case of a pure point spectrum, $\mathrm{spec}_{\mathrm{ess}}(L) = \varnothing$, in particular finite-dimensional systems is also covered by considerations below.

The stability of the spectrum, $\Re\mathrm{spec}(L) < 0$, implies that the associated semi-group e^{tL} is exponentially stable. In fact, the converse is also true [11]. For kinetic systems, this corresponds to the increase of the quadratic Lyapunov function $x \mapsto -\langle x, x \rangle$, which is interpreted as the entropy of the system. In particular, any solution converges to zero as $t \to \infty$, but might exhibit large temporal transients before decaying.

In the following, we will mostly be interested in the case when L is non-normal, i.e., when $[L, L^\dagger] \ne 0$. For a treatment of the general spectral properties of non-normal operators derived as perturbations to multiplication operators, we refer to [32]. This operator class is of particular interest in kinetic theory as regular perturbations to the transport operator [30].

The eigenvectors $\{\hat{x}_j\}_{1 \le j \le n}$ associated to the discrete eigenvalues, the slow modes compared to the essential spectrum, span an invariant hyperplane of the linear dynamical system (1). Henceforth, we will call the hyperplane spanned by the eigenvectors the *slow manifold* and denote it as $\mathcal{M}_{\mathrm{slow}}^n$. The separation of the slow manifold $\mathcal{M}_{\mathrm{slow}}^n$ from the dynamics associated to the essential spectrum is characterized by different time scales: The contributions to a general solution of the essential spectrum will decay faster as compared to the contributions of the point spectrum.

Within the full slow manifold, the ordering of the eigenvalues by real part induces itself a hierarchy of time scales, i.e., attraction rates, and implies the existence of a nested family of invariant, slow manifolds,

$$\mathcal{M}_{\mathrm{slow}}^1 \subset \mathcal{M}_{\mathrm{slow}}^2 \subset ... \subset \mathcal{M}_{\mathrm{slow}}^n. \qquad (3)$$

The knowledge of the reduced model induced by operator L on a slow manifold is only the first step, as one needs to know *how* to project a general initial condition onto the slow manifold. As elaborated in the introduction, general solutions to (1) under the stability assumption (2) may exhibit large transient dynamics for non-normal operators L, i.e., short- to medium-time growth properties and deviations from equilibrium before decaying exponentially fast. In general, reduced models are incapable of capturing these transient dynamics accurately. Below, we will derive a projection of a generic initial condition onto the slow invariant manifold by minimizing the total-in-time deviation of a full trajectory to a solution on the slow manifold in a mean-square sense.

III. DYNAMICALLY OPTIMAL PROJECTION FROM A VARIATIONAL PRINCIPLE

Let $\{\hat{x}_j\}_{1 \le j \le n}$ be a collection of slow modes spanning the slow manifold $\mathcal{M}_{\mathrm{slow}}^n$. The total dynamical error of a trajectory starting on the slow manifold compared to a general trajectory is given by,

$$\mathcal{E}(x_0, \xi) = \frac{1}{2} \int_0^\infty \|e^{tL} x_0 - e^{tL} x_{\mathrm{slow}}(\xi)\|^2 \, dt, \qquad (4)$$

for any initial condition $x_0 \in H$ and a parameter vector $\xi \in \mathbb{C}^n$, where

$$x_{\mathrm{slow}}(\xi) = \sum_{j=1}^n \xi_j \hat{x}_j, \qquad (5)$$

is an initial condition on the slow manifold for the vector $\xi = (\xi_1, ..., \xi_n)$. The reduced dynamics on the slow manifold for this parametrization is simply given by,

$$\frac{d\xi_j}{dt} = \lambda_j \xi_j. \qquad (6)$$

The question to be addressed here is: *Which initial condition on the slow manifold corresponds to a generic initial condition of the system such that the error* (4) *will be minimal?*

Expression (4) is a well-defined, convex quadratic form in ξ, see Appendix A, and hence has a unique minimum,

$$\xi^{\min}(x_0) := \underset{\xi \in \mathbb{C}^n}{\arg\min} \, \mathcal{E}(x_0, \xi). \qquad (7)$$
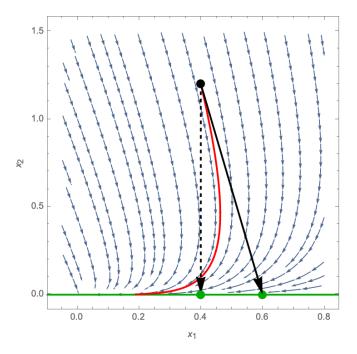
FIG. 1. Illustration of the dynamical projection for system (13) ($\alpha = 5$, $\gamma = 1$) applied to the initial condition $x_0 = (0.4, 1.2)$ for which the trajectory to the full system is shown in red. The dynamically optimal projection (solid arrow) moves the initial condition on the slow manifold (x-axis in green) farther away from the global equilibrium compared to the orthogonal projection (dashed arrow).

Let us define the $n \times n$ spectrally-weighted Gramian matrix G with the elements,

$$G_{ij} = \frac{\langle \hat{x}_i, \hat{x}_j \rangle}{\lambda_i + \lambda_j^*}, \ 1 \le i, j \le n, \tag{8}$$

and mention its Hermitian symmetry, $G^\dagger = G$. Moreover, matrix G is negative semi-definite: $\xi \cdot G\xi^* \le 0$, for $\xi \in \mathbb{C}^n$, see Appendix A.

With the spectrally-weighted Gramian matrix (8), the explicit solution to the minimization problem (7) is found as follows:

Proposition 1. *The unique minimizer* (7) *reads,*

$$\xi_i^{\min}(x_0) = \sum_{j=1}^n \left[\left(G^T \right)^{-1} \right]_{ij} \langle (L + \lambda_j^*)^{-1} x_0, \hat{x}_j \rangle, \tag{9}$$

for $1 \le i \le n$.

The proof of Proposition 1 is given in Appendix B. Note that, the minimizer (9) is well-defined thanks to the stability assumption on the spectrum, $-\lambda^* \notin \mathrm{spec}(L)$.

With the explicit form of the minimizer $\xi^{\min}$ (9), we can define a linear operator on H as follows:

$$\mathbb{P}_{\mathrm{DOP}} x = \sum_{i=1}^n \sum_{j=1}^n \hat{x}_i \left[\left(G^T \right)^{-1} \right]_{ij} \langle (L + \lambda_j^*)^{-1} x, \hat{x}_j \rangle. \tag{10}$$

Moreover, we have the following:

Proposition 2. *Linear operator* $\mathbb{P}_{\mathrm{DOP}}$ (10) *is a projection,*

$$\mathbb{P}_{\mathrm{DOP}}^2 = \mathbb{P}_{\mathrm{DOP}}. \tag{11}$$

The projection property (11) is proved in Appendix C.

We call $\mathbb{P}_{\mathrm{DOP}}$ (10) the *dynamically optimal projection* (DOP) associated to the slow manifold. Clearly, the range of $\mathbb{P}_{\mathrm{DOP}}$ is the whole slow manifold (5) by construction, while its kernel is given by,

$$\ker \mathbb{P}_{\mathrm{DOP}} = \{x \in H : (L + \lambda_j^*)^{-1} x \perp \hat{x}_j, \ 1 \le j \le n\}. \tag{12}$$

Let us summarize the properties of DOP: Given a general initial condition x_0 to system (1), the dynamical projection (10) returns the unique initial condition $\mathbb{P}_{\mathrm{DOP}} x_0$ on the slow manifold that minimizes the cumulative difference of the full trajectory to its counterpart on the slow manifold. Intuitively, the solution of the optimization problem induces a projection by the minimality of the cumulative error of a full trajectory compared to a point of the parametrization on the slow manifold. Indeed, if the trajectory starts with an initial condition on the slow manifold, then the cumulative error (4) vanishes.

Let us take a closer look at the two ingredients of $\mathbb{P}_{\mathrm{DOP}}$. First, we emphasize that for a generic closed operator L, the spectrally-weighted Gramian (8) does not simplify further since eigenvectors corresponding to different eigenvalues need not be orthogonal to each other. However, if L is normal, we have, $L^\dagger \hat{x}_j = \lambda_j^* \hat{x}_j$, and eigenvectors corresponding to different eigenvalues are orthogonal [9]. In this case, the spectrally-weighted Gramian G (8) reduces to a diagonal matrix, while the dynamically optimal projection becomes the canonical orthogonal projection, see Appendix D.

Finally, let us comment that, for a general, not necessarily normal operator L, the projection $\mathbb{P}_{\mathrm{DOP}}$ (10) does not commute with the operator L, since the vectors $\hat{x}_i$ are not necessarily the eigenvectors of the adjoint operator $L^\dagger$. This distinguishes the DOP from the classical Riesz projection [22] which was recently used to project the (non-normal) linearized kinetic Boltzmann operator onto hydrodynamic modes in [30]. Indeed, the Riesz projection onto a hyperplane spanned by eigenvectors of L can be characterized as the unique projection that commutes with the operator L. A discussion of the Riesz projection through commutativity is provided in Appendix E. The dynamically optimal projection can thus be viewed as an alternative to the Riesz projection which differs from it for non-normal operators L.

IV. EXAMPLES

In this section, we discuss two examples of the dynamic projection: a two-dimensional shear flow and the linear three-component Grad's moment system. For these examples, the spectrum is known analytically and the dynamic projection can be calculated explicitly.
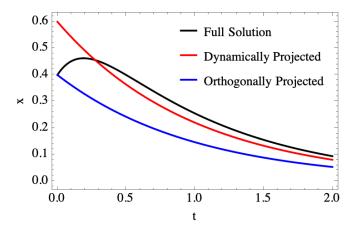
FIG. 2. Comparison of the x-components of the full solution (black) to the dynamically projected trajectory (red) and the orthogonally projected trajectory (blue) of system (13).

A. Two-Dimensional Shear System

As a first toy example, consider the two-dimensional linear system,

$$\frac{dx}{dt} = Lx, \quad L = \begin{pmatrix} -1 & \gamma \\ 0 & -\alpha \end{pmatrix}, \tag{13}$$

where $x = (x_1, x_2)^T$, for a damping parameter $\alpha > 1$ and a shear parameter $\gamma > 0$. The eigenvectors of the non-normal matrix L are,

$$\text{eig}(L) = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{(1-\alpha)^2 + \gamma^2}} \begin{pmatrix} \gamma \\ 1 - \alpha \end{pmatrix} \right\}, \tag{14}$$

corresponding to the eigenvalues -1 and $-\alpha$, respectively. The slow manifold, associated to the eigenvalue -1, is therefore simply given by the x_1-axis. The resolvent appearing in (10) reads

$$(L-1)^{-1} = \frac{(-1)}{2(1+\alpha)} \begin{pmatrix} 1+\alpha & \gamma \\ 0 & 2 \end{pmatrix}, \tag{15}$$

while the spectrally-weighted Gramian (8) reduces to a constant, $G = -1/2$. Thus, the dynamic projection (10) takes the explicit form,

$$\mathbb{P}_{\text{DOP}} x = \left(x_1 + \frac{\gamma}{1+\alpha} x_2 \right) \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{16}$$

For $\gamma = 0$, the projection (16) reduces to the standard orthogonal projection as expected. The dynamical projection for system (13) is shown in Figure 1. Figure 2 shows a direct comparison of the x_1-component of a full trajectory of system (13) to the dynamically projected and the orthogonally projected trajectories. We emphasize that the dynamically projected trajectory follows the overall trend towards equilibrium of the full solution much more closely over the full range of the time evolution. Geometrically, the functional (4) for this two-dimensional
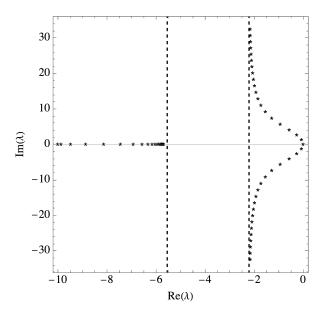

FIG. 3. Spectrum of the linear Grad system (17) for $\varepsilon = 0.1$ and various wave numbers. The fast real modes (star symbol on the real axis) accumulate at $-\frac{5}{9\varepsilon}$ (vertical dashed line at $\Re\lambda = -50/9$), while the slow complex conjugated modes (bell-shaped star symbol) accumulate at $\Re\lambda = -\frac{2}{9\varepsilon}$ (dashed line at $\Re\lambda = -20/9$) as $k \to \infty$.

example seeks an initial condition on the slow manifold such that the weighted area enclosed by a full solution and the projected solution is minimal.

B. Three-Component Grad's Moment System

Consider the three-component Grad's system,

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\frac{5}{3} \frac{\partial u}{\partial x}, \\ \frac{\partial u}{\partial t} &= -\frac{\partial p}{\partial x} - \frac{\partial \sigma}{\partial x}, \\ \frac{\partial \sigma}{\partial t} &= -\frac{4}{3} \frac{\partial u}{\partial x} - \frac{1}{\varepsilon} \sigma, \end{aligned} \tag{17}$$

for the pressure p, the velocity u and the stress σ, while $\epsilon > 0$ is a parameter representative of the Knudsen number. System (17) is a classical model in kinetic theory [18] whose Chapman–Enskog series has been summed explicitly in [16]. Equation (17) is defined on the whole real line with suitable decay at infinity. In the frequency space, system (17) reads,

$$\frac{d}{dt} \begin{pmatrix} \hat{p} \\ \hat{u} \\ \hat{\sigma} \end{pmatrix} = L_k \begin{pmatrix} \hat{p} \\ \hat{u} \\ \hat{\sigma} \end{pmatrix} \tag{18}$$

for the wave-number dependent matrix,

$$L_k = \begin{pmatrix} 0 & -\frac{5}{3}ik & 0 \\ -ik & 0 & -ik \\ 0 & -\frac{4}{3}ik & -\frac{1}{\varepsilon} \end{pmatrix}, \tag{19}$$

where i $= \sqrt{-1}$. We analyze the system (19) on the Hilbert space $H = \mathbb{C}^3$ with the complex inner product $\langle v, w \rangle = v \cdot w^*$ for $v, w \in \mathbb{C}^3$. The matrix L_k is not normal since

$$[L_k, L_k^{\dagger}] = \frac{1}{9\varepsilon} \begin{pmatrix} 16k^2\epsilon & 0 & 11k^2\varepsilon \\ 0 & -23k^2\epsilon & 21ik \\ 11k^2\epsilon & -21ik & 7k^2\varepsilon \end{pmatrix}. \qquad (20)$$

For a detailed spectral analysis of (19) along with its implications for the exact summation of the Chapman–Enskog series [27], we refer to [28]. The characteristic polynomial of (19),

$$P_k(\lambda) = -\lambda^3 - \frac{1}{\varepsilon}\lambda^2 - 3k^2\lambda - \frac{5}{3}\frac{k^2}{\varepsilon}, \qquad (21)$$

has a pair of complex conjugated eigenvalues $\{\lambda_{\rm ac}, \lambda_{\rm ac}^*\}$, called acoustic modes, and one real eigenvalue $\lambda_{\rm diff}$, called generalized diffusion mode. All eigenvalues have negative real part, see Figure 3. In particular, $\mathrm{spec}(L_k) = \mathrm{spec}(L_k^{\dagger})$. The eigenvectors of L_k are given by

$$Q = \begin{pmatrix} -1 - a_1 b_1 & -1 - a_2 b_2 & -1 - a_3 b_3 \\ ib_1 & ib_2 & ib_3 \\ 1 & 1 & 1 \end{pmatrix}, \qquad (22)$$

for

$$a_j = \frac{\lambda_j}{k}, \quad b_j = \frac{3}{4\varepsilon k}(1 + \varepsilon\lambda_j), \quad j = 1, 2, 3, \qquad (23)$$

where $\lambda_1 = \lambda_{\rm ac}$, $\lambda_2 = \lambda_{\rm ac}^*$ and $\lambda_3 = \lambda_{\rm diff}$. The slow manifold at each wave number is then given explicitly as the plane spanned by the two acoustic modes (the first

two columns of (22)),

$$\mathcal{M}_{\rm slow} = \mathrm{span}\left\{ \begin{pmatrix} -1 - a_1 b_1 \\ ib_1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 - a_2 b_2 \\ ib_2 \\ 1 \end{pmatrix} \right\}. \qquad (24)$$

The reduced dynamics on the slow manifold in spectral coordinates (6) is written,

$$\frac{d\xi_{\rm ac}}{dt} = \lambda_{\rm ac}\xi_{\rm ac}, \quad \frac{d\xi_{\rm ac}^*}{dt} = \lambda_{\rm ac}^*\xi_{\rm ac}^*, \qquad (25)$$

while the coordinate transform from spectral coordinates to the (p, u)-coordinates takes the form,

$$H = \begin{pmatrix} -1 - a_1 b_1 & -1 - a_2 b - 2 \\ ib_1 & ib_2, \end{pmatrix}. \qquad (26)$$

corresponding to the p- and u-components of the acoustic eigenvectors and its conjugate. The reduced model for the pressure and the velocity field at each wave number is then given by,

$$\frac{d}{dt}\begin{pmatrix} \hat{p}_k \\ \hat{u}_k \end{pmatrix} = T \begin{pmatrix} \hat{p}_k \\ \hat{u}_k \end{pmatrix}, \qquad (27)$$

for the transport matrix,

$$T = H^{-1}\Lambda H, \quad \Lambda = \begin{pmatrix} \lambda_{\rm ac} & 0 \\ 0 & \lambda_{\rm ac}^* \end{pmatrix}. \qquad (28)$$

The system (27) was earlier derived by exact summation of the Chapman–Enskog expansion in [16]. We also refer to [30] for a general treatment of the hydrodynamic closure for kinetic systems, including details on the coordinate change from spectral to physical coordinates and its relation to the Riesz projection. The resolvent at the negative eigenvalues can be readily calculated and takes the explicit form,

$$(L_k + \lambda^*)^{-1} = \frac{1}{6(\lambda^*)^2 + 10\varepsilon^2} \begin{pmatrix} -4\varepsilon k^2 + 4\lambda^* - 3\varepsilon(\lambda^*)^2 & 5ik - 3ik\lambda^* & 5\varepsilon k^2 \\ 3ik - 3i\varepsilon\lambda^* & 3\lambda^* - 3\varepsilon(\lambda^*)^2 & -3i\varepsilon k\lambda^* \\ 4\varepsilon k^2 & -4i\varepsilon k\lambda^* & -5\varepsilon k^2 - 3\varepsilon(\lambda^*)^2 \end{pmatrix}, \qquad (29)$$

for $\lambda \in \{\lambda_{\rm ac}, \lambda_{\rm ac}^*\}$, while the Gramian (8) reads,

$$G = \begin{pmatrix} \frac{1}{2\lambda_{\rm ac}}(|1 + a_1 b_1|^2 + |b_1|^2 + 1) & \frac{1}{2\Re\lambda_{\rm ac}}[(1 + a_1 b_1)^2 + b_1^2 + 1] \\ \frac{1}{2\Re\lambda_{\rm ac}}[(1 + a_2 b_2)^2 + b_2^2 + 1] & \frac{1}{2\lambda_{\rm ac}^*}(|1 + a_2 b_2|^2 + |b_2|^2 + 1) \end{pmatrix}, \qquad (30)$$

where we have used that $a_1^* = a_2$ and $b_1^* = b_2$. Combining (29) with (30) according to formula (10) then gives the dynamic projector for the three-component Grad system.

Figure 4 shows a comparison of pressure trajectories for the full system, the orthogonally projected initial condition and the dynamically projected initial condition. As
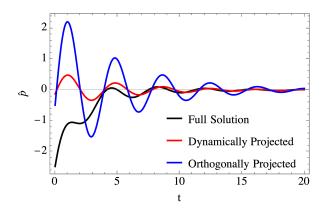
FIG. 4. Comparison of the pressure components of the full solution (black) to the dynamically projected trajectory (red) and the orthogonally projected trajectory (blue) of system (17) with initial condition orthogonal to the slow manifold at wave number $k = 1$. Even though the difference in initial pressure is comparably small, the orthogonally projected pressure trajectory exhibits much larger oscillations compared to the dynamically projected one.

expected, the dynamically projected pressure trajectory follows the transient oscillations of the full solution much more closely.

V. DISCUSSION

We introduced a novel method to optimally project dynamics onto a slow manifold in linear systems using a temporal variational principle. Unlike traditional projection methods, which may fail to capture transient behavior, the presented approach minimizes the cumulative deviation between the full system's trajectory and its reduced counterpart over time. This method accounts for transient dynamics, particularly in systems governed by non-normal operators where transients can be significant. The projection reduces to the familiar orthogonal projection in normal systems, ensuring consistency, and it offers a mathematically explicit formulation based on spectral properties of the system's operator.

We demonstrated the effectiveness of the dynamically optimal projection through explicit examples, including a two-dimensional shear flow and the linear three-

component Grad's system from kinetic theory. These case studies highlight how the new projection captures transient behaviors more accurately than standard projections, especially when dealing with systems exhibiting strong non-normality. Beyond its theoretical development, the projection has practical applications in reduced-order modeling, data assimilation, and the analysis of complex dynamical systems where transient dynamics play a critical role. Future work is anticipated to extend these ideas to non-invariant constraints and broader nonlinear contexts.

Potential applications of the dynamically optimal projection include problems where information about an entire subset of the full phase space is needed, such as sampling over a set of initial conditions or studying the evolution of specific regions of phase space. This is particularly relevant for optimal reduction on non-invariant submanifolds. A forthcoming paper will explore the application of the DOP to non-invariant constraints and their approximate reduced dynamics in greater detail.

A particularly relevant example of non-normal operators satisfying the assumptions in (2) are kinetic Boltzmann-type operators,

$$\mathcal{L}_{\boldsymbol{k}} = -\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{v} + L_Q, \tag{31}$$

where $L_Q^\dagger = L_Q$ is the linear collision operator and $\boldsymbol{k}$ is wave vector. Spectral properties of operators (31) are discussed in, e.g., [10]. We stress that (31) is non-normal and the degree of non-normality scales with wave number, $[\mathcal{L}_{\boldsymbol{k}}, \mathcal{L}_{\boldsymbol{k}}^\dagger] = \mathcal{O}(|\boldsymbol{k}|)$. The three-component Grad systems discussed in Section IV B is a toy model for kinetic operators of the form (31), mimicking some of its pertinent properties. Recently, detailed spectral information for the various linear kinetic operators, including the linear BGK and Shakhov kinetic models have been derived [29, 31]. These can serve as a basis for the application of the theory presented in this work.

[1] F. Achleitner, A. Arnold, V. Mehrmann, and E. A. Nigsch. Hypocoercivity in Hilbert spaces. *Journal of Functional Analysis*, 288(2):110691, 2025.

[2] A. C. Antoulas. *Approximation of large-scale dynamical systems*. SIAM, 2005.

[3] J. S. Baggett, T. A. Driscoll, and L. N. Trefethen. A mostly linear model of transition to turbulence. *Physics of Fluids*, 7(4):833–838, 1995.

[4] X. Cabré, E. Fontich, and R. De la Llave. The param-

eterization method for invariant manifolds I: manifolds associated to non-resonant subspaces. *Indiana University mathematics journal*, pages 283–328, 2003.

[5] X. Cabré, E. Fontich, and R. De la Llave. The parameterization method for invariant manifolds II: regularity with respect to parameters. *Indiana University mathematics journal*, pages 329–360, 2003.

[6] X. Cabré, E. Fontich, and R. De La Llave. The parameterization method for invariant manifolds III: overview

and applications. *Journal of Differential Equations*, 218(2):444–515, 2005.

[7] M. Cenedese, J. Axås, B. Bäuerlein, K. Avila, and G. Haller. Data-driven modeling and prediction of non-linearizable dynamics via spectral submanifolds. *Nature communications*, 13(1):872, 2022.

[8] C. Cercignani. *The Boltzmann Equation*. Springer, 1988.

[9] J. B. Conway. *A course in functional analysis*, volume 96. Springer, 2019.

[10] R. S. Ellis and M. A. Pinsky. The first and second fluid approximations to the linearized Boltzmann equation. *J. Math. Pures Appl*, 54(9):125–156, 1975.

[11] K. Engel. One-parameter semigroups for linear evolution equations, 2000.

[12] B. F. Farrell and P. J. Ioannou. Generalized stability theory. part I: Autonomous operators. *Journal of Atmospheric Sciences*, 53(14):2025–2040, 1996.

[13] N. Fenichel. Geometric singular perturbation theory for ordinary differential equations. *Journal of differential equations*, 31(1):53–98, 1979.

[14] C. W. Gear, T. J. Kaper, I. G. Kevrekidis, and A. Zagaris. Projecting to a slow manifold: Singularly perturbed systems and legacy codes. *SIAM Journal on Applied Dynamical Systems*, 4(3):711–732, 2005.

[15] A. N. Gorban. Model reduction in chemical dynamics: slow invariant manifolds, singular perturbations, thermodynamic estimates, and analysis of reaction graph. *Current Opinion in Chemical Engineering*, 21:48–59, 2018.

[16] A. N. Gorban and I. V. Karlin. Short-wave limit of hydrodynamics: a soluble example. *Physical review letters*, 77(2):282, 1996.

[17] A. N. Gorban and I. V. Karlin. Uniqueness of thermodynamic projector and kinetic basis of molecular individualism. *Physica A: Statistical Mechanics and its Applications*, 336(3):391–432, 2004.

[18] H. Grad. On the kinetic theory of rarefied gases. *Communications on pure and applied mathematics*, 2(4):331–407, 1949.

[19] S. Gugercin, A. C. Antoulas, and C. Beattie. H_2 model reduction for large-scale linear dynamical systems. *SIAM journal on matrix analysis and applications*, 30(2):609–638, 2008.

[20] G. Haller and S. Ponsioen. Nonlinear normal modes and spectral submanifolds: existence, uniqueness and use in model reduction. *Nonlinear dynamics*, 86:1493–1534, 2016.

[21] S. He and J. Jackson. A study of turbulence under conditions of transient flow in a pipe. *Journal of Fluid Mechanics*, 408:1–38, 2000.

[22] P. D. Hislop and I. M. Sigal. *Introduction to spectral theory: With applications to Schrödinger operators*, volume 113. Springer Science & Business Media, 2012.

[23] P. J. Holmes, J. L. Lumley, G. Berkooz, J. C. Mattingly, and R. W. Wittenberg. Low-dimensional models of coherent structures in turbulence. *Physics Reports*, 287(4):337–384, 1997.

[24] P. Ilg and I. Karlin. Validity of a macroscopic description in dilute polymeric solutions. *Physical Review E*, 62(1):1441, 2000.

[25] P. Ilg, I. V. Karlin, and H. C. Öttinger. Canonical distribution functions in polymer dynamics.(I). dilute solutions of flexible polymers. *Physica A: Statistical Mechanics and its Applications*, 315(3-4):367–385, 2002.

[26] J. John Soundar Jerome, J.-M. Chomaz, and P. Huerre. Transient growth in Rayleigh-Bénard-Poiseuille/Couette convection. *Physics of Fluids*, 24(4), 2012.

[27] I. V. Karlin and A. N. Gorban. Hydrodynamics from Grad's equations: What can we learn from exact solutions? *Annalen der Physik*, 514(10-11):783–833, 2002.

[28] F. Kogelbauer. Slow hydrodynamic manifolds for the three-component linearized Grad system. *Continuum Mechanics and Thermodynamics*, 32:1141–1146, 2020.

[29] F. Kogelbauer and I. Karlin. Exact hydrodynamic manifolds for the linear Boltzmann BGK equation I: spectral theory. *Continuum Mechanics and Thermodynamics*, 36(6):1685–1709, 2024.

[30] F. Kogelbauer and I. Karlin. Rigorous hydrodynamics from linear boltzmann equations and viscosity-capillarity balance. *Physical Review E*, 110(5):055105, 2024.

[31] F. Kogelbauer and I. Karlin. Spectral analysis and hydrodynamic manifolds for the linearized Shakhov model. *Physica D: Nonlinear Phenomena*, 459:134014, 2024.

[32] V. È. Ljance. Completely regular perturbation of a continuous spectrum. *Mathematics of the USSR-Sbornik*, 11(1):115, 1970.

[33] E. N. Lorenz. The slow manifold—what is it? *Journal of Atmospheric Sciences*, 49(24):2449–2451, 1992.

[34] R. MacKay. Slow manifolds. In *Energy localisation and Transfer*, pages 149–192. World Scientific, 2004.

[35] S. E. Otto, G. R. Macchio, and C. W. Rowley. Learning nonlinear projections for reduced-order modeling of dynamical systems using constrained autoencoders. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 33(11), 2023.

[36] H. Risken. The Fokker–Planck equation, 1996.

[37] A. Roberts. Appropriate initial conditions for asymptotic descriptions of the long term evolution of dynamical systems. *The ANZIAM Journal*, 31(1):48–75, 1989.

[38] A. Roberts. Computer algebra derives correct initial conditions for low-dimensional dynamical models. *Computer physics communications*, 126(3):187–206, 2000.

[39] J. C. Robinson and C. Pierre. Infinite-dimensional dynamical systems: An introduction to dissipative parabolic PDEs and the theory of global attractors. *Appl. Mech. Rev.*, 56(4):B54–B55, 2003.

[40] C. W. Rowley. Model reduction for fluids, using balanced proper orthogonal decomposition. *International Journal of Bifurcation and Chaos*, 15(03):997–1013, 2005.

[41] C. W. Rowley, T. Colonius, and R. M. Murray. Model reduction for compressible flows using POD and Galerkin projection. *Physica D: Nonlinear Phenomena*, 189(1-2):115–129, 2004.

[42] R. Temam. *Infinite-dimensional dynamical systems in mechanics and physics*, volume 68. Springer Science & Business Media, 2012.

[43] A. Tikhonov. On the dependence of the solutions of differential equations on a small parameter. *Matematicheskii sbornik*, 64(2):193–204, 1948.

[44] F. Waleffe. Transition in shear flows. Nonlinear normality versus non-normal linearity. *Physics of Fluids*, 7(12):3060–3066, 1995.

Appendix A: Properties of the quadratic form $\mathcal{E}$ (4)

In this section, we expand and simplify the temporal integral expressions in (4) and prove the existence of a unique minimizer. We stress that the integral expression in (4) exists because of the stability assumption of the spectrum of L and the resolvent estimate for the corresponding semi-group. Indeed, we recall that an eventually norm-continuous semigroup is exponentially stable if and only if $\Re\mathrm{spec}(L) < 0$, see [11]. For two vectors $\xi, \eta \in \mathbb{C}^n$ with $\xi = (\xi_1, ..., \xi_n)$ and $\eta = (\eta_1, ..., \eta_n)$, we denote the real standard inner product as

$$\xi \cdot \eta = \sum_{j=1}^{n} \xi_j \eta_j. \tag{A1}$$

The cumulative error contains two relevant contributions. The first one corresponds to the internal dynamics on the slow manifold,

$$\mathcal{E}_{\mathrm{inter}}(\xi) = \frac{1}{2} \int_0^\infty \|e^{tL} x_{\mathrm{slow}}(\xi)\|^2 \, dt, \tag{A2}$$

while the second contribution, given by an integrated inner product,

$$\mathcal{E}_{\mathrm{trans}}(x_0, \xi) = -\Re \int_0^\infty \langle e^{tL} x_0, e^{tL} x_{\mathrm{slow}}(\xi) \rangle \, dt, \tag{A3}$$

corresponds to interaction of the dynamics on the slow manifold with a general trajectory of (1). The full functional (4) is then given by,

$$\mathcal{E}(x_0, \xi) = \mathcal{E}_{\mathrm{inter}}(\xi) + \mathcal{E}_{\mathrm{trans}}(x_0, \xi) + \frac{1}{2} \int_0^\infty \|e^{tL} x_0\|^2 \, dt. \tag{A4}$$

Since the last summand in (A4) does not depend on ξ, the minimizer of $\mathcal{E}$ only depends on $\mathcal{E}_{\mathrm{inter}}$ and $\mathcal{E}_{\mathrm{trans}}$.

First, let us simplify the expressions (A2) and (A3) by resolving the temporal integrals. The internal dynamical contribution can be expanded as,

$$\begin{aligned}
\mathcal{E}_{\mathrm{inter}}(\xi) &= \frac{1}{2} \int_0^\infty \|e^{tL} x_{\mathrm{slow}}(\xi)\|^2 \, dt \\
&= \frac{1}{2} \sum_{j,k=1}^{n} \int_0^\infty \langle \xi_j e^{tL} \hat{x}_j, \xi_k e^{tL} \hat{x}_k \rangle \, dt \\
&= \frac{1}{2} \sum_{j,k=1}^{n} \int_0^\infty \langle \hat{x}_j, \hat{x}_k \rangle \xi_j \xi_k^* e^{(\lambda_j + \lambda_k^*) t} \, dt \\
&= \frac{1}{2} \sum_{j,k=1}^{n} \langle \hat{x}_j, \hat{x}_k \rangle \xi_j \xi_k^* \left[\frac{e^{(\lambda_j + \lambda_k^*) t}}{\lambda_j + \lambda_k^*} \right]_0^\infty \\
&= -\frac{1}{2} \sum_{j,k=1}^{n} \xi_j \left[\frac{\langle \hat{x}_j, \hat{x}_k \rangle}{\lambda_j + \lambda_k^*} \right] \xi_k^* \\
&= -\frac{1}{2} \xi \cdot G \xi^*,
\end{aligned} \tag{A5}$$

where

$$G_{jk} = \frac{\langle \hat{x}_j, \hat{x}_k \rangle}{\lambda_j + \lambda_k^*}, \tag{A6}$$

is a spectrally-weighted Gramian matrix associated to the eigenvectors of L as defined in (8). Note that we have only used the assumption $\Re\mathrm{spec}(L) < 0$ in the above calculations.

To evaluate the dynamical interaction contribution, we expand (A3) as

$$\begin{aligned}
\mathcal{E}_{\mathrm{trans}}(x_0, \xi) &= -\Re \int_0^\infty \langle e^{tL} x_0, e^{tL} x_{\mathrm{slow}}(\xi) \rangle \, dt \\
&= -\sum_{j=1}^{n} \Re \int_0^\infty \langle e^{tL} x_0, e^{\lambda_j t} \hat{x}_j \rangle \xi_j^* \, dt \\
&= -\sum_{j=1}^{n} \Re \int_0^\infty \langle x_0, e^{\lambda_j t} e^{tL^\dagger} \hat{x}_j \rangle \xi_j^* \, dt \\
&= -\sum_{j=1}^{n} \Re[\langle x_0, \eta_j \rangle \xi_j^*],
\end{aligned} \tag{A7}$$

where we have defined the vector η with the components,

$$\eta_j = \int_0^\infty e^{tL^\dagger} \hat{x}_j e^{\lambda_j t} \, dt. \tag{A8}$$

Integration by parts in (A8) gives,

$$\begin{aligned}
\eta_j &= \int_0^\infty e^{tL^\dagger} \hat{x}_j e^{\lambda_j t} \, dt \\
&= \frac{1}{\lambda_j} \left[e^{tL^\dagger} \hat{x}_j e^{\lambda_j t} \right]_0^\infty - \frac{1}{\lambda_j} \int_0^\infty L^\dagger e^{tL^\dagger} \hat{x}_j e^{\lambda_j t} \, dt \\
&= -\frac{\hat{x}_j}{\lambda_j} - \frac{1}{\lambda_j} L^\dagger \eta_j,
\end{aligned} \tag{A9}$$

where we have again used that $\Re\mathrm{spec}(L) < 0$. From (A9), we infer the following linear equation for η_j:

$$(L^\dagger + \lambda_j) \eta_j = -\hat{x}_j. \tag{A10}$$

Because of the stability assumption $\Re\mathrm{spec}(L) < 0$ and since $\mathrm{spec}(L^\dagger) = \mathrm{spec}(L)^*$, the complex number $-\lambda_j$ does not belong to the spectrum of $L^\dagger$. Hence, we can solve equation (A10) to give,

$$\eta_j = -(L^\dagger + \lambda_j)^{-1} \hat{x}_j. \tag{A11}$$

Thus, expression (A3) consequently simplifies as,

$$\begin{aligned}
\mathcal{E}_{\mathrm{trans}}(x_0, \xi) &= \sum_{j=1}^{n} \Re[\langle x_0, (L^\dagger + \lambda_j)^{-1} \hat{x}_j \rangle \xi_j^*] \\
&= \sum_{j=1}^{n} \Re[\langle (L + \lambda_j^*)^{-1} x_0, \hat{x}_j \rangle \xi_j^*] \\
&= \Re[I \cdot \xi^*],
\end{aligned} \tag{A12}$$

where we have defined a vector $I(x_0)$ with the components,

$$I_j(x_0) = \langle (L + \lambda_j^*)^{-1} x_0, \hat{x}_j \rangle, \ 1 \leq j \leq n. \quad \text{(A13)}$$

Clearly, from (A5) and (A12), the function $\xi \mapsto \mathcal{E}(x_0, \xi)$ defines a quadratic form for any $x_0 \in H$. To see that $\mathcal{E}$ is a convex function in ξ, we observe that the second-order term of $\mathcal{E}$ is given by (A2), which is strictly positive as the integral of a norm for any $\xi \in \mathbb{C}^n$. In particular, the spectrally-weighted Gramian matrix G is negative semi-definite, as it follows from (A5). Hence the quadratic part of $\mathcal{E}$ is positive definite and $\mathcal{E}$ is convex.

Appendix B: Explicit expression of the minimizer

In this section, we derive an explicit expression of the unique minimizer of the quadratic form $\xi \mapsto \mathcal{E}(x_0, \xi)$ on the slow manifold. To evaluate the complex derivative $\partial \mathcal{E} / \partial \xi$, we recall the following lemma (proof is included for completeness):

Lemma B.1. *Consider a quadratic form,*

$$f(\xi) = \Re[w \cdot \xi^*] - \frac{1}{2} \xi \cdot G \xi^*, \quad \text{(B1)}$$

for a Hermitian matrix $G = G^\dagger$ and complex vectors $w = w_R + \mathrm{i} w_I$, $\xi = \xi_R + \mathrm{i} \xi_I$. Then the complex derivative of f is given by

$$\frac{\partial f}{\partial \xi} = w - G^T \xi. \quad \text{(B2)}$$

Proof. First, we expand

$$\begin{aligned}
\Re[w \cdot \xi^*] &= w_R \cdot \xi_R + w_I \cdot \xi_I, \\
\xi \cdot G \xi^* &= \xi_R \cdot G \xi_R + \xi_I \cdot G \xi_I \\
&\quad - \mathrm{i} \xi_R \cdot G \xi_I + \mathrm{i} \xi_I \cdot G \xi_R,
\end{aligned} \quad \text{(B3)}$$

and set

$$G_S = \frac{1}{2}(G + G^T), \ G_A = \frac{1}{2}(G - G^T), \quad \text{(B4)}$$

to obtain,

$$\begin{aligned}
\frac{\partial f}{\partial \xi_R} &= w_R - G_S \xi_R + \mathrm{i} G_A \xi_I, \\
\frac{\partial f}{\partial \xi_I} &= w_I - G_S \xi_I - \mathrm{i} G_A \xi_R.
\end{aligned} \quad \text{(B5)}$$

Consequently, the complex derivative of (B1) is given by

$$\begin{aligned}
\frac{\partial f}{\partial \xi} &= \frac{\partial f}{\partial \xi_R} + \mathrm{i} \frac{\partial f}{\partial \xi_I} \\
&= (w_R - G_S \xi_R + \mathrm{i} G_A \xi_I) + \mathrm{i}(w_I - G_S \xi_I - \mathrm{i} G_A \xi_R) \\
&= w - G^T \xi_R - \mathrm{i} G^T \xi_I \\
&= w - G^T \xi.
\end{aligned} \quad \text{(B6)}$$

$\square$

Applying Lemma B.1 to $w = I(x_0)$ and G as defined in (A6), it follows that the critical point of the function (4) satisfies the equation

$$I(x_0) - G^T \xi = 0. \quad \text{(B7)}$$

As shown before, the critical point indeed defines a minimum thanks to the convexity of $\mathcal{E}$. The unique minimizer of (4) is then given explicitly as the solution to (B7):

$$\xi^{\min}(x_0) = (G^T)^{-1} I(x_0). \quad \text{(B8)}$$

Appendix C: Projection property of DOP (10)

Let us prove that parameter value that minimizes the cumulative dynamical error,

$$\xi^{\min}(x_0) := \underset{\xi \in \mathbb{C}^n}{\arg \min} \, \mathcal{E}(x_0, \xi), \quad \text{(C1)}$$

indeed, defines a projection onto the slow manifold as

$$\mathbb{P}_{\mathrm{DOP}} x = \sum_{i=1}^{n} \sum_{j=1}^{n} \hat{x}_i \left[(G^T)^{-1} \right]_{ij} \langle (L + \lambda_j^*)^{-1} x, \hat{x}_j \rangle. \quad \text{(C2)}$$

To ease notation, we bundle the eigenvectors into a vector of Hilbert space elements, $\hat{x} = (\hat{x}_1, ..., \hat{x}_n)$; elements of the slow manifold (5) are written more compactly using the real standard inner product (A1),

$$x_{\mathrm{slow}} = \xi \cdot \hat{x}. \quad \text{(C3)}$$

First, we evaluate the vector (A13) on the elements of the slow manifold (C3),

$$\begin{aligned}
I_j(\xi \cdot \hat{x}) &= \sum_{k=1}^{n} \xi_k \langle (L + \lambda_j^*)^{-1} \hat{x}_k, \hat{x}_j \rangle \\
&= \sum_{k=1}^{n} \left[\frac{\langle \hat{x}_k, \hat{x}_j \rangle}{\lambda_k + \lambda_j^*} \right] \xi_k \\
&= [G^T \xi]_j.
\end{aligned} \quad \text{(C4)}$$

This can be written in a form of an identity,

$$\xi = \left[G^T \right]^{-1} I(\xi \cdot \hat{x}), \quad \text{(C5)}$$

for any $\xi \in \mathbb{C}^n$, while operator (C2) takes the form,

$$\mathbb{P}_{\mathrm{DOP}} x = \hat{x} \cdot \left[G^T \right]^{-1} I(x). \quad \text{(C6)}$$

Second, applying the operator (C6) twice and using (C5) for $\xi = \left[G^T \right]^{-1} I(x)$, we find,

$$\begin{aligned}
\mathbb{P}_{\mathrm{dyn}}^2 x &= \mathbb{P}_{\mathrm{DOP}} \hat{x} \cdot \left[G^T \right]^{-1} I(x) \\
&= \hat{x} \cdot \left[G^T \right]^{-1} I \left(\hat{x} \cdot \left[G^T \right]^{-1} I(x) \right) \\
&= \hat{x} \cdot \left[G^T \right]^{-1} G^T \left[G^T \right]^{-1} I(x) \\
&= \hat{x} \cdot \left[G^T \right]^{-1} I(x) \\
&= \mathbb{P}_{\mathrm{DOP}} x.
\end{aligned} \quad \text{(C7)}$$

This proves Proposition 2.

Appendix D: DOP reduces to the canonical orthogonal projection for L normal

In this section, we assume that operator L is normal, $[L, L^\dagger] = 0$, and recall that eigenvectors of a normal operator corresponding to different eigenvalues are orthogonal. For the sake of presentation, we also assume that the eigenvalues λ_j are simple; thus, $\langle \hat{x}_i, \hat{x}_j \rangle = \delta_{ij}$. The case of repeated eigenvalues and generalized eigenvectors leads to Jordan blocks in the matrix G and can be treated analogously.

As an immediate consequence, the spectrally-weighted Gramian matrix (8) becomes diagonal,

$$G_{ij} = \frac{1}{2\Re\lambda_i}\delta_{ij}, \tag{D1}$$

where δ_{ij} is Kronecker's delta. Let us recall that, if $\hat{x}_j$ is unit-length eigenvector of L with eigenvalue λ_j, then it is also the eigenvector of the adjoint $L^\dagger$ with the eigenvalue λ_j^*,

$$L^\dagger \hat{x}_j = \lambda_j^* \hat{x}_j. \tag{D2}$$

Consequently, we find in (A13),

$$
\begin{aligned}
I_j(x) &= \langle x, (L^\dagger + \lambda_j)^{-1}\hat{x}_j \rangle \\
&= \langle x, (\lambda_j^* + \lambda_j)^{-1}\hat{x}_j \rangle \\
&= \frac{\langle x, \hat{x}_j \rangle}{\lambda_j + \lambda_j^*} \\
&= \frac{\langle x, \hat{x}_j \rangle}{2\Re\lambda_j}.
\end{aligned} \tag{D3}
$$

With (D1) and (D3), we obtain in (10),

$$
\begin{aligned}
\mathbb{P}_{\text{DOP}}x &= \sum_{i=1}^{n}\sum_{j=1}^{n} \hat{x}_i \, (2\Re\lambda_i \delta_{ij}) \, \frac{\langle x, \hat{x}_j \rangle}{2\Re\lambda_j} \\
&= \sum_{j=1}^{n} \langle x, \hat{x}_j \rangle \hat{x}_j.
\end{aligned} \tag{D4}
$$

Thus, the dynamically optimal projection becomes the standard orthogonal projection if operator L is normal.

Appendix E: Characterization of the Riesz projection through commutativity

In this section, we discuss properties of the Riesz projection onto a finite set of eigenvectors associated to isolated finite-order eigenvalues of a closed operator L. For completeness, we will show that the Riesz projection can be characterized through the commutation property with L. More specifically, we show that if a projection onto a finite set of linearly independent vectors commutes with L, then the projection is defined by eigenvectors of the adjoint of L.

Let $\hat{x}_1, ..., \hat{x}_n$ be linearly independent unit-length eigenvectors of a closed operator,

$$L\hat{x}_j = \lambda_j \hat{x}_j. \tag{E1}$$

For any compact subset of the spectrum, $\Lambda \subseteq \text{spec}(L)$, consisting only of isolated eigenvalues of finite multiplicity, the Riesz projection is defined as

$$\mathbb{P}_\Lambda = -\frac{1}{2\pi i} \oint_{\Gamma(\Lambda)} (z - L)^{-1} \, dz, \tag{E2}$$

where $\Gamma(\Lambda)$ is any simple contour encircling the set Λ once in the positive direction and lying entirely in the resolvent set of L. The linear operator (E2) indeed defines a projection [22], its range is exactly the linear subspace spanned by the eigenvectors associated to Λ. Moreover, the Riesz projection (E2) commutes with the operator L,

$$[\mathbb{P}_\Lambda, L] = 0. \tag{E3}$$

On the other hand, a general linear projection onto the subspace spanned by $\{\hat{x}_1, ..., \hat{x}_n\}$ is defined through a linear mapping $M \in L(H, \mathbb{C}^n)$ such that

$$\mathbb{P}_M x = \sum_{j=1}^{n} \hat{x}_j (Mx)_j. \tag{E4}$$

By the Riesz representation theorem [9], there exist vectors $\theta_1, .., \theta_n \in H$ such that

$$\mathbb{P}_M x = \sum_{j=1}^{n} \hat{x}_j \langle x, \theta_j \rangle. \tag{E5}$$

The linear mapping (E4) defines a projection (involution) provided that

$$\langle \hat{x}_i, \theta_j \rangle = \delta_{ij}, \tag{E6}$$

i.e., if $\{\theta_j\}$ is a dual set of $\{\hat{x}_j\}$. Moreover, the following commutation property characterizes the projection uniquely:

Lemma E.1. *Let $\mathbb{P}$ be a linear projection onto the linearly independent eigenvectors $\hat{x}_1, ..., \hat{x}_n$ of a closed operator L. Assume that $\mathbb{P}$ commutes with L. Then the vectors $\theta_1, .., \theta_n$ in the representation (E5) are eigenvectors of the adjoint of L,*

$$L^\dagger \theta_j = \lambda_j^* \theta_j. \tag{E7}$$

Proof. From

$$
\begin{aligned}
\mathbb{P}Lx &= \sum_{j=1}^{n} \langle Lx, \theta_j \rangle \hat{x}_j \\
&= \sum_{j=1}^{n} \langle x, L^\dagger \theta_j \rangle \hat{x}_j,
\end{aligned} \tag{E8}
$$

and

$$LPx = \sum_{j=1}^{n} \langle x, \theta_j \rangle L\hat{x}_j$$
$$= \sum_{j=1}^{n} \lambda_j \langle x, \theta_j \rangle \hat{x}_j, \quad \text{(E9)}$$

where we have used that $\hat{x}_j$'s are eigenvectors of L (E1), it follows from the linear independence of the $\hat{x}_j$'s that

$$\langle x, L^\dagger \theta_j \rangle = \lambda_j \langle x, \theta_j \rangle = \langle x, \lambda_j^* \theta_j \rangle. \quad \text{(E10)}$$

Since x is arbitrary, the statement (E7) follows. $\quad\square$

Thus, the commutation property (E3) defines the Riesz projection uniquely as the projection with the null-space $\ker \mathbb{P} = \{x \in H : x \perp \theta_j, \ L^\dagger \theta_j = \lambda_j^* \theta_j, \ 1 \le j \le n\}$ and range $\operatorname{im} \mathbb{P} = \operatorname{span}\{\hat{x}_j, \ 1 \le j \le n\}$.

Finally, we note that the Riesz representation (E5) of the DOP (10) corresponds to the following dual set,

$$\theta_i = \sum_{j=1}^{n} \left[(G^T)^{-1}\right]_{ij} (L^\dagger + \lambda_j)^{-1} \hat{x}_j, \ 1 \le i \le n. \quad \text{(E11)}$$

Indeed,

$$\langle \hat{x}_k, \theta_i \rangle = \sum_{j=1}^{n} \left[(G^T)^{-1}\right]_{ij} \langle \hat{x}_k, (L^\dagger + \lambda_j)^{-1} \hat{x}_j \rangle$$
$$= \sum_{j=1}^{n} \left[(G^T)^{-1}\right]_{ij} \langle (L + \lambda_j^*)^{-1} \hat{x}_k, \hat{x}_j \rangle$$
$$= \sum_{j=1}^{n} \left[(G^T)^{-1}\right]_{ij} \left[\frac{\langle \hat{x}_k, \hat{x}_j \rangle}{\lambda_k + \lambda_j^*}\right] \quad \text{(E12)}$$
$$= \sum_{j=1}^{n} \left[(G^T)^{-1}\right]_{ij} G_{kj}$$
$$= \sum_{j=1}^{n} \left[(G^T)^{-1}\right]_{ij} \left[G^T\right]_{jk}$$
$$= \delta_{ik}, \ 1 \le i, k \le n.$$

From the above considerations, it becomes apparent that DOP (10) is different from the Riesz projection (E2) for non-normal operators L, while at the same time, the DOP reduces to the Riesz projection and becomes the standard orthogonal projection (D4) when L is normal.