2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

A max-plus based fundamental solution for a class of infinite dimensional Riccati equations

Peter M. Dower and William M. McEneaney

***Abstract*— A new fundamental solution for a specific class of infinite dimensional Riccati equations is developed. This fundamental solution is based on the max-plus dual of the dynamic programming solution operator (or semigroup) of an associated control problem. By taking the max-plus dual of this semigroup operator, the kernel of a dual-space integral operator may be obtained. This kernel is the dual-space Riccati solution propagation operator. Specific initial conditions for the Riccati equation correspond to the associated growth rates of the control problem terminal payoffs. Propagation of the solution of the Riccati equation from these initial conditions proceeds in the dual-space, via a max-plus convolution operation utilizing the aforementioned Riccati solution propagation operator.**

I. Introduction

The objective in this paper is to generalize the approach of [5] to a class of infinite dimensional Riccati equations [2], leading to a new fundamental solution for that class of equations. As in the finite dimensional case [5], this fundamental solution is based on the max-plus dual of the dynamic programming evolution operator (or semigroup) of an associated control problem. The fundamental solution developed is for a specific class of infinite dimensional Riccati equations that was originally motivated by a related problem concerning the amplification of optical signals in optical networks [3]. However, it is expected that the principle demonstrated in this paper can easily be extended to other infinite dimensional Riccati equations.

The theory yielding the aforementioned fundamental solution proceeds by considering an associated finite horizon, infinite dimensional, optimal control problem that generalizes that of [5]. This optimal control problem is constructed such that the associated value function exhibits quadratic growth with respect to the state variable, where this growth is determined by the solution of the Riccati equation in question. Specifically, the kernel of the integral operator that generates this quadratic growth is the solution of the Riccati equation. Hence, by solving the optimal control problem, the solution of the Riccati equation may be obtained.

In order to solve the constructed infinite dimensional optimal control problem, dynamic programming is applied. Specifically, it is noted that the value of the optimal control problem may be propagated to a longer time horizon by the application of a max-plus linear, dynamic programming evolution operator $\mathcal{S}_t$, where non-negative t denotes the increase in the time horizon sought. Taking the max-plus dual leads to an analogous evolution operator $\mathcal{B}_t$ in the dual space. That is, $\mathcal{B}_t$ propagates the dual of the value of the optimal control problem, and hence the solution of the Riccati equation, to a horizon increased by t. As $\mathcal{B}_t$ is a max-plus integral operator, this dual space propagation may be described in terms of the propagation of the kernel B_t of $\mathcal{B}_t$. An operation on this kernel may be obtained that very efficiently propagates B_t itself, leading to significant a reduction in the time required to compute the solution of the Riccati equation, when compared with standard integration techniques. The steps involved are as follows:

1) Encode the Riccati equation initial condition in terms of a quadratic functional;
2) Take the Legendre/Fenchel transform of this quadratic functional;
3) Propagate this in time by any duration t, by operating on it with $\mathcal{B}_t$;
4) Take the inverse Legendre/Fenchel transform to obtain the propagated quadratic functional; and
5) Recover the Riccati solution at t from the functional.

In terms of organization, notation and various preliminaries are provided in Section II, while the specific Riccati equation to be solved is stated in Section III. Analysis of a particular solution is presented in Section IV, along with the development of the dual space operator for propagating any solution. The nature of this solution propagation is considered in Section V, followed by a simple example. Some brief conclusions are provided in Section VI. Proofs are omitted due to space limitations.

II. Preliminaries

This section provides a summary of notation, definitions and concepts from functional analysis (for example, [1]) employed in this paper. Readers wishing to proceed directly to problem formulation may skip to Section III.

A. Function spaces

Given an open subset $\mathcal{I}$ of Euclidean space and a Banach space Z, adopt the notation $C(\mathcal{I};Z)$, $C^1(\mathcal{I};Z)$ and $\mathscr{L}_2(\mathcal{I};Z)$ for the spaces of continuous, continuously differentiable and Lebesgue-square integrable functions with the indicated domain and range spaces, respectively. Also, let $C^{01}(\mathcal{I};Z) \doteq C\left(\overline{\mathcal{I}};Z\right) \cap C^1(\mathcal{I};Z)$. Let $\Lambda \doteq (0,L) \subset \mathbb{R}_{\geq 0}$ denote a spatial interval, where $L \in \mathbb{R}_{>0}$, and $\Lambda_0 \doteq [0,L)$ and $\Lambda_L \doteq (0,L]$. Symbol ∂ is used to denote spatial

This research was supported by grants from AFOSR, NSF, and the Australian Research Council (ARC).

P.M. Dower is with the Department of Electrical & Electronic Engineering, University of Melbourne, Victoria 3010, Australia pdower@unimelb.edu.au

W.M. McEneaney is with the Department of Mechanical & Aerospace Engineering, University of California (San Diego), 9500 GIlman Dr., La Jolla, CA 92093-0411, USA. wmceneaney@eng.ucsd.edu

differentiation of functions of one spatial variable on Λ. Where two spatial variables are present, ∂_1 and ∂_2 are used to represent differentiation with respect to the first and second variable respectively. For example, for $\xi \in C^1(\Lambda;\mathbb{R}^n)$ and $F \in C^1(\Lambda\times\Lambda;\mathbb{R}^{n\times n})$,

$$(\partial\,\xi)(\zeta) \doteq \frac{\partial\xi}{\partial\zeta}(\zeta)\,, \qquad \begin{array}{l} (\partial_1 F)(\eta,\zeta) \doteq \dfrac{\partial F}{\partial\eta}(\eta,\zeta)\,, \\ (\partial_2 F)(\eta,\zeta) \doteq \dfrac{\partial F}{\partial\zeta}(\eta,\zeta)\,. \end{array} \tag{1}$$

Define the spaces

$$\begin{aligned} \mathscr{X} &\doteq \mathscr{L}_2(\Lambda;\mathbb{R}^n)\,, \\ \mathscr{X}_\zeta &\doteq \left\{ x\in\mathscr{X} \,\middle|\, \begin{array}{c} x \text{ absolutely continuous on } \Lambda_\zeta, \\ x(\zeta)=0,\ \partial x\in\mathscr{X} \end{array} \right\}, \\ \mathscr{W} &\doteq \mathscr{L}_2(\Lambda;\mathbb{R}^m)\,, \\ \star[r,t] &\doteq \mathscr{L}_2([r,t];\star)\,, \quad \star\in\{\mathscr{W},\mathscr{X}_0,\mathscr{X}_L\}\,, \\ \mathscr{W}_0[r,t] &\doteq \left\{ w\in\mathscr{W}[r,t] \,\middle|\, \begin{array}{c} [w(\cdot)](0)\in\mathscr{L}_2([r,t];\mathbb{R}^m)\,, \\ \partial w\in\mathscr{W}[r,t] \end{array} \right\}, \end{aligned} \tag{2}$$

in which $\zeta\in\{0,L\}$ and $[r,t]\subset\mathbb{R}_{\geq 0}$ is a time interval. Given $x,\xi\in\mathscr{X}$, $\langle x,\,\xi\rangle_{\mathscr{X}}$ denotes the standard inner product on $\mathscr{L}_2(\Lambda;\mathbb{R}^n)$, while $\|x\|_{\mathscr{X}}$ denotes the corresponding norm. $B_{\mathscr{X}}(x;R)$ is used to denote a ball of radius $R\in\mathbb{R}_{\geq 0}$, centre x in $\mathscr{X}$. Given two Banach spaces $\mathscr{X}$ and $\mathscr{Z}$, $\mathscr{L}(\mathscr{X};\mathscr{Z})$ denotes the space of bounded linear operators mapping $\mathscr{X}$ to $\mathscr{Z}$. The domain of an operator Π is denoted by $\mathsf{dom}(\Pi)$. Where $\mathscr{X}$ and $\mathscr{Z}$ are Hilbert spaces, the adjoint $\Pi'\in\mathscr{L}(\mathscr{Z};\mathscr{X})$ of $\Pi\in\mathscr{L}(\mathscr{X};\mathscr{Z})$ exists if

$$\langle z,\,\Pi\,x\rangle_{\mathscr{Z}} = \langle \Pi'\,z,\,x\rangle_{\mathscr{X}} \tag{3}$$

for all $x\in\mathsf{dom}(\Pi)\subset\mathscr{X}$, $z\in\mathsf{dom}(\Pi')\subset\mathscr{Z}$. By definition, the spatial differentiation operation ∂ of (1) is bounded on $\mathscr{X}_0$ and $\mathscr{X}_L$, but not on $\mathscr{X}$. For convenience, define the spatial differentiation operator ∇ to be the restriction of ∂ to $\mathscr{X}_0$. Then, its adjoint ∇' exists on $\mathscr{X}_L$. Specifically,

$$\nabla : \mathsf{dom}(\nabla)\subset\mathscr{X}\mapsto\mathscr{X}\,, \quad \left\{ \begin{array}{l} \mathsf{dom}(\nabla) = \mathscr{X}_0\,, \\ \;[\nabla\xi](\lambda) = [\partial\xi](\lambda)\,. \end{array} \right. \tag{4}$$

$$\nabla' : \mathsf{dom}(\nabla')\subset\mathscr{X}\mapsto\mathscr{X}\,, \left\{ \begin{array}{l} \mathsf{dom}(\nabla') = \mathscr{X}_L\,, \\ \;[\nabla'\xi](\lambda) = -[\partial\xi](\lambda)\,. \end{array} \right. \tag{5}$$

B. Functionals

A functional $f:\mathscr{X}\mapsto\mathbb{R}$ is nonnegative if $f(x)\geq 0$ for all $x\in\mathscr{X}$. f is positive if it is nonnegative and $f(x)>0$ for all $x\neq 0$. Given two such functionals $f,\,g:\mathscr{X}\mapsto\mathbb{R}$, $f>g$ means that $f-g$ is positive (respectively, $\geq$ and nonnegative). A symbol of the form $\mathcal{F}$ is used to denote an integral operator on $\mathscr{X}$, defined with respect to a kernel F. Specifically,

$$\mathcal{F}\,x = (\mathcal{F}\,x)(\cdot) \doteq \int_\Lambda F(\cdot,\zeta)\,x(\zeta)\,d\zeta\,. \tag{6}$$

Lemma 1: Let $f:\mathscr{X}\mapsto\mathbb{R}$ denote a quadratic functional of the form $f(x)\doteq\frac{1}{2}\langle x,\,\mathcal{F}\,x\rangle_{\mathscr{X}}$, in which $\mathcal{F}$ denotes a bounded integral operator of the form (6).

(i) f is closed;

(ii) f is convex if and only if f is nonnegative.

The Fréchet derivative of a functional $V\in C^{01}(\mathscr{X};\mathbb{R})$ at $x\in\mathscr{X}$ in direction $h\in\mathscr{X}$ is denoted by $\nabla_x V(x)[h]\in\mathbb{R}$. The Gateaux derivative of $V\in C^{01}(\mathscr{X};\mathbb{R})$ at $x\in\mathscr{X}$ is denoted by $\nabla_x V(x)\in\mathscr{X}$. As $\mathscr{X}$ is a Hilbert space, the Riesz-Fréchet representation theorem requires that for all $x,\,h\in\mathscr{X}$,

$$\nabla_x V(x)[h] = \lim_{\epsilon\to 0^+}\frac{V(x+\epsilon\,h)-V(x)}{\epsilon} = \langle\nabla_x V(x),\,h\rangle_{\mathscr{X}}\,. \tag{7}$$

C. Integral operators

An operator $\mathcal{F}$ of the form (6) satisfies $\mathcal{F}\,x=0$ for all $x\in\mathscr{X}$ if and only if the functional $f:\mathscr{X}\mapsto\mathbb{R}$ defined by $f(x)\doteq\frac{1}{2}\langle x,\,\mathcal{F}\,x\rangle_{\mathscr{X}}$ is identically zero. Equivalently, the kernel F of $\mathcal{F}$ is identically the zero matrix in $\mathbb{R}^{n\times n}$. $\mathcal{F}:\mathscr{X}\mapsto\mathscr{X}$ is nonnegative (respectively, positive) if the aforementioned functional $f:\mathscr{X}\mapsto\mathbb{R}$ is nonnegative (positive). Integral operators of the form (6) may be composed with a constant matrix multiplication operator to yield another integral operator of the form (6). Specifically, given a constant matrix $A\in\mathbb{R}^{n\times n}$,

$$\left\{ \begin{array}{l} (\mathcal{F}\,A)\,x \doteq \mathcal{F}\,(A\,x) = \displaystyle\int_\Lambda [F(\cdot,\zeta)\,A]\,x(\zeta)\,d\zeta\,, \\ (A'\,\mathcal{F})\,x \doteq A'\,(\mathcal{F}\,x) = \displaystyle\int_\Lambda [A'\,F(\cdot,\zeta)]\,x(\zeta)\,d\zeta\,, \end{array} \right. \tag{8}$$

so that the kernels are respectively $F\,A$ and $A'\,F$. The composition of integral operators $\mathcal{F}\,\mathcal{G}$ is also an integral operator. For convenience, the kernel of this composition is denoted by $F\circledast G$. That is, $F\circledast G:\Lambda\times\Lambda\mapsto\mathbb{R}^{n\times n}$, where

$$(F\circledast G)(\eta,\zeta) \doteq \int_\Lambda F(\eta,\rho)\,G(\rho,\zeta)\,d\rho\,. \tag{9}$$

The adjoint of $\mathcal{F}$ is denoted by $\mathcal{F}'$ and defined by

$$\mathcal{F}'x \doteq \int_\Lambda F(\zeta,\cdot)'\,x(\zeta)\,d\zeta\,. \tag{10}$$

$\mathcal{F}$ is a self-adjoint operator if $\mathsf{dom}(\mathcal{F})=\mathsf{dom}(\mathcal{F}')$ and $\mathcal{F}\,x=\mathcal{F}'\,x$ for all $x\in\mathsf{dom}(\mathcal{F})$. Equivalently, in terms of the kernel F,

$$F(\eta,\zeta) = F(\zeta,\eta)'\,, \quad \forall\;\eta,\,\zeta\in\Lambda\,. \tag{11}$$

Given a symmetric matrix $\Sigma\in\mathbb{R}^{n\times n}$ and a self-adjoint operator $\mathcal{F}$, note that the kernel of the composed operator $\mathcal{F}\,\Sigma\,\mathcal{F}$ is $F\circledast(\Sigma\,F)\equiv(F\,\Sigma)\circledast F$. Operator $\mathcal{F}\,\Sigma\,\mathcal{F}$ is self-adjoint by (11).

The spatial differentiation operators ∇ and ∇' of (4) and (5) may be formally applied to an integral operator $\mathcal{F}$ or its adjoint $\mathcal{F}'$, respectively (6) and (10). Note specifically that

$$(\nabla'\mathcal{F})\,x \doteq \nabla'\,(\mathcal{F}\,x) = -\int_\Lambda \partial_1 F(\cdot,\zeta)\,x(\zeta)\,d\zeta\,,$$

where $\mathcal{F}\,x\in\mathsf{dom}(\nabla')$. The domain condition on the right-hand side implies a specific boundary condition for the kernel

F, in particular, that $F(L,\cdot) = 0^{n\times n}$ if $\mathcal{F}\, x \in \mathsf{dom}\,(\nabla')$. Note that by (3), (10) and integration by parts,

$$
\begin{aligned}
(\nabla'\mathcal{F})'\, x &= \mathcal{F}'\, \nabla x = \int_\Lambda F(\zeta,\cdot)'\, \partial x(\zeta)\, d\zeta \\
&= F(L,\cdot)'\, x(L) - F(0,\cdot)'\, x(0) - \int_\Lambda \partial_1 F(\zeta,\cdot)'\, x(\zeta)\, d\zeta\,.
\end{aligned}
$$

Where $x \in \mathscr{X}_0$ and $\mathcal{F}\, x \in \mathsf{dom}\,(\nabla')$, note that the first two terms on the right-hand side are zero. Furthermore, if $\mathcal{F}$ is self-adjoint, then $\nabla'\mathcal{F}$ and $(\nabla'\mathcal{F})'$ are of similar form, with

$$
\left\{
\begin{aligned}
\nabla'\mathcal{F}\, x &= \int_\Lambda \left[-\partial_1 F(\cdot,\zeta)\right]\, x(\zeta)\, d\zeta\,, \\
(\nabla'\mathcal{F})'\, x &= \int_\Lambda \left[-\partial_2 F(\cdot,\zeta)\right]\, x(\zeta)\, d\zeta\,.
\end{aligned}
\right.
$$

Where an integral operator of the form (6) is time indexed, the time derivative is

$$
\dot{\mathcal{F}}_t\, x \doteq \lim_{\epsilon\to 0} \left[\frac{\mathcal{F}_{t+\epsilon}\, x - \mathcal{F}_t\, x}{\epsilon}\right] = \int_\Lambda \dot{F}_t(\cdot,\zeta)\, x(\zeta)\, d\zeta\,, \quad (12)
$$

which is also of the form (6). Finally, boundedness of $\mathcal{F}$ may be inferred via an induced norm on the kernel F. The details of this are omitted for brevity.

III. Problem statement

The infinite dimensional Riccati equation of interest is an integro-differential equation with initial and boundary data. A solution $P_t : \Lambda\times\Lambda \mapsto \mathbb{R}^{n\times n}$ of this equation is a time-dependent matrix-valued function of two spatial variables, each defined on Λ. Specifically, a solution P_t of the equations

$$
\begin{aligned}
&\dot{P}_t = P_t\, A + A'\, P_t + \partial_1 P_t + \partial_2 P_t + (P_t\, \sigma) \circledast (\sigma'\, P_t) + C\,, \\
&P_0 = M\,, \quad 0 = \mathcal{B}_1 P_t = \mathcal{B}_2 P_t\,,
\end{aligned}
\quad (13)
$$

is sought for all $t \in [0,T]$, for some $T \in \mathbb{R}^+ \doteq R_{>0} \cup \{+\infty\}$ fixed. There, $\mathcal{B}_{1,2}$ denote the boundary value operators

$$
\begin{aligned}
\mathcal{B}_1 P &= (\mathcal{B}_1 P)\,(\cdot) \doteq P(L,\cdot)\,, \\
\mathcal{B}_2 P &= (\mathcal{B}_2 P)\,(\cdot) \doteq P(\cdot,L)\,,
\end{aligned}
\quad (14)
$$

while the notation ∂_1, ∂_2, and $\circledast$ is as per (1) and (9). Equation (13) represent the kernel form of the Riccati equation of interest. The corresponding operator form of this Riccati equation is given by

$$
\dot{\mathcal{P}}_t = \mathcal{P}_t\, A + A'\, \mathcal{P}_t - \nabla'\mathcal{P}_t - (\nabla'\mathcal{P}_t)' + \mathcal{P}_t\, \sigma\, \sigma'\, \mathcal{P}_t + \mathcal{C}\,,
$$

subject to the initialization $\mathcal{P}_0 = \mathcal{M}$. Here, operators $\mathcal{P}_t$, $\mathcal{M}$, and $\mathcal{C}$ are of the form (6), with kernels P_t, M, and C respectively. For brevity, attention is restricted to the kernel form (13). There, the initial condition $M : \Lambda\times\Lambda \mapsto \mathbb{R}^{n\times n}$ is restricted according to

$$
M \in \mathscr{M}\,, \quad (15)
$$

where

$$
\mathscr{M} \doteq \left\{ M \;\middle|\; \begin{array}{c} M : \Lambda\times\Lambda \mapsto \mathbb{R}^{n\times n} \text{ is the kernel of any} \\ \text{self-adjoint and invertible operator } \mathcal{M} \\ \text{satisfying (6) and } \mathcal{M}\, x \in \mathscr{X}_L \;\forall\, x \in \mathscr{X}_0 \end{array} \right\}.
$$

Existence and uniqueness of solutions for (13) is invoked via the following condition on $M \in \mathscr{M}$:

Given $M \in \mathscr{M}$, there exists a $T \in \mathbb{R}^+$ such that (13) has a unique solution P_t for all $t \in [0,T)$. Furthermore, P_t defines an integral operator $\mathcal{P}_t$ of the form (6) (with P_t as its kernel) with the following properties:
(i) $\mathcal{P}_t > \mathcal{M}$;
(ii) $\left\|\mathcal{M}^{-1}\, \mathcal{P}_t\right\| < 1$. (16)

In order to conveniently discuss a particular solution of (13) that satisfies (16), as parameterized by a specific initial condition M, define

$$
\begin{aligned}
T_M &\doteq \sup \left\{ T > 0 \;\middle|\; \begin{array}{c} \text{(16) holds for} \\ \text{the given } M \end{array} \right\}, \\
\mathsf{Ric}(M) &\doteq P \;\left|\; \begin{array}{c} P_t \text{ satisfies (13) } \forall\, t \in [0,T_M)\,, \\ T_M > 0\,,\ M \text{ given}\,. \end{array} \right.
\end{aligned}
$$

Throughout, it will be assumed that a particular solution with the appropriate properties discussed above exists. This is summarized as follows.

Assumption 2: $\exists\, M \in \mathscr{M}$ such that $T_M > 0$.

The objective is to develop an efficient method for computing a solution $\mathsf{Ric}(\widetilde{M})$ given a particular solution $\mathsf{Ric}(M)$, for which the condition

$$
\widetilde{\mathcal{M}} > \mathcal{M} \quad (17)
$$

holds. This method will be based on a generalization of the max-plus approach reported in [5].

IV. The particular solution

An optimal control problem is posed whose value function incorporates the particular solution $\mathsf{Ric}(M)$ for M as per Assumption 2. By formulating the dynamic programming principle for this optimal control problem, a semigroup evolution operator may be developed for the value function, and hence the particular solution $\mathsf{Ric}(M)$. A dual space representation for this evolution operator is obtained that is fundamental to computing both $\mathsf{Ric}(M)$ and $\mathsf{Ric}(\widetilde{M})$.

A. Optimal control problem

The optimal control problem of interest is defined with respect to the linear infinite dimensional plant dynamics and initial state given by

$$
\dot{\xi}(t) = A\, \xi(t) - \nabla\xi(t) + \sigma\, w(t)\,, \quad \xi(0) = \xi_\circ\,, \quad (18)
$$

for all $t \in [0,T_M)$, where M is as per Assumption 2. Here, $\xi(t)$ and $w(t)$ respectively denote the state and input at time t (both are functions on Λ for every t), $\xi_\circ$ denotes the initial state, and ∇ denotes the spatial differentiation operator (4). Symbols A and σ denote fixed real matrices, with $A \in \mathbb{R}^{n\times n}$ and $\sigma \in \mathbb{R}^{n\times m}$, that multiply respectively $[\xi(t)]\,(\lambda)$ and $[\nabla\xi(t)]\,(\lambda)$ on the left for every $\lambda \in \Lambda$. Solutions to (18) are further restricted by the requirement that

$$
\xi_\circ \in \mathscr{X}_0\,, \quad w \in \mathscr{W}_0[0,t] \quad \forall\, t \in [0,T_M)\,. \quad (19)
$$

For convenience, define the operator valued function $\mathcal{T}$ by

$$(\mathcal{T}_t\,\xi)\,(\eta) = \begin{cases} \exp\,(A\,t)\;\xi(\eta - t) & \text{if } t\in[0,\eta]\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

$\mathcal{T}$ is of fundamental importance in defining solutions of (18). In particular, $\mathcal{T}$ is a C_0-semigroup generated by the operator $A - \nabla$ in (18). It allows solutions of the abstract Cauchy problem (18) to be represented in terms of operations on the initial state and input. The key result of importance here is that such a solution exists under appropriate conditions, and resides in a particular space.

Theorem 3: Given any initial state $\xi_\circ \in \mathscr{X}_0$, and any input $w \in \mathscr{W}_0[0,t]$, the mild solution ξ of the initial value problem (18) satisfies the following for all $t \in [0, T_M)$:

$$\xi(t) = \mathcal{T}(t-r)\,\xi_\circ + \int_r^t \mathcal{T}(t-s)\,\sigma\,w(s)\,ds\,, \tag{20}$$

$$\xi(t) \in \mathscr{X}_0\,. \tag{21}$$

An optimal control problem may be defined with respect to the dynamics of (18) in order to capture the solution of (13). Specifically, for each $z \in \mathscr{X}_0$, $t \in [0,T_M)$, define the payoff $J^z : [0,t]\times\mathscr{X}\times\mathscr{W}_0[0,t] \mapsto \mathbb{R}_{\ge 0}$ and value $\overline{W}^z : [0,t]\times\mathscr{X}\mapsto\mathbb{R}$ by

$$J^z(t,x;w) \doteq \int_0^t \tfrac{1}{2}\langle \xi(s),\,\mathcal{C}\,\xi(s)\rangle_{\mathscr{X}} - \tfrac{1}{2}\|w(s)\|^2_{\mathscr{W}}\,ds + \psi(\xi(t),z)\,, \tag{22}$$

$$\overline{W}^z(t,x) \doteq \sup_{w\in\mathscr{W}_0[0,t]} J^z(t,x;w)\,, \tag{23}$$

where $\xi(0) = x \in \mathscr{X}$, and $\psi : \mathscr{X}\times\mathscr{X}\mapsto\mathbb{R}_{\ge 0}$ denotes the terminal cost given by

$$\psi(x,z) \doteq \tfrac{1}{2}\langle x-z, \mathcal{M}\,(x-z)\rangle_{\mathscr{X}}\,. \tag{24}$$

Here, $\mathcal{C}$ and $\mathcal{M}$ denote self-adjoint integral operators of the form (6), with the kernel of $\mathcal{M}$ satisfying (15) by Assumption 2. By inspection, note that for all $x,\,z\in\mathscr{X}$

$$\overline{W}^z(0,x) = \psi(x,z)\,. \tag{25}$$

The value (23) naturally satisfies a dynamic programming principle, as stated in the following theorem.

Theorem 4 (DPP): Given any $z \in \mathscr{X}_0$, the value $\overline{W}^z$ of (23) satisfies the dynamic programming principle (DPP)

$$\overline{W}^z = \mathcal{D}_\tau\,\overline{W}^z \tag{26}$$

for all $\tau\in[0,t]$, $t\in[0,T_M)$, where $\mathcal{D}_\tau$ is the evolution operator defined by

$$[\mathcal{D}_\tau\,\phi]\,(t,x) \doteq \sup_{w\in\mathscr{W}_0[0,\tau]} \left\{ \begin{array}{c|c} \displaystyle\int_0^\tau \tfrac{1}{2}\langle\xi(s),\,\mathcal{C}\,\xi(s)\rangle_{\mathscr{X}} & \\ \quad -\tfrac{1}{2}\|w(s)\|^2_{\mathscr{W}}\,ds & \begin{array}{c}\text{(18) holds with}\\ \xi(0)=x\end{array} \\ \quad\quad + \phi(t-\tau,\xi(\tau)) & \end{array} \right\}. \tag{27}$$

A verification theorem for the optimal control problem (23) is key to demonstrating that an explicit solution exists. This theorem states that if a solution to a particular partial differential equation exists, then that solution corresponds to the value (23) of the control problem.

Theorem 5: Suppose $W^z \in C^{01}((0,T_M)\times\mathscr{X};\mathbb{R})$ satisfies

$$0 = -\frac{\partial W^z}{\partial t}(t,x) + \tfrac{1}{2}\langle x,\,\mathcal{C}\,x\rangle_{\mathscr{X}} + \langle \nabla_x W^z(t,x),\,A\,x\rangle_{\mathscr{X}} - \langle \nabla'\nabla_x W^z(t,x),\,x\rangle_{\mathscr{X}} + \tfrac{1}{2}\left\|\sigma'\,\nabla_x W^z(t,x)\right\|^2_{\mathscr{W}} \tag{28}$$

$$\nabla_x W^z(t,x) \in \mathsf{dom}\,(\nabla') = \mathscr{X}_L \tag{29}$$

$$W^z(0,x) = \psi(x,z)\,, \tag{30}$$

for all $t\in[0,T_M)$, $x\in\mathscr{X}_0$, where $z\in\mathscr{X}_0$ fixed. Then, $W^z(t,x) \ge J^z(t,x;w)$ for all $x\in\mathscr{X}_0$, $w\in\mathscr{W}_0[0,t]$. Furthermore, if there exists a mild solution ξ^* as per (20) corresponding to a distributed input w^* with feedback characterization

$$w^*(s) \doteq \tfrac{1}{2}\sigma'\,\nabla W^z(t-s,\xi^*(s))\,,\quad s\in[0,t]\,, \tag{31}$$

such that $\xi^*(s)\in\mathscr{X}_0$ for all $s\in[0,t]$, then $W^z(t,x) = J^z(t,x;w^*)$, and consequently $W^z(t,x) = \overline{W}^z(t,x)$.

With a view to applying the verification Theorem 5, for each $z\in\mathscr{X}_0$ and $t\in[0,T_M)$, explicitly define the functional $\widehat{W}^z : [0,t]\times\mathscr{X}\mapsto\mathbb{R}$ by

$$\widehat{W}^z(t,x) \doteq \tfrac{1}{2}\langle x,\,\mathcal{P}_t\,x\rangle_{\mathscr{X}} + \langle x,\,\mathcal{Q}_t\,z\rangle_{\mathscr{X}} + \tfrac{1}{2}\langle z,\,\mathcal{R}_t\,z\rangle_{\mathscr{X}}\,, \tag{32}$$

where $\mathcal{P}_t$, $\mathcal{Q}_t$, $\mathcal{R}_t$ denote time-indexed integral operators of the form (6) whose kernels P_t, Q_t, R_t satisfy respectively (13) and

$$\begin{cases} \dot{Q}_t = A'\,Q_t + \partial_1 Q_t + (P_t\,\sigma)\circledast(\sigma'\,Q_t)\,, \\ Q_0 = -M\,, \\ \quad 0 = \mathcal{B}_1 Q_t = \mathcal{B}_2 Q_t\,, \end{cases} \tag{33}$$

$$\begin{cases} \dot{R}_t = (Q_t'\,\sigma)\circledast(\sigma'\,Q_t)\,, \\ R_0 = M\,, \\ \quad 0 = \mathcal{B}_1 Q_t = \mathcal{B}_2 Q_t\,, \end{cases} \tag{34}$$

in which $P_t,\,Q_t,\,R_t : \Lambda\times\Lambda\mapsto\mathbb{R}^{n\times n}$ are time-dependent matrix-valued functions of two independent spatial variables (both defined on the interval Λ), and M is the kernel of the self-adjoint operator $\mathcal{M}$ of (24) satisfying Assumption 2.

Theorem 6: The explicit functional $\widehat{W}^z$ of (32) satisfies for all $t\in[0,T_M)$, $x,\,z\in\mathscr{X}_0$.

$$\widehat{W}^z(t,x) = \overline{W}^z(t,x) = [\mathcal{S}_t\,\psi(\cdot,z)]\,(x)\,, \tag{35}$$

in which ψ is the terminal cost (24), and $\mathcal{S}_t$ is the evolution operator defined by

$$[\mathcal{S}_t\,\phi]\,(x) \doteq \sup_{w\in\mathscr{W}_0[0,t]} \left\{ \begin{array}{c|c} \displaystyle\int_0^t \tfrac{1}{2}\langle\xi(s),\,\mathcal{C}\,\xi(s)\rangle_{\mathscr{X}} & \text{(18) holds} \\ \quad -\tfrac{1}{2}\|w(s)\|^2_{\mathscr{W}}\,ds & \text{with} \\ \quad + \phi(\xi(t)) & \xi(0)=x \end{array} \right\}, \tag{36}$$

which satisfies the semigroup property

$$\mathcal{S}_{\tau+t}\,\phi = \mathcal{S}_\tau\,\mathcal{S}_t\,\phi \tag{37}$$

for all $t\in[0,T_M)$, $\tau\in[0,T_M-t)$.

B. Dual-space representation for $\mathsf{Ric}(M)$

The semigroup property of (37) describes how the particular solution $\mathsf{Ric}(M)$ can be propagated forward in time. By appealing to semiconvex duality, this evolution can be represented in a dual space. Analysis of this dual space representation ultimately leads to the definition of a dual space evolution operator that can be applied to propagate *any* solution. That is, it can be used to find another solution $\mathsf{Ric}(\widetilde{M})$ corresponding to initial conditions $\widetilde{M}$ different from M. Here, a dual-space representation of the propagated particular solution is developed. The details of this development rely on concepts and results from convex analysis.

Semiconvex duality [6] will be introduced using operators defined with respect to the max-plus algebra (e.g. [4]). This algebra is a commutative semifield over $\mathbb{R}^- \doteq \mathbb{R} \cup \{-\infty\}$ equipped with the addition and multiplication operations $\oplus$ and $\ominus$ are defined by $a \oplus b \doteq \max(a, b)$ and $a \otimes b \doteq a + b$. The max-plus algebra is an idempotent semifield as $\oplus$ is idempotent operation (i.e. $a \oplus a = a$) with no inverse.

Recall that a functional $f : \mathscr{X} \mapsto \mathbb{R}$ is semiconvex if there exists a self-adjoint integral operator $\mathcal{K}$ of the form (6) such that $f(x) + \frac{1}{2}\langle x, \mathcal{K}\, x\rangle_{\mathscr{X}}$ is convex. A space of such functionals is denoted respectively by $\mathscr{S}^{\mathcal{K}}(\mathscr{X})$, where

$$\mathscr{S}^{\mathcal{K}}(\mathscr{X}) \doteq \left\{ f : \mathscr{X} \mapsto \mathbb{R}^n \;\middle|\; \begin{array}{c} f(\cdot) + \frac{1}{2}\langle \cdot, \mathcal{K}\, \cdot\rangle_{\mathscr{X}} \\ \text{is convex on } \mathscr{X} \end{array} \right\}.$$

It may be shown that $\mathscr{S}^{\mathcal{K}}(\mathscr{X})$ is a max-plus vector space. It is convenient to define a max-plus integral operator $\mathcal{D}_\psi$ in which $\psi : \mathscr{X} \times \mathscr{X} \mapsto \mathbb{R}$ of (24) plays the role of kernel. In particular, define the *semi-convex dual operator* [6]

$$\mathcal{D}_\psi\, \phi = [\mathcal{D}_\psi\, \phi](\cdot) = -\int_{\mathscr{X}}^{\oplus} \psi(x, \cdot) \otimes [-\phi(x)]\, dx\,, \quad (38)$$

where $\int_{\mathscr{X}}^{\oplus} f(z)\, dz \doteq \sup_{z \in \mathscr{X}} f(z)$.

Theorem 7: Let $\phi \in \mathscr{S}^{\mathcal{K}}(\mathscr{X})$ be closed, where $\mathcal{K}$ is a self-adjoint integral operator of the form (6) satisfying $\mathcal{K} < -\mathcal{M}$, with $\mathcal{M}$ defined by kernel M. Then, for all $x, z \in \mathscr{X}$,

$$\phi(x) = [\mathcal{D}_\psi^{-1}\, a](x)\,, \quad a(z) = [\mathcal{D}_\psi\, \phi](z)\,, \quad (39)$$

where

$$\mathcal{D}_\psi^{-1}\, a = [\mathcal{D}_\psi^{-1}\, a](\cdot) \doteq \int_{\mathscr{X}}^{\oplus} \psi(\cdot, z) \otimes a(z)\, dz\,, \quad (40)$$

Theorem 7 may be applied to show that the semiconvex dual of $\mathcal{S}_t\, \psi(\cdot, z)$ is well-defined for each $z \in \mathscr{X}_0$. To this end, define the operator $\mathcal{K}_t \doteq -\alpha\, \mathcal{P}_t - (1-\alpha)\, \mathcal{M}$, with $\alpha \in (0, 1)$ fixed, where $\mathcal{P}_t$ and $\mathcal{M}$ are as defined by $\mathsf{Ric}(M)$. Then, by assertion (i) of (16), $\mathcal{P}_t > \mathcal{M}$, so that

$$\begin{aligned} \mathcal{P}_t + \mathcal{K}_t &= (1-\alpha)\, (\mathcal{P}_t - \mathcal{M}) > 0\,, \\ -\mathcal{K}_t - \mathcal{M} &= \alpha\, (\mathcal{P}_t - \mathcal{M}) > 0\,. \end{aligned}$$

Note also that $\mathcal{K}_t$ is self-adjoint by definition of $\mathcal{P}_t$ and $\mathcal{M}$. That is,

$$-\,\mathcal{P}_t < \mathcal{K}_t < -\mathcal{M}\,, \quad \mathcal{K}_t \text{ self-adjoint}\,. \quad (41)$$

By Theorem 6, i.e. (32) and (35),

$$\begin{aligned} &[\mathcal{S}_t\, \psi(\cdot, z)]\,(x) + \tfrac{1}{2}\, \langle x, \mathcal{K}_t\, x\rangle_{\mathscr{X}} = \widetilde{W}^z(t, x) + \tfrac{1}{2}\, \langle x, \mathcal{K}_t\, x\rangle_{\mathscr{X}} \\ &= \tfrac{1}{2}\, \langle x, (\mathcal{P}_t + \mathcal{K}_t)\, x\rangle_{\mathscr{X}} + \langle x, \mathcal{Q}_t\, z\rangle_{\mathscr{X}} + \tfrac{1}{2}\, \langle z, \mathcal{R}_t\, z\rangle_{\mathscr{X}}\,. \end{aligned}$$

Note that $\mathcal{P}_t + \mathcal{K}_t$ is positive by (41). Hence, the right-hand side of the above equation is convex with respect to $x \in \mathscr{X}_0$ by Lemma 1. That is, for any $t \in [0, T_M)$,

$$[\mathcal{S}_t\, \psi(\cdot, z)]\,(\cdot) \in \mathscr{S}^{\mathcal{K}_t}(\mathscr{X})\,, \quad (42)$$

where $\mathcal{K}_t < -\mathcal{M}$ is a self-adjoint integral operator. Hence, by Theorem 7, the semiconvex dual of $\mathcal{S}_t\, \psi(\cdot, z)$ is well-defined for any $z \in \mathscr{X}_0$. This dual is denoted by the functional $B_t(\cdot, z) : \mathscr{X}_0 \mapsto \mathbb{R}$, and is explicitly related to $\mathcal{S}_t\, \psi(\cdot, z)$ by (39). That is,

$$[\mathcal{S}_t\, \psi(\cdot, z)]\,(x) = [\mathcal{D}_\psi^{-1}\, B_t(\cdot, z)]\,(x)\,, \quad (43)$$

$$B_t(x, z) = [\mathcal{D}_\psi\, \mathcal{S}_t\, \psi(\cdot, z)]\,(x)\,. \quad (44)$$

An explicit expression for (44) exists as a consequence of Theorem 6.

Lemma 8: $B_t(y, z)$ is given explicitly by

$$\begin{aligned} B_t(y, z) &= \tfrac{1}{2}\, \langle y, \mathcal{M}\, \mathcal{H}_t\, y\rangle_{\mathscr{X}} + \langle y, (\mathcal{I} + \mathcal{H}_t')\, \mathcal{Q}_t\, z\rangle_{\mathscr{X}} \\ &\quad + \tfrac{1}{2}\, \langle z, (\mathcal{R}_t + \mathcal{Q}_t'\, \mathcal{M}^{-1}\, (\mathcal{I} + \mathcal{H}_t')\, \mathcal{Q}_t)\, z\rangle_{\mathscr{X}}\,, \end{aligned} \quad (45)$$

for any $y, z \in \mathscr{X}_0$, where $\mathcal{H}_t$ is the well-defined operator

$$\mathcal{H}_t \doteq \mathcal{N}_t + \mathcal{N}_t\, \mathcal{N}_t + \mathcal{N}_t\, \mathcal{N}_t\, \mathcal{N}_t + \ldots\,, \qquad \mathcal{N}_t \doteq \mathcal{M}^{-1}\, \mathcal{P}_t\,. \quad (46)$$

This demonstrates that the dual $B_t(y, z)$ is a quadratic functional in $y, z \in \mathscr{X}_0$. This functional may be used as the kernel in the definition of a max-plus integral operator $\mathcal{B}_t$, given by

$$[\mathcal{B}_t\, a]\,(x) \doteq \int_{\mathscr{X}}^{\oplus} B_t(x, z) \otimes a(z)\, dz\,. \quad (47)$$

This operator is instrumental in the construction of the fundamental solution of the Riccati equation (13) of interest.

V. The Fundamental Solution

By definition, a solution $\mathsf{Ric}(\widetilde{M})$ corresponding to the initial condition $\widetilde{M} \in \mathscr{M}$ exists on the the open interval $[0, T_{\widetilde{M}})$. Consequently, both $\mathsf{Ric}(M)$ and $\mathsf{Ric}(\widetilde{M})$ exist on the interval $[0, \widehat{T}]$, where

$$\widehat{T} \doteq \min\left(T_M, T_{\widetilde{M}}\right) \in \mathbb{R}^+\,. \quad (48)$$

Define the functional $\widetilde{\psi} : \mathscr{X} \mapsto \mathbb{R}$ by

$$\widetilde{\psi}(x) \doteq \tfrac{1}{2}\, \langle x, \widetilde{\mathcal{M}}\, x\rangle_{\mathscr{X}}\,. \quad (49)$$

By replacing the terminal cost ψ of (24) with this functional, note that $\mathsf{Ric}(\widetilde{M})$ may be characterized via an optimal control problem in an analogous way to (35) and (37). That is, $\mathsf{Ric}(\widetilde{M})$ is encapsulated via the propagated value $\mathcal{S}_t\, \widetilde{\psi}$ for all $t \in [0, T_{\widetilde{M}})$. Furthermore, this propagated value can be represented in terms of the max-plus integral operator $\mathcal{B}_t$ of (47), as expressed in the following result.

Theorem 9: Consider an initial condition $\widetilde{M} \in \mathscr{M}$ for $\mathsf{Ric}(\widetilde{M})$ such that (17) holds. Let $\widetilde{\psi} : \mathscr{X} \mapsto \mathbb{R}$ denote the functional (49), and $\widehat{T} \in \mathbb{R}^+$ as per (48). Then,

$$\mathcal{S}_t\,\widetilde{\psi} = \mathcal{D}_\psi^{-1}\,\mathcal{B}_t\,\mathcal{D}_\psi\,\widetilde{\psi}\,, \tag{50}$$

for all $t \in [0, \widehat{T})$, in which $\mathcal{D}_\psi^{-1}$, $\mathcal{B}_t$ and $\mathcal{D}_\psi$ are the operators (40), (47), and (38).

Theorem 9 provides a representation for the solution $\mathsf{Ric}(\widetilde{M})$ in terms of operators defined with respect to the particular solution $\mathsf{Ric}(M)$. By inspection of (37) and (50),

$$\begin{aligned}\mathcal{S}_{\tau+t}\,\widetilde{\psi} = \mathcal{S}_\tau\,\mathcal{S}_t\,\widetilde{\psi} &= \mathcal{D}_\psi^{-1}\,\mathcal{B}_\tau\,\mathcal{D}_\psi\,\mathcal{D}_\psi^{-1}\,\mathcal{B}_t\,\mathcal{D}_\psi\,\widetilde{\psi}\\ &= \mathcal{D}_\psi^{-1}\,\mathcal{B}_\tau\,\mathcal{B}_t\,\mathcal{D}_\psi\,\widetilde{\psi}\,,\\ &\equiv \mathcal{D}_\psi^{-1}\,\mathcal{B}_{\tau+t}\,\mathcal{D}_\psi\,\widetilde{\psi}\,.\end{aligned}$$

With $a \doteq \mathcal{D}_\psi\,\widetilde{\psi}$, the above equivalence requires that $\mathcal{B}_{\tau+t}\,a = \mathcal{B}_\tau\,\mathcal{B}_t\,a$. As the terminal cost $\widetilde{\psi}$ is arbitrary in the sense that $\widetilde{M} \in \mathscr{M}$ in (49) is arbitrary modulo (17), this suggests that a semigroup property for $\mathcal{B}_t$ holds. Under appropriate conditions, this is indeed the case. As this semigroup property can be used to propagate any solution, as indicated by (50), $\mathcal{B}_t$ is referred to as a *fundamental solution semigroup*.

Theorem 10: The integral operator (47) satisfies the semigroup property

$$\mathcal{B}_{\tau+t}\,a = \mathcal{B}_\tau\,\mathcal{B}_t\,a\,, \quad a \doteq \mathcal{D}_\psi^{-1}\widetilde{\psi}\,, \tag{51}$$

for any $\widetilde{\psi}$ as per (49) such that $T_{\widetilde{M}} > 0$.

Corollary 11: The kernel B_t of the fundamental solution semigroup $\mathcal{B}_t$ satisfies

$$B_{\tau+t}(y,z) = \int_{\mathscr{X}}^{\oplus} B_\tau(y,\xi) \otimes B_t(\xi,z)\,d\xi \tag{52}$$

for all $t \in [0, \widehat{T})$, $\tau \in [0, \widehat{T} - t)$, y, $z \in \mathscr{X}_0$.

Theorem 10 implies that the solution $\mathsf{Ric}(\widetilde{M})$, encapsulated by $\mathcal{S}_t\,\widetilde{\psi}$ of (50), may be propagated forward in time via Theorem 9 and Corollary 11. Key to the realization of this propagation is the computation of the kernel convolution of (52). Recalling Lemma 8, kernel B_t enjoys an explicit quadratic form (45). Consequently, if the convolution $B_{\tau+t}$ of (52) can be shown to retain this form, then the evolution of B_t can be computed extremely efficiently by conducting repeated explicit maximizations as per (52). This is indeed the case, as the operator encapsulating the quadratic ξ dependence in (52) can be shown to remain invertible through subsequent convolutions. This may be formalized using two auxiliary time-indexed operators $\mathcal{Y}$ and $\mathcal{Z}$ given by

$$\begin{aligned}\mathcal{Y}_\tau &\doteq \mathcal{Q}_\tau'\,(\mathcal{I} + \mathcal{H}_\tau)\,,\\ \mathcal{Z}_{\tau,t} &\doteq \mathcal{R}_\tau + \mathcal{Q}_\tau'\,\mathcal{M}^{-1}(\mathcal{I} + \mathcal{H}_\tau')\,\mathcal{Q}_\tau + \mathcal{M}\,\mathcal{H}_t\,.\end{aligned}$$

Theorem 12: Suppose there exists a $\tau \in [0, \widehat{T})$ such that $\mathcal{Z}_{\tau,\tau} < 0$. Then, the kernel convolution of (52) with $t = k\,\tau$ yields $B_{(k+1)\,\tau}$ of the same form as (45). In particular, the quadratic dependence on ξ in (52) is given by $\frac{1}{2}\langle \xi,\, \mathcal{Z}_{\tau,(k+1)\tau}\,\xi\rangle_{\mathscr{X}}$, where

$$\mathcal{Z}_{\tau,(k+1)\tau} = \mathcal{Z}_{\tau,\tau} - \mathcal{Y}_\tau'\,\mathcal{Z}_{\tau,k\tau}^{-1}\,\mathcal{Y}_\tau \quad k \in \mathscr{N},\ (k+1)\tau < \widehat{T}\,.$$

A. Example

A scalar-valued Riccati equation of the form (13) is considered (i.e. $n = 1$), in which $A = -2$, $\sigma = 1/\sqrt{2}$, $C = 1/3$. The spatial interval Λ is defined by $L = 2$. The Riccati PDE (13) is solved numerically via a Runge-Kutta method (RK45) and via the dual-space propagation of Theorems 9, 10 and 12. Approximation errors are computed with respect to a fine grid RK45 computation. These errors are illustrated in Figure 1, where a considerable computational advantage of the dual-space propagation is demonstrated.

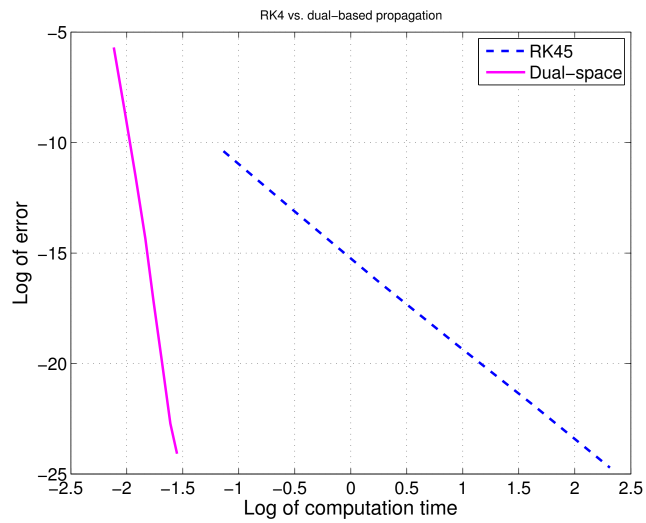

Fig. 1. Approximation error versus computation time for standard (RK45) and dual-propagation methods.

VI. Conclusion

By exploiting connections between a Riccati equation and a specific optimal control problem, dynamic programming can be used to develop an evolution operator for propagating solutions of this equation. By examining this evolution in a dual space (via semiconvexity and the Legendre / Fenchel transform), a dual space evolution operator may be defined that can be applied to any propagate the solution of the Riccati equation corresponding to any initial condition in a particular class. A max-plus method based on this propagation leads to a substantial computational performance improvement in a simple example.

References

[1] G. Bachman and L. Narici. *Functional Analysis*. Dover (orig. Acad. Press, 1966), 2000.

[2] R.F. Curtain and H.J. Zwart. *An introduction to infinite-dimensional linear systems theory*, volume 21 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1995.

[3] P.M. Dower, P.M. Farrell, and D. Nesic. Extremum seeking control of cascaded raman optical amplifiers. *IEEE Transactions on Control Systems Technology*, 16(3):396–407, 2008.

[4] W.M. McEneaney. *Max-plus methods for nonlinear control and estimation*. Systems & Control: Foundations & Application. Birkhauser, 2006.

[5] W.M. McEneaney. A new fundamental solution for differential Riccati equations arising in control. *Automatica*, 44:920–936, 2008.

[6] R.T. Rockafellar. Conjugate duality and optimization. *SIAM Regional Conf. Series in Applied Math.*, 16, 1974.