On the Convergence of an Efficient Algorithm for Kullback–Leibler Approximation of Spectral Densities

Augusto Ferrante, Federico Ramponi, *Member, IEEE*, and Francesco Ticozzi

***Abstract*—This paper deals with a method for the approximation of a spectral density function among the solutions of a generalized moment problem à la Byrnes/Georgiou/Lindquist. The approximation is pursued with respect to the Kullback–Leibler pseudo-distance, which gives rise to a convex optimization problem. After developing the variational analysis, we discuss the properties of an efficient algorithm for the solution of the corresponding dual problem, based on the iteration of a nonlinear map in a bounded subset of the dual space. Our main result is the proof of local convergence of the latter, established as a consequence of the central manifold theorem. Supported by numerical evidence, we conjecture that, in the mentioned bounded set, the convergence is actually *global*.**

***Index Terms*—Kullback–Leibler pseudo-distance, spectral estimation.**

I. Introduction

DURING the last decade a broad research program on the interplay between *(generalized) moment problems* and *analytic interpolation problems* with complexity constraints, *robust control*, *approximation* and *estimation* of spectral density functions has been carried out by C. I. Byrnes, T. Georgiou, and A. Lindquist and their collaborators and epigones [3]–[12], [15], [17], [18], [20]–[26], [28], [29], [31], [32]. Moment problems have a long history and have been at the heart of many mathematical and engineering problems in the past century, see, e.g., [1], [33] and the references therein. Only with recent developments of the above-mentioned research program, however, the parametrization of solutions in the presence of additional constraints on the complexity has been satisfactorily addressed [11]. This result, that has been possible thanks to a suitable variational formulation, is of key interest in control engineering. In fact, the well-known relation between moment problems and Nevanlinna-Pick interpolation problems allows for solutions of $\mathcal{H}_\infty$ control problems that include a bound on the complexity of the controller, which is of paramount practical importance [3], [8], [19]. Similar considerations apply to the *covariance extension* problem [9], [26]. Among practical applications, we mention in particular signal and image processing [22]–[25], and Biomedical Engineering [30]. These applications are based on a *spectral estimation* procedure that hinges on *optimal approximation* of a given spectral density Ψ with linear *integral constraints*. The latter may be viewed as constraints on a finite number of "generalized" moments of the spectrum and represent an *a posteriori* knowledge on the steady-state covariance of a bank of filters that is designed to estimate the unknown spectral density Φ. The spectral density Ψ, on the other hand, represents a prior knowledge on Φ, see Section II for more details. As discussed in [6], [26], [31], this optimal approximation leads to a tunable spectral estimation procedure that provides high resolution estimates in prescribed frequency bands even in presence of a short observation record. The degrees of freedom provided by both the presence of Ψ and the structure of the bank of filters — which determines high-resolution regions — make the above-mentioned method particularly flexible. As we shall see, a key feature of the method is that the primal optimization problem is strictly convex and can be solved in closed form. Furthermore, as long as the prior spectral density Ψ is rational, it yields a solution which is also rational, and comes with an *a priori* bound on its complexity.

On the other hand the *dual problem*, which is finite-dimensional, poses a numerical challenge. In fact, the dual variable is an Hermitian matrix and, as discussed in [26], its reparametrization may lead to a loss of convexity. Moreover, the gradient of the dual functional is unbounded in the neighborhood of the boundary, leading to serious numerical difficulties in practical implementations. As a consequence any *gradient-based* numerical method is severely affected by heavy computational burden, due to a large number of back-stepping iterations.

In order to avoid this major computational problem, a nonlinear matricial iteration has been introduced in [31]. The latter has manifold good properties, namely it is surprisingly simple, it is reliable, it is not computationally demanding (it does not involve either back-stepping or the computation and inversion of Hessians), and it *seems* to converge globally — in a prescribed set — to an optimal solution of the dual problem, with linear rate of convergence. Providing any explanation for its convergence, however, has so far been a challenging open problem.

Our main contribution is the proof that this iteration is locally convergent to the manifold of solutions of the dual problem. In

Manuscript received May 28, 2009; revised November 26, 2009; accepted June 10, 2010. Date of publication July 15, 2010; date of current version March 09, 2011. This work was supported in part by the Italian Ministry for Education and Research (MIUR) under PRIN grant "Identification and Robust Control of Industrial Systems", by the Dept. Ingegneria dell'Informazione (University of Padova), under project "QUINTET" and by the University of Padova under Projects CPDA080209/08 and "Quantum Future." Recommended by Associate Editor Z. Wang.

A. Ferrante and F. Ticozzi are with the Dipartimento di Ingegneria dell'Informazione, Università di Padova, Padova 35131, Italy (e-mail: augusto@dei.unipd.it; ticozzi@dei.unipd.it).

F. Ramponi is with the Institut für Automatik, ETH Zürich, Zürich 8092, Switzerland (e-mail: ramponif@control.ee.ethz.ch).

Digital Object Identifier 10.1109/TAC.2010.2057171

0018-9286/$26.00 © 2010 IEEE

doing this, we analyze in depth the dynamical properties of this matricial iteration and establish other relevant results that clarify its connection with the structures of the dual space and some features of the dual solutions. The second main contribution of the paper is to show that, by resorting to a spectral factorization method, the proposed iteration may be implemented in an effective way. In fact, by exploiting the system-theoretic features of the iteration, we establish a procedure in which each iteration of the algorithm requires only the solution of a Riccati equation and of a Lyapunov equation, for which robust and efficient algorithms are available. Our results lay a solid ground for efficiently applying this spectral approximation methods to signal analysis and to various fields of control engineering.

The paper is organized as follows. In Section II we give a proper mathematical statement of the problem, and proceed by recalling some relevant facts from the literature as well as establishing some preliminary results. Section III contains our main result: Local convergence of the matricial iteration to one of the fixed points that solve the dual problem. The proof is rather articulated and involves many different tools from linear and non-linear systems theory, including the *Center Manifold Theorem*. Subdivision of Section III into many subsections provides a roadmap of the various parts of the proof. As a byproduct, we also obtain many relevant results on the iteration and its linearization. In Section IV we describe how the proposed iteration may be implemented in an effective numerical way. Section V illustrates some results obtained from simulations and a conjecture, and concludes the paper with final remarks and future perspective.

Notation

We denote by $\mathbb{H}_n$ the set of Hermitian matrices of dimension n. Given a complex matrix A, A^* denotes the transpose conjugate of A, while, for a matrix valued function $\chi(z)$ in the complex variable z, $\chi^*(z)$ denotes the analytic continuation of the function that for $|z| = 1$ equals the transpose conjugate of $\chi(z)$. Thus, for a matrix-valued rational function $\chi(z) = H(zI - F)^{-1}G + J$, we have $\chi^*(z) = G^*(z^{-1}I - F^*)^{-1}H^* + J^*$. We denote by $\mathbb{T}$ the unit circle in the complex plane $\mathbb{C}$ and by $C(\mathbb{T})$ the set of complex-valued continuous functions on $\mathbb{T}$. $C_+(\mathbb{T})$ denotes the subset of $C(\mathbb{T})$ whose elements are real-valued positive functions. Elements in $C_+(\mathbb{T})$ will be thought of as *spectral densities*.

II. Problem Formulation and Background Material

Consider the transfer function $G(z) = (zI - A)^{-1}B$, of the system

$$x(t+1) = Ax(t) + By(t)$$

where $A \in \mathbb{C}^{n\times n}$ is a stability matrix, i.e., has all its eigenvalues in the open unit disc, $B \in \mathbb{C}^{n\times 1}$, and (A, B) is a *reachable pair*.

The transfer function $G(z)$ models a bank of filters fed by a stationary process $y(t)$ of unknown spectral density $\Phi(z)$. We assume that we know (or that we can reliably estimate) the steady-state covariance Σ of the state x of the filter. Based on Σ and on an *a priori* information in the form of a prior spectral density $\Psi(z)$, we want to estimate the spectral density $\Phi(z)$.

We will consider the Kullback-Leibler index as a measure of the difference between spectral densities Ψ and Φ in $C_+(\mathbb{T})$

$$\mathbb{S}(\Psi\|\Phi) = \int \Psi \log\left(\frac{\Psi}{\Phi}\right) = \int_{-\pi}^{\pi} \Psi(\mathrm{e}^{\mathrm{j}\vartheta}) \log\left(\frac{\Psi(\mathrm{e}^{\mathrm{j}\vartheta})}{\Phi(\mathrm{e}^{\mathrm{j}\vartheta})}\right) \frac{d\vartheta}{2\pi}.$$

The above notation, where integration takes place on the unit circle and with respect to the normalized Lebesgue measure, is used throughout the whole paper.

As in [26], we consider the following

Problem 2.1: (Approximation problem) Let $\Psi \in C_+(\mathbb{T})$, and let $\Sigma \in \mathbb{C}^{n\times n}$ satisfy $\Sigma = \Sigma^* > 0$. Find $\hat{\Phi}$ that solves

$$\text{minimize} \quad \mathbb{S}(\Psi\|\Phi) \tag{1}$$

$$\text{over} \quad \left\{\Phi \in C_+(\mathbb{T}) \mid \int G\Phi G^* = \Sigma\right\}. \tag{2}$$

Remark 2.1: Notice that in order to guarantee that $\mathbb{S}(\cdot\|\cdot) \geq 0$ we need that the zeroth-order moment of its arguments is the same, i.e., if $\int \Psi = \int \Phi$ then $\mathbb{S}(\Psi\|\Phi) \geq 0$. For the minimization problem to make sense as an *approximation*, we need precisely this condition. In [26] it is shown that, when A is singular, the zeroth-order moment of all the spectra Φ compatible with the constraint $\int G\Phi G^* = \Sigma$ is constant, say $\int \Phi \equiv \alpha$. Without either of the singularity of A or the equality of the zeroth-order moments of Ψ *and* all of the Φ's, it is not clear at all if $\mathbb{S}$ serves as a pseudo-distance, even if the minimization problem continues to be valid. In view of this consideration, we require from now on that A has at least one eigenvalue at the origin, and that Ψ is rescaled accordingly in order to obtain $\int \Psi = \alpha$. Hence, what we approximate is the "shape" of Ψ, not Ψ itself. If A is non-singular, it is still possible to consider a weighted version of the Kullback-Leibler pseudo-distance in such a way that the problem maintains the meaning of spectral approximation.

Remark 2.2: Notice that minimizing $\mathbb{S}(\Psi\|\Phi)$ may appear unusual with respect to the minimization of $\mathbb{S}(\Phi\|\Psi)$, which is more common in the literature on statistics, probability, and information theory. In our setting, however, this choice has the advantage of leading to an optimal solution having the form of a rational function with an *a priori* bound on the McMillan degree. Besides, by selecting $\Psi \equiv 1$, this choice includes, as special case, the maximum entropy spectral density [22].

Remark 2.3: To simplify the writing we can, without loss of generality, normalize Σ and Ψ. Indeed, if $\Sigma \neq I$, it suffices to replace G by $G' := \Sigma^{-1/2}G$ and (A, B) with $(A' = \Sigma^{-1/2}A\Sigma^{1/2}, B' = \Sigma^{-1/2}B)$ to obtain an equivalent problem where $\Sigma = I$. In a similar fashion, if $\int \Phi \equiv \int \Psi = \alpha \neq 1$ (compare with Remark 2.1), let $\Psi' := \Psi/\alpha$ and $G' = \sqrt{\alpha}\, G$. Then, to any solution Φ to the moment problem $\int G\Phi G^* = \Sigma$ there corresponds a solution Φ' to the problem $\int G'\Phi' G'^* = \Sigma$, where $\Phi' = \Phi/\alpha$. It is immediate to check that $\mathbb{S}(\Psi'\|\Phi') = \mathbb{S}(\Psi\|\Phi)/\alpha$, which ensures that the positivity of the pseudo-distance is preserved. Therefore, we can assume that $\int \Psi = 1$.

The first issue one needs to worry about is the existence of $\Phi \in C_+(\mathbb{T})$ satisfying constraint (2). It has been shown that the following conditions are equivalent [26]:

1) The family of Φ satisfying constraint (2) is nonempty.
2) there exists $H \in \mathbb{C}^{1\times n}$ such that $I - AA^* = BH + H^*B^*$.

A third equivalent condition is based on the linear operator Γ defined as follows. Let $\mathbb{H}_n$ be the space of Hermitian matrices of dimension n, and consider

$$\begin{aligned}\Gamma: \quad C(\mathbb{T}) &\longrightarrow \mathbb{H}_n \\ \Phi &\mapsto \int G\Phi G^*. \end{aligned} \tag{3}$$

It is clear (recall that $\Sigma = I$) that there exists $\Phi \in C(\mathbb{T})$ satisfying (2) if and only if

$$I \in \operatorname{Range}\Gamma. \tag{4}$$

Indeed, it has been shown [17] that (4) is necessary and sufficient for the family of Φ in (2) to be nonempty. Thus, (4) will be a standing assumption for this paper. We endow the space $\mathbb{H}_n$ of Hermitian matrices with the inner product $\langle P, Q\rangle := \operatorname{tr}(PQ)$. The orthogonal complement of $\operatorname{Range}\Gamma$ (defined with respect to this inner product) has been shown in [17] to be given by

$$\operatorname{Range}\Gamma^{\perp} = \left\{X \in \mathbb{H}_n : G^*(e^{j\vartheta})XG(e^{j\vartheta}) = 0 \ \forall \vartheta \in [0, 2\pi]\right\}.$$

Notice that, in view of (4), we have $\operatorname{tr}(X) = \langle X, I\rangle = 0$, $\forall X \in \operatorname{Range}\Gamma^{\perp}$. Given (4) and hence the existence of spectral densities satisfying the constraints, we next focus on the minimization Problem 2.1.

A. Variational Analysis

To solve Problem 2.1, we consider a *matrix Lagrange multiplier* $\Lambda \in \mathbb{H}_n$ satisfying $G^*\Lambda G > 0$ on all of $\mathbb{T}$, and define the *Lagrangian functional*

$$\begin{aligned} L(\Phi, \Lambda) &:= \mathbb{S}(\Psi\|\Phi) + \left\langle \Lambda, \int G\Phi G^* - I \right\rangle \\ &= \mathbb{S}(\Psi\|\Phi) + \int G^*\Lambda G\Phi - \operatorname{tr}(\Lambda). \end{aligned} \tag{5}$$

This functional is easily seen to be strictly convex. Therefore *unconstrained* minimization of $L(\Phi, \Lambda)$ can be achieved if there exists a Φ such that the directional derivative $D(L(\Phi, \Lambda), \delta\Phi)$ is annihilated along all directions $\delta\Phi \in C(\mathbb{T})$. Such a Φ indeed exists and is given by the following expression in terms of the Lagrange multiplier Λ

$$\Phi_{\Lambda} = \frac{\Psi}{G^*\Lambda G}. \tag{6}$$

Now, it is clear that if $\Lambda_\circ = \Lambda_\circ^*$ satisfies

$$G^*\Lambda_\circ G > 0, \quad \forall e^{j\vartheta} \in \mathbb{T}, \tag{7a}$$

$$\int G\frac{\Psi}{G^*\Lambda_\circ G}G^* = I \tag{7b}$$

then

$$\Phi_\circ := \Phi_{\Lambda_\circ} = \frac{\Psi}{G^*\Lambda_\circ G} \tag{8}$$

is optimal for Problem 2.1. As for many optimization problems, the most delicate issue is *existence* of a $\Lambda_\circ$ satisfying (7). This issue has been addressed in [16], [26] where the following result has been proven.

Theorem 2.1: There exist matrices $\Lambda_\circ = \Lambda_\circ^*$ such that (7) hold. For any such a $\Lambda_\circ$, $\Phi_\circ$ given by (8) is the unique solution of the Approximation Problem (2.1).

Remark 2.4: Since Problem (2.1) admits a unique solution, if $\Lambda_\circ$ and $\Lambda_\circ'$ are two matrices satisfying (7), then $\frac{\Psi}{G^*\Lambda_\circ G} = \frac{\Psi}{G^*\Lambda_\circ' G}$ so that we clearly have $G^*(\Lambda_\circ - \Lambda_\circ')G \equiv 0$, or equivalently, $\Lambda_\circ - \Lambda_\circ' \in \operatorname{Range}\Gamma^{\perp}$. Conversely, it is clear that if $\Lambda_\circ$ satisfies conditions (7) then $(\Lambda_\circ + X)$ also satisfies conditions (7) for any $X \in \operatorname{Range}\Gamma^{\perp}$. Thus, the family $\mathcal{L}_\circ$ of all solutions of (7) is an affine space that may be parametrized in terms of an arbitrary solution $\Lambda_\circ$, as

$$\mathcal{L}_\circ = \{\Lambda_\circ + X;\ X \in (\operatorname{Range}\Gamma)^{\perp}\}. \tag{9}$$

Moreover, for any $\Lambda_\circ$ satisfying (7), we have $\operatorname{tr}\int G\frac{\Psi}{G^*\Lambda_\circ G}G^*\Lambda_\circ = \operatorname{tr}\Lambda_\circ$ and, using the cyclic property of the trace we immediately get $\operatorname{tr}\Lambda_\circ = 1$.

By duality theory, a $\Lambda_\circ$ satisfying (7) may be computed by maximization of the dual functional $\Lambda \mapsto \inf\{L(\Phi, \Lambda)|\Phi \in C_+(\mathbb{T})\}$. The latter may be explicitly written as

$$\Lambda \mapsto L\left(\frac{\Psi}{G^*\Lambda G}, \Lambda\right) = \int \Psi \log G^*\Lambda G - \operatorname{tr}(\Lambda) + \int \Psi. \tag{10}$$

Consider now the maximization of the dual functional (10) over the set

$$\mathcal{L}_+ := \{\Lambda = \Lambda^* | G^*\Lambda G > 0, \forall e^{j\vartheta} \in \mathbb{T}\}. \tag{11}$$

Let $\mathbb{J}_\Psi(\Lambda) := -\int \Psi \log G^*\Lambda G + \operatorname{tr}(\Lambda)$. The dual problem is then equivalent to

$$\text{minimize} \quad \{\mathbb{J}_\Psi(\Lambda)|\Lambda \in \mathcal{L}_+\}. \tag{12}$$

As discussed in the Introduction, the bottleneck of the whole theory and of its numerous applications is now the numerical computation of a $\Lambda_\circ$ satisfying (7). To this aim the following algorithm has been proposed in [31] and further discussed in [16].

B. Iterative Algorithm

For $\Lambda \geq 0$, let

$$\Theta(\Lambda) := \int \Lambda^{1/2} G\left[\frac{\Psi}{G^*\Lambda G}\right] G^*\Lambda^{1/2}. \tag{13}$$

It has been shown in [31] that Θ is a map from *density matrices* to density matrices, i.e., if Λ is a positive semi-definite Hermitian matrix with trace equal to 1, then $\Theta(\Lambda)$ has the same properties. Density matrices have long been studied in statistical quantum mechanics, representing quantum states in the presence of uncertainty [34]. Moreover, Θ maintains positive definiteness, i.e., if $\Lambda > 0$, then $\Theta(\Lambda) > 0$. In addition to this, the following holds:

Proposition 2.1: The matrix $\Theta(\Lambda)$ has the same rank and the same kernel of the matrix Λ.

Proof: By taking into account that $\ker(\Lambda^{1/2}) = \ker(\Lambda)$, it is immediate to check that if $v \in \ker[\Lambda]$ then $v \in \ker[\Theta(\Lambda)]$. Conversely, it is sufficient to prove that

$$\int G\left[\frac{\Psi}{G^*\Lambda G}\right]G^* > 0. \tag{14}$$

Indeed, if this is the case and $v \in \ker[\Theta(\Lambda)]$ then $\Lambda^{1/2}v = 0$, so that $v \in \ker[\Lambda]$. To prove (14), we observe that $\frac{\Psi}{G^*\Lambda G}$ is continuous and strictly positive on $\mathbb{T}$ and hence has a positive minimum there. It is therefore sufficient to show that $\int GG^*$ is positive definite. The latter integral is the steady-state covariance of the filter G driven by normalized white noise, i.e., the unique solution Ξ of the discrete-time Lyapunov equation $\Xi - A\Xi A^* = BB^*$. In view of the controllability of the pair (A, B), it is clear that Ξ is positive definite. □

Consider the sequence $\{\Lambda_k\}$ produced by the following iteration:

$$\Lambda_{k+1} = \Theta(\Lambda_k) \tag{15}$$

with an arbitrary initial condition $\Lambda_0 > 0$. Notice that, since each Λ_k is positive definite and has trace equal to 1, we also have $\Lambda_k \leq I\ \forall k > 0$. If the sequence $\{\Lambda_k\}$ converges to a limit point $\hat{\Lambda} > 0$ then such a $\hat{\Lambda}$ is a fixed point for the map Θ in (13)

$$\hat{\Lambda} := \int \hat{\Lambda}^{1/2}G\left[\frac{\Psi}{G^*\hat{\Lambda}G}\right]G^*\hat{\Lambda}^{1/2}. \tag{16}$$

By multiplying the latter by $\hat{\Lambda}^{-1/2}$ on both sides, it is clear that $\hat{\Lambda}$ satisfies (7) and hence provides a solution of Problem 2.1.

Notice that, even if all Λ_k are positive definite, it may happen that the sequence $\{\Lambda_k\}$ converges to a limit point $\hat{\Lambda}_s$ which is singular. In this case, it is not guaranteed that $\hat{\Lambda}_s$ satisfies (7).

We observe that if $\Lambda_\circ > 0$ is a fixed point of Θ, then for any $\Lambda_\perp \in (\mathrm{Range}\,\Gamma)^\perp$, $\Lambda_\circ + \Lambda_\perp$ is also a fixed point of Θ, as long as $\Lambda_\circ + \Lambda_\perp \geq 0$. In fact, in view of Remark 2.4, we have

$$\Theta(\Lambda_\circ+\Lambda_\perp) = (\Lambda_\circ+\Lambda_\perp)^{1/2}I(\Lambda_\circ+\Lambda_\perp)^{1/2} = \Lambda_\circ+\Lambda_\perp. \tag{17}$$

In a wide series of simulations, we have observed that Λ_k always converges to a limit point. In only one case such a limit point was a singular matrix. Also in that case, however, the limit point satisfied (7) and hence provided a solution of Problem 2.1.

III. Proof of Local Convergence

In this section we prove the following result that is the main contribution of the paper.

Theorem 3.1: The intersection between the set of solutions of (7) and the cone of positive definite matrices is a non-empty manifold that is locally asymptotically stable for the iteration (15).

A. Existence of a Positive Definite $\Lambda_\circ$

Once again, the first issue that must be addressed is an existence result: We have to show that

$$\mathcal{L}_{\circ+} := \{\Lambda_\circ = \Lambda_\circ^* > 0 : \Lambda_\circ \in \mathcal{L}_\circ\} \neq \varnothing \tag{18}$$

where $\mathcal{L}_\circ$ is defined in (9). To this aim we need a preliminary result in the same vein of Lemma 9 in [26]. The latter has been established in a slightly different setting and using an abstract functional-analytic approach. We will instead use a direct algebraic approach that provides a constructive proof.

Lemma 3.1: If $G^*\Lambda_\circ G > 0$, $\forall\, e^{j\vartheta} \in \mathbb{T}$, then there exists a vector $C_\circ \in \mathbb{C}^{n\times 1}$ such that $G^*\Lambda_\circ G = G^*C_\circ C_\circ^* G$.

Proof: As shown in Lemma A.1 in the Appendix, we can obtain a decomposition $G^*\Lambda G = W^*W$, where the (right) spectral factor $W(z)$ is given by (59). Denoting by P the stabilizing solution of the Riccati equation (60), by using (63), W may be explicitly expressed in the form

$$W = (B^*PB)^{-1/2}B^*P\left(A(zI-A)^{-1}+I\right)B. \tag{19}$$

It is immediate to check that $A(zI-A)^{-1}+I = z(zI-A)^{-1}$ so that

$$W = z(B^*PB)^{-1/2}B^*P(zI-A)^{-1}B \tag{20}$$

and thus $G^*\Lambda G = W^*W = W_1^*W_1$, with $W_1 := z^{-1}W = (B^*PB)^{-1/2}B^*P(zI-A)^{-1}B$. Therefore, the vector $C_\circ$ exists and may be explicitly written as $C_\circ = \left((B^*PB)^{-1/2}B^*P\right)^*$. □

Theorem 3.2: The set $\mathcal{L}_{\circ+}$ defined by (18) is nonempty and it is an open convex subset of the affine space $\mathcal{L}_\circ$.

Proof: Let $\Lambda_\circ \in \mathcal{L}_\circ$ (recall that $\mathcal{L}_\circ \neq \varnothing$ due to Theorem 2.1) so that $G^*\Lambda_\circ G > 0$, $\forall\, e^{j\vartheta} \in \mathbb{T}$. From Lemma 3.1 we know that this implies the existence of a vector $C_\circ \in \mathbb{C}^{n\times 1}$ such that $G^*\Lambda_\circ G = G^*C_\circ C_\circ^* G$. On $\mathbb{T}$, $G^*\Lambda_\circ G$ is continuous and positive. Thus, $\mu := \min\{G(z)^*\Lambda_\circ G(z) : z \in \mathbb{T}\} > 0$. Similarly, on $\mathbb{T}$, G^*G is continuous. Thus, $\nu := \max\{G(z)^*G(z) : z \in \mathbb{T}\}$ is finite. Let $\varepsilon := \mu/4\nu$. Clearly,

$$\begin{aligned} G(z)^*\left(\frac{1}{2}\Lambda_\circ - \varepsilon I\right)G(z) &= \frac{1}{2}G(z)^*\Lambda_\circ G(z) - \varepsilon G(z)^*G(z) \\ &\geq \frac{\mu}{2} - \frac{\mu}{4\nu}\nu = \frac{\mu}{4} > 0, \quad \forall\, z \in \mathbb{T}. \end{aligned} \tag{21}$$

Hence, exploiting again Lemma 3.1, we conclude that there exists $C_1 \in \mathbb{C}^{n\times 1}$ such that

$$G^*\left(\frac{1}{2}\Lambda_\circ - \varepsilon I\right)G = G^*C_1C_1^*G.$$

Therefore we have

$$\begin{aligned} G^*\Lambda_\circ G &= \frac{1}{2}G^*C_\circ C_\circ^* G + \frac{1}{2}G^*C_\circ C_\circ^* G + \varepsilon G^*G - \varepsilon G^*G \\ &= G^*\left(\frac{1}{2}C_\circ C_\circ^* + \varepsilon I\right)G + \frac{1}{2}G^*\Lambda_\circ G - \varepsilon G^*G \\ &= G^*\left(\frac{1}{2}C_\circ C_\circ^* + \varepsilon I\right)G + G^*\left(\frac{1}{2}\Lambda_\circ - \varepsilon I\right)G \\ &= G^*\left(\frac{1}{2}C_\circ C_\circ^* + \varepsilon I + C_1C_1^*\right)G = G^*\Lambda_{\circ+}G \end{aligned}$$

where $\Lambda_{\circ+} := (1/2)C_\circ C_\circ^* + \varepsilon I + C_1C_1^*$ is positive definite so that $\Lambda_{\circ+} \in \mathcal{L}_{\circ+}$. Thus, $\mathcal{L}_{\circ+}$ is non-empty. The fact that $\mathcal{L}_{\circ+}$ is an open convex subset of $\mathcal{L}_\circ$ is an immediate consequence of the fact that the cone of positive definite matrices is open and convex together with the fact that $\mathcal{L}_\circ$ is an affine space. □

B. Linearization

Given that $\mathcal{L}_{\circ+}$ is non-empty, we can now pick a point $\Lambda_\circ \in \mathcal{L}_{\circ+}$ and analyze the map Θ in a neighborhood of $\Lambda_\circ$. To this aim we linearize the map Θ, namely we compute the directional derivative of Θ at $\Lambda_\circ$ in the direction specified by an arbitrary Hermitian matrix X

$$D(\Theta(\Lambda_\circ), X) := \lim_{\varepsilon \to 0} \frac{\Theta(\Lambda_\circ + \varepsilon X) - \Theta(\Lambda_\circ)}{\varepsilon}.$$

In order to find an explicit form for this derivative, we first need an expression for $D(\Lambda_\circ^{1/2}, X)$.

1) Derivative of the Matrix Square Root: For a given function $\varrho : \mathbb{H}_n \to \mathbb{H}_n$, let us take the directional derivative of $(\varrho(\Lambda))^2$ in the direction X. The chain rule gives

$$D(\varrho(\Lambda)^2; X) = D(\varrho(\Lambda); X)\, \varrho(\Lambda) + \varrho(\Lambda)\, D(\varrho(\Lambda); X).$$

Now if $\varrho(\Lambda) = \Lambda^{1/2}$, we have $\varrho(\Lambda)^2 = \Lambda$ so that clearly $D(\varrho(\Lambda)^2; X) = D(\Lambda; X) = X$. In conclusion, we get that the derivative $D(\Lambda^{1/2}; X)$ is the solution of the following Lyapunov equation:

$$D(\Lambda^{1/2}; X)\, \Lambda^{1/2} + \Lambda^{1/2}\, D(\Lambda^{1/2}; X) = X. \tag{22}$$

2) Derivative of Θ: Let us take the variation of (13) in a direction X. By applying the chain rule we get

$$\begin{aligned} D(\Theta(\Lambda); X) =& D(\Lambda^{1/2}; X) \int \frac{G\Psi G^*}{G^*\Lambda G}\Lambda^{1/2} \\ &+ \Lambda^{1/2} D\left(\int \frac{G\Psi G^*}{G^*\Lambda G}; X\right)\Lambda^{1/2} \\ &+ \Lambda^{1/2}\int \frac{G\Psi G^*}{G^*\Lambda G}\, D(\Lambda^{1/2}; X) \\ =& D(\Lambda^{1/2}; X) \int \frac{G\Psi G^*}{G^*\Lambda G}\Lambda^{1/2} \\ &- \Lambda^{1/2}\int \frac{G\Psi G^*}{G^*\Lambda G}\frac{G^*XG}{G^*\Lambda G}\Lambda^{1/2} \\ &+ \Lambda^{1/2}\int \frac{G\Psi G^*}{G^*\Lambda G}\, D(\Lambda^{1/2}; X). \end{aligned} \tag{23}$$

We now compute the latter expression at $\Lambda = \Lambda_\circ$ and take (7b) into account. This yields

$$\begin{aligned} D(\Theta(\Lambda_\circ); X) = D(\Lambda_\circ^{1/2}; X) I\, \Lambda_\circ^{1/2} + \Lambda_\circ^{1/2}\, I\, D(\Lambda_\circ^{1/2}; X) \\ -\Lambda_\circ^{1/2}\int \frac{G\Psi G^*}{G^*\Lambda_\circ G}\frac{G^*XG}{G^*\Lambda_\circ G}\Lambda_\circ^{1/2} \end{aligned} \tag{24}$$

which, by property (22), may be rewritten as

$$D(\Theta(\Lambda_\circ); X) = \mathcal{M}(X) := X - \Lambda_\circ^{1/2}\int \frac{G\Psi G^*}{G^*\Lambda_\circ G}\frac{G^*XG}{G^*\Lambda_\circ G}\Lambda_\circ^{1/2} \tag{25}$$

where we have defined the linear map $\mathcal{M} : \mathbb{H}_n \longrightarrow \mathbb{H}_n$ that is therefore the derivative of Θ computed at a given fixed point $\Lambda_\circ$.

Adopting a system-theoretic approach, we can consider the sequence of increments $\{X_k\}$ with $X_k = \Lambda_k - \Lambda_\circ$ and the *linear* system $X_{k+1} = \mathcal{M}(X_k)$ as the linear approximation of the nonlinear system $\Lambda_{k+1} = \Theta(\Lambda_k)$ in the neighborhood of its equilibrium point $\Lambda_\circ$. If all the eigenvalues of $\mathcal{M}$ lied in the open unit circle, then we could immediately conclude that $\Lambda_\circ$ is asymptotically stable. However, this is *not* the case. In fact, it is immediate to check that

$$\mathcal{M}(X_\perp) = X_\perp, \quad \forall\, X_\perp \in \mathrm{Range}\,\Gamma^\perp \tag{26}$$

so that $\mathcal{M}$ acts on $\mathrm{Range}\,\Gamma^\perp$ as the identity operator. We thus need a more sophisticated analysis.

C. Properties and Spectrum of $\mathcal{M}$

First, notice that $\mathcal{M}$ maps $\mathbb{H}_n$ in itself but it is not self-adjoint (with respect to the inner product defined in $\mathbb{H}_n$ by $\langle X, Y\rangle = \mathrm{tr}(XY)$). Indeed, given $X, Y \in \mathbb{H}_n$, it may happen that

$$\langle \mathcal{M}(X), Y\rangle \neq \langle X, \mathcal{M}(Y)\rangle$$

so it is not *a priori* true that the eigenvalues of $\mathcal{M}$ are real and that the eigenmatrices of $\mathcal{M}$ span the whole space $\mathbb{H}_n$.

A second observation is stated in the following result.

Lemma 3.2: For any $X \in \mathbb{H}_n$, $\mathrm{tr}\,\mathcal{M}(X) = 0$.

Proof: We have

$$\begin{aligned} \mathrm{tr}\,\mathcal{M}(X) &= \mathrm{tr}\,X - \int \Psi \frac{G^*\Lambda_\circ G}{G^*\Lambda_\circ G}\frac{G^*XG}{G^*\Lambda_\circ G} \\ &= \mathrm{tr}\,X - \int \Psi \frac{G^*XG}{G^*\Lambda_\circ G} \\ &= \mathrm{tr}\,X - \mathrm{tr}\int \frac{G\Psi G^*}{G^*\Lambda_\circ G}X = \mathrm{tr}\,X - \mathrm{tr}\,I\,X = 0. \end{aligned} \tag{27}$$

□

We are now ready to analyze the spectrum of $\mathcal{M}$. Let Y be an eigenmatrix of $\mathcal{M}$ and α be the corresponding eigenvalue, namely Y is a non-zero Hermitian matrix such that $\mathcal{M}(Y) = \alpha\, Y$. Due to (27), it must be $\alpha\ \mathrm{tr}\,Y = 0$. Thus, we have the following corollary.

Corollary 3.1: Let Y be any eigenmatrix of $\mathcal{M}$ and assume $\mathrm{tr}\,Y \neq 0$. Then, the corresponding eigenvalue is zero.

Notice that $\Lambda_\circ$ is one such eigenmatrix. Indeed

$$\begin{aligned} \mathcal{M}(\Lambda_\circ) &= \Lambda_\circ - \Lambda_\circ^{1/2}\int \frac{G\Psi G^*}{G^*\Lambda_\circ G}\frac{G^*\Lambda_\circ G}{G^*\Lambda_\circ G}\,\Lambda_\circ^{1/2} \\ &= \Lambda_\circ - \Lambda_\circ^{1/2}\int \frac{G\Psi G^*}{G^*\Lambda_\circ G}\,\Lambda_\circ^{1/2} \\ &= \Lambda_\circ - \Lambda_\circ^{1/2}\, I\, \Lambda_\circ^{1/2} = 0. \end{aligned}$$

Let Y be an eigenmatrix of $\mathcal{M}$ and α be the corresponding eigenvalue. We want to compute bounds for α. In view of Corollary 3.1, we can assume $\mathrm{tr}\,Y = 0$. We have

$$\alpha Y = Y - \Lambda_\circ^{1/2}\int \frac{G\Psi G^*}{G^*\Lambda_\circ G}\frac{G^*YG}{G^*\Lambda_\circ G}\,\Lambda_\circ^{1/2}$$

or, equivalently

$$(1-\alpha)Y = \Lambda_\circ^{1/2}\int \frac{G\Psi G^*}{G^*\Lambda_\circ G}\frac{G^*YG}{G^*\Lambda_\circ G}\,\Lambda_\circ^{1/2}.$$

Since $\Lambda_\circ > 0$, we can multiply both members by $\Lambda_\circ^{-1/2}$ on the left side, and by $\Lambda_\circ^{-1/2}Y$ on the right side. This yields

$$(1-\alpha)\Lambda_\circ^{-1/2}Y\Lambda_\circ^{-1/2}Y = \int \frac{G\Psi G^*}{G^*\Lambda_\circ G}\frac{G^*YG}{G^*\Lambda_\circ G}Y$$

which, by taking the trace on both members and exploiting the cyclic property of the trace, implies

$$(1-\alpha)\,\mathrm{tr}\left[\left(\Lambda_\circ^{-1/4}Y\Lambda_\circ^{-1/4}\right)^2\right] = \int \Psi\frac{(G^*YG)^2}{(G^*\Lambda_\circ G)^2}.$$

We now observe that the trace in the left-hand side is strictly positive, being the square of the Frobenius norm of the nonzero matrix $\Lambda_\circ^{-1/4}Y\Lambda_\circ^{-1/4}$. In conclusion we get

$$(1-\alpha) = \frac{\int \Psi\frac{(G^*YG)^2}{(G^*\Lambda_\circ G)^2}}{\mathrm{tr}\left[\left(\Lambda_\circ^{-1/4}Y\Lambda_\circ^{-1/4}\right)^2\right]}. \tag{28}$$

The right-hand side of (28) is clearly real and non-negative. Indeed, it vanishes if and only if $G(z)^*YG(z)$ is identically zero on $\mathbb{T}$ or, equivalently, if and only if $Y \in (\mathrm{Range}\,\Gamma)^\perp$. The following theorem is thus proven.

Theorem 3.3: All the eigenvalues of the map $\mathcal{M}$ are real. For any eigenmatrix of $\mathcal{M}$ that is *not* in $(\mathrm{Range}\,\Gamma)^\perp$, the corresponding eigenvalue is strictly smaller than 1. On the space $(\mathrm{Range}\,\Gamma)^\perp$, $\mathcal{M}$ acts as the identity operator.

Remark 3.1: The above theorem may be interpreted as follows. Let $\{S_i\}_{i=1}^{n^2}$ be a basis for the (real) vector space $\mathbb{H}_n$ and define $\mathrm{hvec}(X) := (x_1,\dots,x_{n^2})^\top$ as the column vector with n^2 real entries $x_i := \mathrm{tr}(S_iX)$. Choose a matrix V whose columns form a basis for the linear space $\{x = \mathrm{hvec}(X) : X \in (\mathrm{Range}\,\Gamma)^\perp\}$. Let W be such that

$$T := [V \mid W] \tag{29}$$

is square and nonsingular. For any $X \in \mathbb{H}_n$, let $x := T^{-1}\mathrm{hvec}(X)$. Clearly, x is a coordinate representation of X. Theorem 3.3 states that, with respect to these coordinates, the linear map $\mathcal{M}$ is represented by a matrix M of dimension $n^2 \times n^2$ with block structure $M = \begin{bmatrix} I_{n_\perp} & M_{12} \\ 0 & B \end{bmatrix}$, where $n_\perp$ is the dimension of $(\mathrm{Range}\,\Gamma)^\perp$ and

$$\sigma(B) \subset (-\infty, 1). \tag{30}$$

Clearly, since $I_{n_\perp}$ and B have disjoint spectra, we can select W in (29) in such a way that $M_{12} = 0$, i.e., M has the structure

$$M = \begin{bmatrix} I_{n_\perp} & 0 \\ 0 & B \end{bmatrix}. \tag{31}$$

It remains to establish a lower bound for the spectrum of B.

D. Eigenvalues of $\mathcal{M}$ are Non-Negative

In order to provide a lower bound for the spectrum, we shall consider the linearized map as the generator of a continuous-time evolution semigroup, for which a key spectral property will be derived.

1) $\mathcal{M}$ is the Opposite of a Lindblad Generator: We observe that the operator $\mathcal{M}$ may be written in the form

$$\begin{aligned}\mathcal{M}(X) &= X - \int\left[\Lambda_\circ^{1/2}G\frac{\Psi^{1/2}}{G^*\Lambda_\circ G}G^*\right]X\left[G\frac{\Psi^{1/2}}{G^*\Lambda_\circ G}G^*\Lambda_\circ^{1/2}\right] \\ &= X - \int LXL^*\end{aligned} \tag{32}$$

where $L := \Lambda_\circ^{1/2}G\frac{\Psi^{1/2}}{G^*\Lambda_\circ G}G^*$. It is immediate to check that $L^*L = G\frac{\Psi}{G^*\Lambda_\circ G}G^*$, so that clearly $\int L^*L = I$. Therefore, we can write $-\mathcal{M}(X)$ as a (generalized) *Lindblad* generator [27]:[1]

$$-\mathcal{M}(X) = \int LXL^* - \frac{1}{2}[XL^*L + L^*LX]. \tag{33}$$

2) A Continuous-Time Evolution: We now consider the following continuous-time linear system :

$$\dot X(t) = -\mathcal{M}(X(t)) = \int LX(t)L^* - \frac{1}{2}[X(t)L^*L + L^*LX(t)] \tag{34}$$

with state space being the set of traceless Hermitian matrices (notice that, since $\mathrm{tr}\,\mathcal{M}(X) = 0$, evolution (34) is trace preserving). This will be helpful in proving the following

Theorem 3.4: All the eigenvalues of the map $\mathcal{M}$ are non-negative.

Proof: Let α be an eigenvalue of $\mathcal{M}$ and Y be the corresponding eigenmatrix, so that the state trajectory generated by system (34) with initial condition $X(0) = Y$ is $X(t) = \mathrm{e}^{-\alpha t}Y$. We denote by $\|Y\|_1$ the sum of the absolute values of the eigenvalues of Y, i.e., $\|Y\|_1 := \sum_{\lambda\in\sigma(Y)}|\lambda|$.

Let $Y_P \geq 0$ and $Y_N \geq 0$ be the *positive and negative parts* of Y defined as follows: Let $T^*YT = D = D_P - D_N$, where $T^* = T^{-1}$, D is a diagonal matrix, D_P is obtained from D by annihilating the negative entries and $D_N := -(D - D_P)$. Define $Y_P := TD_PT^*$ and $Y_N := TD_NT^*$. Define also the orthogonal projection $\Pi_P := TO_PT^*$ ($\Pi_N := TO_NT^*$), where O_P (O_N) is the matrix obtained from D_P (D_N) by setting to 1 all the non-zero entries. Clearly

$$\|Y\|_1 = \mathrm{tr}[Y_P + Y_N], \quad Y = Y_P - Y_N. \tag{35}$$

Moreover

$$Y_P = \Pi_PY, \quad Y_N = -\Pi_NY, \quad Y_PY_N = Y_NY_P = 0. \tag{36}$$

Recall now that, in view of Corollary 3.1, we can assume $\mathrm{tr}[Y] = 0$. Thus, taking into account (35) and (36), we have $0 = \mathrm{tr}[Y] = \mathrm{tr}[(\Pi_P + \Pi_N)Y]$ so

[1]It is remarkable to notice that, in the framework of quantum statistical mechanics, it has been shown by Lindblad [27] that any trace-preserving, strongly-continuous semigroup of completely positive maps has a generator which can be written as the sum of a Hamiltonian (Liouvillian) term and a number of terms of the form of the integrand in (33). Such Markov semigroups have long been studied for their relevance to many aspects of quantum theory and thermodynamics (see *e.g.*,[2]). In that setting, their spectral properties have been investigated from an operator-theoretic perspective. In order to avoid to overburden this paper with an unnecessary and rather technical detour, we choose here to prove the needed results by means of linear algebraic tools.

that $0 = \mathrm{tr}[X(t)] = \mathrm{tr}[(\Pi_P + \Pi_N)X(t)]$. Hence, $\mathrm{tr}[\Pi_P \frac{d}{dt}X(t)] = -\mathrm{tr}[\Pi_N \frac{d}{dt}X(t)]$. It is now easy to see that

$$\begin{aligned}\frac{d}{dt}\|X(t)\|_1 &= \frac{d}{dt}\|\mathrm{e}^{-\alpha t}Y\|_1 = \frac{d}{dt}\,\mathrm{tr}[\mathrm{e}^{-\alpha t}(Y_P + Y_N)] \\ &= \mathrm{tr}\left[\frac{d}{dt}(\mathrm{e}^{-\alpha t}(\Pi_P - \Pi_N)Y)\right] \\ &= \mathrm{tr}\left[(\Pi_P - \Pi_N)\frac{d}{dt}X(t)\right] = 2\ \mathrm{tr}\left[\Pi_P \frac{d}{dt}X(t)\right] \\ &= 2\ \mathrm{tr}[-\Pi_P \mathcal{M}(X(t))] \\ &= \int \mathrm{tr}[\Pi_P[2LX(t)L^* - X(t)L^*L - L^*LX(t)]].\end{aligned} \tag{37}$$

Define $X_P(t) := \Pi_P X(t)$, and $X_N(t) := \Pi_N X(t)$. Notice that

$$X_P(t) = X_P(t)\Pi_P = \Pi_P X_P(t) \tag{38}$$

and a similar equality holds for $X_N(t)$. We also have $X_P(t) - X_N(t) = X(t)$ so that, by linearity, the integrand $\mathcal{I}$ of the last member of (37) may be written as $\mathcal{I} = \mathcal{I}_P + \mathcal{I}_N$ with

$$\begin{aligned}\mathcal{I}_P &:= \mathrm{tr}\,[\Pi_P[2LX_P(t)L^* - X_P(t)L^*L - L^*LX_P(t)]] \\ &= 2\ \mathrm{tr}[L^*\Pi_P LX_P(t) - L^*LX_P(t))] \\ &= 2\ \mathrm{tr}[L^*(\Pi_P - I)LX_P(t)] \\ &= 2\ \mathrm{tr}[[X_P(t)]^{1/2}L^*(\Pi_P - I)L[X_P(t)]^{1/2}] \le 0\end{aligned} \tag{39}$$

where we have used the cyclic property of the trace operator, equality (38), the fact that $X_P(t) = \mathrm{e}^{-\alpha t}Y_P \ge 0$ and eventually that $\Pi_P - I \le 0$. As for $\mathcal{I}_N$, we have

$$\begin{aligned}\mathcal{I}_N &:= \mathrm{tr}\,[\Pi_P[-2LX_N(t)L^* + X_N(t)L^*L + L^*LX_N(t)]] \\ &= -2\ \mathrm{tr}[\Pi_P LX_N(t)L^*\Pi_P] \le 0\end{aligned} \tag{40}$$

where we have used the fact that $\Pi_P X_N(t) = 0$, the cyclic property of the trace operator, the fact that $\Pi_P = \Pi_P^2$ and eventually that $X_N(t) \ge 0$. In conclusion we have that $\mathcal{I} \le 0$ so that $(d/dt)\|X(t)\|_1 = \int \mathcal{I} \le 0$. On the other hand

$$0 \ge \frac{d}{dt}\|X(t)\|_1 = \frac{d}{dt}\|\mathrm{e}^{-\alpha t}Y\|_1 = -\alpha \mathrm{e}^{-\alpha t}\|Y\|_1 \tag{41}$$

so that $\alpha \ge 0$ (recall that, since Y is an eigenmatrix, $\|Y\|_1 > 0$). □

Theorem 3.4 and (30) allow us to conclude that the matrix B in (31) is such that $\sigma(B) \subset [0, 1)$ and hence it is a discrete-time stability matrix.

E. Center Manifold Theory

Let us go back to the original non-linear map Θ. The iteration (15) may be incrementally represented as $\Lambda_{k+1} - \Lambda_\circ = \Theta(\Lambda_k) - \Lambda_\circ$ and, by Taylor series expansion, as

$$X_{k+1} = \mathcal{M}(X_k) + m(X_k) \tag{42}$$

where we have defined $X_k := \Lambda_k - \Lambda_\circ$ and m is the residue function that vanishes with its first derivatives at the origin. Moreover, notice that, from (17), (26), and (42), we immediately get

$$m(X_\perp) = 0 \quad \forall\, X_\perp \in (\mathrm{Range}\,\Gamma)^\perp. \tag{43}$$

Theorem 3.5: The set $\mathcal{L}_{\circ+}$ is locally asymptotically stable for Θ.

Proof: We resort again to the coordinate representation of X introduced in Remark 3.1. Moreover, we partition x in the form $x = \begin{bmatrix} x^\perp \\ x^r \end{bmatrix}$, where $x^\perp \in \mathbb{C}^{n_\perp}$ is the component of x corresponding to $(\mathrm{Range}\,\Gamma)^\perp$, and $x^r \in \mathbb{C}^{n_r}$. In these coordinates the incremental evolution (42) is represented by

$$\begin{cases} x^\perp_{k+1} = x^\perp_k + f(x^\perp_k, x^r_k) \\ x^r_{k+1} = Bx^r_k + g(x^\perp_k, x^r_k), \end{cases} \tag{44}$$

where, as already discussed, B is a stability matrix. We are now in the setting of Center Manifold theory, see [13, pp. 34–35]. The first and, in general, most difficult step to apply this theory is to find a center manifold, i.e., a $\mathcal{C}^2$ function $h : \mathbb{C}^{n_\perp} \to \mathbb{C}^{n_r}$ that vanishes with its first order derivatives at the origin, and such that the center manifold equation

$$h(x^\perp + f(x^\perp, h(x^\perp)) = Bh(x^\perp) + g(x^\perp, h(x^\perp)) \tag{45}$$

is satisfied. In our situation, however, this equation admits a solution that may be computed very easily. In fact, in view of (43), it is immediate to check that

$$f(x^\perp, 0) = 0,\ g(x^\perp, 0) = 0, \quad \forall x^\perp. \tag{46}$$

Therefore, we may choose as a solution to (45) the identically zero function h. The asymptotic behavior of trajectories of (44) originating in a neighborhood of the origin is determined by the flow on the center manifold whose dynamics is governed by the equation

$$u_{k+1} = u_k + f(u_k, h(u_k)) = u_k + f(u_k, 0) = u_k. \tag{47}$$

Clearly, the zero solution of (47) is stable[2] and thus, as stated in [13, Theorem 8, page 35]:

1) The zero solution of (44) is stable.
2) If the norm of $x_0 := \begin{bmatrix} x^\perp_0 \\ x^r_0 \end{bmatrix}$ is sufficiently small, then there exists a solution $u_k = u$ of (47) and two positive constants κ and $\beta < 1$, such that

$$|x^\perp_k - u| \le \kappa\beta^k, \quad |x^r_k| = |x^r_k - h(u)| \le \kappa\beta^k. \tag{48}$$

Notice that, since in our case the dynamics of (47) is constant, u does not depend on k, but, in general, *depends on the initial condition* x_0. Notice also that the whole argument holds for an arbitrary choice of $\Lambda_\circ \in \mathcal{L}_{\circ+}$ as the reference fixed point, so that we may always assume that the norm of $x^\perp_0$ is arbitrarily small (or even zero). Therefore, for point 2) above to hold it is only necessary to assume that the norm of x^r_0 is sufficiently small.

[2]Here, following [13, page 35], we use the term *stable* in the sense of *neutrally stable* as opposed to *asymptotically stable*.

In conclusion, if the initial condition of (44) is sufficiently close to a point of the vector representation (according to Remark 3.1) of $\mathcal{L}_{\circ+}$, then $x_k = \begin{bmatrix} x_k^{\perp} \\ x_k^{r} \end{bmatrix} \longrightarrow \begin{bmatrix} \bar{x}^{\perp} \\ 0 \end{bmatrix}$, i.e., x_k converges to a state representing an element of $(\mathrm{Range}\,\Gamma)^{\perp}$. This is equivalent to say that for any $\Lambda_\circ \in \mathcal{L}_{\circ+}$ (defined in (18)) there exists a neighborhood $\mathcal{B}(\Lambda_\circ)$ such that all trajectories $\{\Lambda_k\}$ generated by (15) and originating from $\mathcal{B}(\Lambda_\circ)$, converge to a matrix in $\mathcal{L}_{\circ+}$. Equivalently, $\mathcal{L}_{\circ+}$ is locally asymptotically stable for Θ. □

Remark 3.2: We recall that $\mathcal{L}_{\circ+}$ is a set of fixed points for Θ: This implies that as soon as we enter the set the iteration stops, *i.e.,* all the "orbits" in the center manifold actually consist of single points.

Remark 3.3: We observe that center manifold reduction is typically carried out for an *isolated* fixed point: Our proof, instead, employes this result to establish local asymptotic stability of the *manifold* $\mathcal{L}_{\circ+}$. Each point $\Lambda_\circ \in \mathcal{L}_{\circ+}$ of this manifold provides a solution of the dual problem and hence the *same* solution of the original problem.

IV. Numerical Implementation

In this section we discuss a numerically efficient implementation of the integral in the iteration (15). We show that it may be computed using very robust and reliable linear algebra algorithms. We want to compute $\int G\frac{\Psi}{G^*\Lambda G}G^*$, where G, Ψ and Λ are given and $G^*\Lambda G$ is positive on $\mathbb{T}$. To this aim, we assume that Ψ is rational and, as a preliminary step we compute, using standard tools, a minimal minimum-phase spectral factor $W_\Psi(z) = H(zI - F)^{-1}G + D$ of Ψ. We also employ the factorization $G^*\Lambda G = W^*(z)W(z)$, with

$$W := (B^*PB)^{-1/2}B^*PA(zI - A)^{-1}B + (B^*PB)^{1/2} \quad (49)$$

derived in Appendix. Thus, we clearly have

$$\int G\frac{\Psi}{G^*\Lambda G}G^* = \int (GW^{-1}W_\Psi)(GW^{-1}W_\Psi)^* \quad (50)$$

so that the integral in (50) is the steady-state output covariance of the filter $GW^{-1}W_\Psi$ driven by normalized white noise. Then, let us compute a state space realization of $GW^{-1}W_\Psi$. First, we observe that

$$W^{-1} = [I - (B^*PB)^{-1}B^*PA(zI - Z)^{-1}B](B^*PB)^{-1/2} \quad (51)$$

where

$$Z := A - B(B^*PB)^{-1}B^*PA \quad (52)$$

is a stability matrix. Hence

$$\begin{aligned} GW^{-1} = &-(zI - A)^{-1}B(B^*PB)^{-1}B^*PA(zI - Z)^{-1} \\ &\times B(B^*PB)^{-1/2} + (zI - A)^{-1}B(B^*PB)^{-1/2}. \end{aligned} \quad (53)$$

Notice that $B(B^*PB)^{-1}B^*PA = A - Z = (zI - Z) - (zI - A)$. Plugging this expression into (53) we get

$$\begin{aligned} GW^{-1} = &-(zI - A)^{-1}B(B^*PB)^{-1/2} \\ &+ (zI - A)^{-1}B(B^*PB)^{-1/2} \\ &+ (zI - Z)^{-1}B(B^*PB)^{-1/2} \\ = &(zI - Z)^{-1}B(B^*PB)^{-1/2}. \end{aligned} \quad (54)$$

Eventually, it is now easy to see that $GW^{-1}W_\Psi$ has the following state space realization:

$$GW^{-1}W_\Psi = [0 \mid I]\left(zI - \hat{F}\right)^{-1}\hat{G} \quad (55)$$

with

$$\hat{F} := \begin{bmatrix} F & 0 \\ B(B^*PB)^{-1/2}H & Z \end{bmatrix}, \hat{G} := \begin{bmatrix} G \\ B(B^*PB)^{-1/2}D \end{bmatrix}.$$

Notice that $\hat{F}$ is a stability matrix so that the following result is a straightforward conclusion.

Proposition 4.1: Let Ξ be the solution of the following discrete-time Lyapunov equation:

$$\Xi = \hat{F}\Xi\hat{F}^* + \hat{G}\hat{G}^*. \quad (56)$$

Then the integral in (50) is the bottom-right block of Ξ, i.e.

$$\int G\frac{\Psi}{G^*\Lambda G}G^* = [0 \mid I]\,\Xi\begin{bmatrix} 0 \\ I \end{bmatrix}. \quad (57)$$

In conclusion, for each iteration of the algorithm (15) we only have to compute: the solution of an algebraic Riccati equation of order n, the solution of a discrete-time Lyapunov equation of order $n + n_\Psi$ (n_Ψ being the state space dimension of $W_\Psi(z)$), and the square root of a positive definite matrix Λ. All these operations are accomplished by standard linear algebra algorithms that may be implemented by numerically efficient and robust routines.

V. Evidence From Simulations and a Convergence Conjecture

In the previous section we have proven a local result. In an extensive campaign of simulations, however, we have always observed that the sequence $\{\Lambda_k\}$ converges very fast to a $\Lambda_\circ$ in the closure $\overline{\mathcal{L}}_{\circ+}$ of $\mathcal{L}_{\circ+}$. Thus, we conjecture that indeed $\overline{\mathcal{L}}_{\circ+}$ is globally asymptotically stable for Θ. To this extent, we can do a few considerations. The function Θ maps the open set $\mathcal{P}_+$ of positive definite matrices with unitary trace to itself. Even if all fixed points in this open set are clearly in $\mathcal{L}_{\circ+}$, we cannot exclude that the sequence $\{\Lambda_k\}$ converges to the boundary of $\mathcal{P}_+$, i.e., to a singular matrix. Indeed, it is easy to see that there is a whole family of singular matrices in the boundary of $\mathcal{P}_+$ that are fixed points of Θ. This is the family of the 1-D orthogonal projections. We have conducted some numerical experiments to understand the behavior of the map Θ in the neighborhood of 1-D orthogonal projections. We have observed that, even if we generate the sequence $\{\Lambda_k\}$ by choosing the initial condition

arbitrarily close to a 1-D orthogonal projection Π_1, the sequence always converged to $\mathcal{L}_{o+}$. For this reason we believe that, except for those in $\overline{\mathcal{L}}_{o+}$, the 1-D orthogonal projections are *unstable* equilibrium points. A formal proof of this fact should probably adopt a (nonlinear) Lyapunov approach. In fact, the derivative of the square root in the neighborhood of a singular matrix (as is an orthogonal projection) is infinite and thus a proof based on linearization does not seem viable.

A second remark concerns the values of $\mathbb{J}_\Psi$ along the trajectory $\{\Lambda_k\}$ generated by iterating Θ. It has been shown in [16] that $\Delta\Lambda_k := \Lambda_{k+1} - \Lambda_k$ is a descent direction for $\mathbb{J}_\Psi$. Indeed, experimental evidence in numerical simulation shows that more is true: $\mathbb{J}_\Psi$ always decreases along trajectories $\{\Lambda_k\}$. This fact, if proven, would be an important step toward a Lyapunov argument for *global convergence*.

Summing up, while a closed form solution of the dual problem (2.1) is not available, an iterative algorithm has been proposed in [16], [31] in order to obtain the dual solution numerically. Necessary conditions for such an algorithm to be of interest are clearly its numerical efficiency and, most important, its convergence features.

Our main result proves that the iteration is locally convergent to the manifold of full-rank solutions for the dual problem. The path to this results is quite tortuous yet provides new insights on the dynamics associated to the iteration.

The iterative algorithm presented in Section II-C is thus proven to be an eligible candidate for being the missing piece towards a satisfactory, feasible solution of the spectral approximation problem in the general case. In fact, we believe it should be possible to prove that convergence is almost global, namely that all the stationary points that are not in $\mathcal{L}_{0+}$ are in fact repulsive. Technical difficulties rule out a linearization approach, suggesting a general Lyapunov analysis as the natural pathway to the desired result. This indeed represents the most challenging yet compelling direction for further work.

APPENDIX

In the following we present a factorization result that is repeatedly used in the paper.

Lemma A.1: Let $G(z) = (zI - A)^{-1}B$ with $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times 1}$, and (A, B) a reachable pair. Let $\Lambda \in \mathbb{H}_n$ be such that $G^*\Lambda G > 0$ on $\mathbb{T}$. Then, the following factorization holds:

$$G^*\Lambda G = W^*W \tag{58}$$

where

$$W := (B^*PB)^{-1/2}B^*PA(zI-A)^{-1}B + (B^*PB)^{1/2} \tag{59}$$

and $P \in \mathbb{H}_n$ is the stabilizing solution of the algebraic Riccati equation

$$\Pi = A^*\Pi A - A^*\Pi B(B^*\Pi B)^{-1}B^*\Pi A + \Lambda \tag{60}$$

so that the spectrum of closed loop matrix Z defined in (52) lies inside the open unit disk.

Proof: Given $G(z) = (zI - A)^{-1}B$ and $\Pi \in \mathbb{H}_n$, the following identity holds [14]:

$$[G^*(z) \mid 1]\begin{bmatrix} A^*\Pi A - \Pi & A^*\Pi B \\ B^*\Pi A & B^*\Pi B \end{bmatrix}\begin{bmatrix} G(z) \\ 1 \end{bmatrix} \equiv 0. \tag{61}$$

Therefore, we have

$$G^*\Lambda G = [G^* \mid 1]\begin{bmatrix} \Lambda + A^*\Pi A - \Pi & A^*\Pi B \\ B^*\Pi A & B^*\Pi B \end{bmatrix}\begin{bmatrix} G \\ 1 \end{bmatrix}. \tag{62}$$

Since $G^*\Lambda G$ is positive on the whole $\mathbb{T}$, there exists the stabilizing solution P of the algebraic Riccati (60). Thus if we set $\Pi = P$, the block matrix on the right hand side of (62) admits the factorization

$$\begin{bmatrix} A^*PB \\ B^*PB \end{bmatrix}(B^*PB)^{-1}[B^*PA \mid B^*PB]$$

so that

$$\begin{aligned} G^*\Lambda G &= [G^*A^*PB + B^*PB](B^*PB)^{-1}[B^*PAG + B^*PB] \\ &= W^*(z)W(z) \end{aligned} \tag{63}$$

with $W(z)$ given by (59). □

ACKNOWLEDGMENT

The authors wish to thank the late C. I. Byrnes, who contributed to their work during his last stay in Padova with many insightful discussions, his endless intellectual curiosity, and his friendly attitude, and M. Pavon, for constant encouragement and helpful suggestions about a preliminary version of the manuscript.

REFERENCES

[1] N. I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis* Transl.:Translated from the Russian by N. Kemmer. New York: Hafner Publishing, Co., 1965.

[2] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications*. Berlin, Germany: Springer-Verlag, 1987.

[3] A. Blomqvist, A. Lindquist, and R. Nagamune, "Matrix-valued Nevanlinna-Pick interpolation with complexity constraint: An optimization approach," *IEEE Trans. Autom. Control*, vol. 48, no. 12, pp. 2172–2190, Dec. 2003.

[4] C. I. Byrnes, P. Enqvist, and A. Linquist, "Cepstral coefficients, covariance lags and pole-zero models for finite data strings," *IEEE Trans. Sig. Proc.*, vol. 50, no. 4, pp. 677–693, Apr. 2001.

[5] C. I. Byrnes, P. Enqvist, and A. Linquist, "Identifiability and well-posedness of shaping-filter parameterizations: A global analysis approach," *SIAM J. Control Optim.*, vol. 41, pp. 23–59, 2002.

[6] C. I. Byrnes, T. Georgiou, and A. Lindquist, "A new approach to spectral estimation: A tunable high-resolution spectral estimator," *IEEE Trans. Sig. Proc.*, vol. 49, no. 11, pp. 3189–3205, Nov. 2000.

[7] C. I. Byrnes, T. Georgiou, and A. Lindquist, "A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint," *IEEE Trans. Aut. Control*, vol. 46, no. 6, pp. 822–839, Jun. 2001.

[8] C. I. Byrnes, T. Georgiou, A. Lindquist, and A. Megretski, "Generalized interpolation in H-infinity with a complexity constraint," *Trans. Amer. Math. Soc.*, vol. 358, no. 3, pp. 965–987, 2006.

[9] C. I. Byrnes, S. Gusev, and A. Lindquist, "A convex optimization approach to the rational covariance extension problem," *SIAM J. Control Opim.*, vol. 37, pp. 211–229, 1999.

[10] C. I. Byrnes, S. Gusev, and A. Lindquist, "From finite covariance windows to modeling filters: A convex optimization approach," *SIAM Rev.*, vol. 43, pp. 645–675, 2001.

[11] C. I. Byrnes and A. Lindquist, "The generalized moment problem with complexity constraint," *Integr. Equations Oper. Theory*, vol. 56, no. 2, pp. 163–180, 2006.

[12] C. I. Byrnes and A. Linquist, "Important moments in systems and control," *SIAM J. Control Optim.*, vol. 47, no. 5, pp. 2458–2469, 2008.
[13] J. Carr, *Applications of Centre Manifold Theory*. New York: Springer-Verlag, 1981.
[14] P. Colaneri and A. Ferrante, "Algebraic Riccati equation and J-spectral factorization for H_∞ estimation," *Syst. Control Lett.*, vol. 51, pp. 383–393, 2004.
[15] P. Enqvist, "A homotopy approach to rational covariance extension with degree constraint," *Int. J. Appl. Math. Comp. Sci.*, vol. 11, pp. 1173–1201, 2001.
[16] A. Ferrante, M. Pavon, and F. Ramponi, "Further results on the Byrnes-Georgiou-Lindquist generalized moment problem," in *Modeling, Estimation and Control: Festschrift in Honor of Giorgio Picci on the Occasion of his Sixty-Fifth Birthday*, A. Chiuso, A. Ferrante, and S. Pinzoni, Eds. New York: Springer-Verlag, 2007, pp. 73–83.
[17] A. Ferrante, M. Pavon, and F. Ramponi, "Hellinger vs. Kullback-Leibler multivariable spectrum approximation," *IEEE Trans. Autom. Control*, vol. 53, no. 4, pp. 954–967, May 2008.
[18] A. Ferrante, M. Pavon, and F. Ramponi, "Constrained approximation in the Hellinger distance," in *Proc. Eur. Control Conf. (ECC'07)*, Kos, Greece, Jul. 2007, pp. 322–327.
[19] T. Georgiou, "A topological approach to Nevanlinna-Pick interpolation," *SIAM J. Math. Anal.*, vol. 18, no. 5, pp. 1248–1260, 1987.
[20] T. Georgiou, "The interpolation problem with a degree constraint," *IEEE Trans. Autom. Control*, vol. 44, no. 3, pp. 631–635, Mar. 1999.
[21] T. Georgiou, "Spectral estimation by selective harmonic amplification," *IEEE Trans. Autom. Control*, vol. 46, no. 1, pp. 29–42, Jan. 2001.
[22] T. Georgiou, "Spectral analysis based on the state covariance: The maximum entropy spectrum and linear fractional parameterization," *IEEE Trans. Autom. Control*, vol. 47, no. 11, pp. 1811–1823, Nov. 2002.
[23] T. Georgiou, "Relative entropy and the multivariable multidimensional moment problem," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 1052–1066, Mar. 2006.
[24] T. Georgiou, "Distances between time-series and their autocorrelation statistics," in *Modeling, Estimation and Control: Festschrift in Honor of Giorgio Picci on the Occasion of his Sixty-Fifth Birthday*, A. Chiuso, A. Ferrante, and S. Pinzoni, Eds. New York: Springer-Verlag, 2007, pp. 123–133.
[25] T. Georgiou, "Distances and Riemannian metrics for spectral density functions," *IEEE Trans. Signal Processing*, vol. 55, no. 8, pp. 3995–4003, Aug. 2007.
[26] T. Georgiou and A. Lindquist, "Kullback-Leibler approximation of spectral density functions," *IEEE Trans. Inform. Theory*, vol. 49, no. 11, pp. 2910–2917, Nov. 2003.
[27] G. Lindblad, "On the generators of quantum dynamical semigroups," *Commun. Math. Phys.*, vol. 48, no. 2, pp. 119–130, 1976.
[28] A. Lindquist, "Prediction-error approximation by convex optimization," in *Modeling, Estimation and Control: Festschrift in Honor of Giorgio Picci on the Occasion of his Sixty-Fifth Birthday*, A. Chiuso, A. Ferrante, and S. Pinzoni, Eds. New York: Springer-Verlag, 2007, pp. 265–275.
[29] R. Nagamune, "A robust solver using a continuation method for Nevanlinna-Pick interpolation with degree constraint," *IEEE Trans. Autom. Control*, vol. 48, no. 1, pp. 113–117, Jan. 2003.
[30] A. Nasiri Amini, E. Ebbini, and T. Georgiou, "Noninvasive estimation of tissue temperature via high-resolution spectral analysis techniques," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 2, pp. 221–228, Feb. 2005.
[31] M. Pavon and A. Ferrante, "On the Georgiou-Lindquist approach to constrained Kullback-Leibler approximation of spectral densities," *IEEE Trans. Autom. Control*, vol. 51, no. 4, pp. 639–644, Apr. 2006.
[32] F. Ramponi, A. Ferrante, and M. Pavon, "A globally convergent matricial algorithm for multivariate spectral estimation," *IEEE Trans. Autom. Control*, vol. 54, no. 10, pp. 2376–2388, Oct. 2009.
[33] J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*. New York: American Mathematical Society, 1943.
[34] J. Von Neumann, *Mathematical Foundations of Quantum Mechanics*. Princeton, NJ: Princeton Univ. Press, 1955.

Augusto Ferrante was born in Piove di Sacco, Italy, on August 5, 1967. He received the "Laurea" degree (with honors) in electrical engineering and the Ph.D. degree in control systems engineering from the University of Padova, Padova, Italy, in 1991 and 1995, respectively.

He has been on the faculty of the Colleges of Engineering, University of Udine, Udine, Italy, and the Politecnico di Milano, Milan, Italy. He is presently a Professor in the Department of Information Engineering, University of Padova. His research interests are in the areas of linear systems, spectral estimation, optimal control and optimal filtering, quantum control, and stochastic realization.

Federico Ramponi (M'10) was born in Verona, Italy, on August 3, 1978. He received the Laurea degree in computer engineering and the Ph.D. degree in control engineering from the University of Padova, Padua, Italy, in 2004 and 2009, respectively.

He is currently a post-doctoral Fellow within the Automatic Control Laboratory, ETH Zürich, Zurich, Switzerland. His current research interests include stochastic model predictive control and stochastic model checking.

Francesco Ticozzi received the "Laurea" degree in management engineering and the Ph.D. degree in automatic control and operations research from the University of Padova, Padova, Italy, in 2002 and 2007, respectively.

Since February 2007, he has been with the Department of Information Engineering, University of Padova, first as a Research Associate, and then as an Assistant Professor (*Ricercatore*). From 2005 to 2010, he held visiting appointments at Dartmouth College, Hanover, New Hampshire, partially supported by an "A. Gini" fellowship. His research interests include modeling and control of quantum systems, protection of quantum information, quantum communications and information-theoretic approaches to control systems.