

## CONTROL STRATEGIES FOR THE FOKKER–PLANCK EQUATION

TOBIAS BREITEN<sup>1,2,a</sup>, KARL KUNISCH<sup>2</sup> AND LAURENT PFEIFFER<sup>1,2</sup>

**Abstract.** Using a projection-based decoupling of the Fokker–Planck equation, control strategies that allow to speed up the convergence to the stationary distribution are investigated. By means of an operator theoretic framework for a bilinear control system, two different feedback control laws are proposed. Projected Riccati and Lyapunov equations are derived and properties of the associated solutions are given. The well-posedness of the closed loop systems is shown and local and global stabilization results, respectively, are obtained. An essential tool in the construction of the controls is the choice of appropriate control shape functions. Results for a two dimensional double well potential illustrate the theoretical findings in a numerical setup.

**Mathematics Subject Classification.** 35Q35, 49J20, 93D05, 93D15.

Received July 6, 2016. Revised June 8, 2017. Accepted June 13, 2017.

### 1. INTRODUCTION

To partially set the stage, let us consider a very large set of dragged Brownian particles, whose motion is described by a stochastic differential equation (SDE) in  $\mathbb{R}^{2n}$  called the Langevin equation:

$$dx(s) = y(s) ds; \quad dy(s) = -\beta y(s) ds + F(x, s) ds + \sqrt{2\beta kT/m} dB(s).$$

Here  $s$  is the time variable,  $\beta > 0$  is a friction parameter,  $m$  the mass of the particle,  $k$  the Boltzmann constant,  $T$  is the temperature, and  $B$  is an  $n$ -dimensional Brownian motion. The force  $F$  is assumed to be related to a potential  $V$ , so that  $F(x, s) = -\nabla V(x, s)$ . For large values of  $\beta$ , the Langevin equation can be approximated by the Smoluchowski equation:

$$dx(t) = -\nabla V(x, t)dt + \sqrt{2\nu} dB_t,$$

where  $t = s/\beta$  and  $\nu = kT/m$ . The probability density function  $\rho$  of the solution to the above equation is the solution to the Fokker–Planck equation:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot J(x, t), \quad \text{where: } J(x, t) = \nu \nabla \rho + \rho \nabla V.$$

---

*Keywords and phrases.* Fokker–Planck equation, bilinear control systems, Lyapunov functions, Riccati equation, Lyapunov equation.

<sup>1</sup> Institute for Mathematics and Scientific Computing, Karl-Franzens-Universität Graz, Heinrichstr. 36, 8010 Graz, Austria.

<sup>a</sup> Corresponding author: [tobias.breiten@uni-graz.at](mailto:tobias.breiten@uni-graz.at);  
[laurent.pfeiffer@uni-graz.at](mailto:laurent.pfeiffer@uni-graz.at)

<sup>2</sup> Institute for Mathematics and Scientific Computing, Karl-Franzens-Universität, Heinrichstr. 36, 8010 Graz, Austria and Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstraße 69, 4040 Linz, Austria.  
[karl.kunisch@uni-graz.at](mailto:karl.kunisch@uni-graz.at)

The variable  $J$  is the probability current. The simplification of the Langevin equation is discussed in ([27], Sect. 10.4, see also [11], Sect. 4.3.4) for details on the connection between SDEs and the Fokker–Planck equation. In this article, we consider the Fokker–Planck equation with reflective boundary conditions:

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \nu \Delta \rho + \nabla \cdot (\rho \nabla V) && \text{in } \Omega \times (0, \infty), \\ 0 &= (\nu \nabla \rho + \rho \nabla V) \cdot \mathbf{n} && \text{on } \Gamma \times (0, \infty), \\ \rho(x, 0) &= \rho_0(x) && \text{in } \Omega, \end{aligned} \quad (1.1)$$

where  $\Omega \subset \mathbb{R}^n$  denotes a bounded domain with smooth boundary  $\Gamma = \partial\Omega$ , and  $\rho_0$  denotes an initial probability distribution with  $\int_{\Omega} \rho_0(x) dx = 1$ . The boundary condition states that the probability current has to vanish in the normal direction on the boundary. This models the fact that any particle reaching the boundary  $\Gamma$  is reflected ([11], Sect. 5.2.3). We refer to [22] for a description of reflected SDEs.

The force  $F$  can be an electric force, created by focusing a laser beam. The obtained structure is called optical tweezer and enables to manipulate microscopic particles, see [16]. We refer to [12] for an overview of feedback control problems in optical trapping. Following the discussion in [14], let us assume that we can interact with the particle by means of an optical tweezer such that the potential  $V$  is of the form

$$V(x, t) = G(x) + \alpha(x)u(t), \quad (1.2)$$

where  $\alpha$  is a control shape function satisfying

$$\alpha \in W^{1,\infty}(\Omega) \cap W^{2,\max(2,n)}(\Omega) \text{ with } \nabla \alpha \cdot \mathbf{n} = 0 \text{ on } \Gamma. \quad (1.3)$$

A more precise characterization of  $\alpha$  will be given in Section 4.2. Thus the control enters in bilinear and separable form into the state equation. While the case that  $G$  is piecewise smooth is certainly of interest, see *e.g.* [27], we focus here on the regular case and assume that  $G \in W^{1,\infty}(\Omega) \cap W^{2,\max(2,n)}(\Omega)$ .

We will consider system (1.1) as an abstract bilinear control system of the form

$$\dot{y} = \mathcal{A}y + u\mathcal{N}y + \mathcal{B}u, \quad y(0) = y_0, \quad (1.4)$$

on an appropriate Hilbert space  $\mathcal{Y}$ . In this setting, the unbounded operator  $\mathcal{A}$  will be the infinitesimal generator of an analytic, strongly continuous semigroup on  $\mathcal{Y}$ . The control objective will be to improve the asymptotic stability of the system to a steady state  $\rho_{\infty}$ . With regard to the design of suboptimal feedback laws for the shifted variable  $y = \rho - \rho_{\infty}$ , we consider two different strategies either of which are based on the linearized version of (1.4). The first feedback law relies on the infinite horizon cost functional and is of the form

$$\mathcal{J}(y, u) = \frac{1}{2} \int_0^{\infty} \langle y, \mathcal{M}y \rangle + |u|^2 dt, \quad (1.5)$$

and is obtained by means of an algebraic operator Riccati equation

$$\mathcal{A}^* \Pi + \Pi \mathcal{A} - \Pi \mathcal{B} \mathcal{B}^* \Pi + \mathcal{M} = 0. \quad (1.6)$$

via  $u = -\mathcal{B}^* \Pi y$ . For the associated nonlinear closed-loop system, we show that for  $\|y_0\| < \varepsilon$ , the system converges to zero with an exponential rate.

As an alternative, we investigate a nonlinear feedback law based on the solution  $\mathcal{T}$  to an operator Lyapunov equation

$$\mathcal{A}^* \mathcal{T} + \mathcal{T} \mathcal{A} + 2\mu I = 0, \quad (1.7)$$

for an appropriately chosen parameter  $\mu > 0$ . Though the control will not be obtained from an optimal control problem, it will be shown to yield a globally, exponentially stabilizing feedback law.

The boundary conditions that we have chosen (for the state equation and for  $\alpha$ ) ensure a mass conservation property. Therefore, the control, which acts inside a differential operator, does not affect the dynamics on a subspace of the state space. As a consequence, we actually have to work with a formulation of (1.4) on the subspace of elements having zero mean and equations (1.6) and (1.7) have to be adapted accordingly. Another

important aspect is the choice of the control potential  $\alpha$  within  $V$ , see (1.2). Our choice is guided by a criterium formulated in the infinite dimensional version of the Hautus criterion.

Besides the large number of publications which consider the Fokker–Planck equations primarily from the stochastic point we mention [20] which gives an analytical framework for Fokker–Planck equations with irregular coefficients, a semigroup approach for Kolmogorov operators with applications to the Fokker–Planck equations [7], and a detailed functions space analysis of steady state solutions in [15]. Concerning stabilization of infinite dimensional systems by means of linearization techniques and the use of Riccati equations to devise feedback mechanisms we refer to *e.g.* [3, 25, 30]. Bilinear control systems arise in the context of parameter estimation problems, for example, and in the control of quantum mechanical equations. Concerning controllability of such systems we refer to the monograph [18], and the references given there.

The construction of suboptimal feedback laws on the basis of applying linear quadratic regulator theory to conveniently defined linearizations has many predecessors. In the context of distributed parameter systems we refer to *e.g.* [5, 25, 26]. In all these papers the control enters linearly into the control system, while it appears in a bilinear fashion in our problem (1.1) with the control entering in the potential  $V$  specified in (1.2). We also stress that the control acts on the differential operator, more precisely on the convection term if (1.1) is considered as a diffusion-convection equation. Hence our problem does not belong to the class of bilinear control problems which was investigated in [4] where the control operator multiplies a bounded term in the state equation.

A brief description of the contents of the paper is given next. Section 2 is devoted to establishing well-posedness of the state equation. We provide the functions space setting in a form which is required for our results on stabilization and as basis for the numerical treatment. Section 3 summarizes some properties of the Fokker–Planck operator with reflecting boundary conditions and provides a succinct splitting of the state equation with respect to the ground state and its complement. A Riccati-based stabilizing feedback mechanism together with an appropriate choice for the control potential is investigated in Section 4. Section 5 provides an alternative which is based on a Lyapunov technique. Loosely speaking, the Riccati-based approach is local and allows an arbitrary decay rate, while the Lyapunov technique is global but it only effects the first eigenspace different from the ground state. Section 6 describes a numerical approach and provides examples which illustrate the theoretical results.

## 2. WELL-POSEDNESS

In this short section we establish basic well-posedness properties of the state equation (1.1). For arbitrary  $T > 0$  we shall refer to  $\rho$  as (variational) solution of (1.1) on  $(0, T)$  if

$$\rho \in W(0, T) = L^2(0, T; H^1(\Omega)) \cap H^1(0, T; (H^1(\Omega))^*)$$

and for a.e.  $t \in (0, T)$

$$\begin{aligned} \langle \rho_t(t), v \rangle + \langle \nu \nabla \rho(t) + \rho(t) \nabla G, \nabla v \rangle + u(t) \langle \rho(t) \nabla \alpha, \nabla v \rangle &= 0 \text{ for all } v \in H^1(\Omega) \\ \rho(0) &= \rho_0. \end{aligned} \tag{2.1}$$

Above  $(H^1(\Omega))^*$  denotes the topological dual of  $H^1(\Omega)$ , with respect to  $L^2(\Omega)$  as pivot space. Let us recall that  $W(0, T) \subset C([0, T], L^2(\Omega))$ , see (*e.g.* [9], Thm. 11.4), so that  $\rho(0)$  is well defined. We also repeat the standing assumption that  $G$  and  $\alpha$  are elements of  $W^{1, \infty}(\Omega) \cap W^{2, \max(2, n)}(\Omega)$ , which in particular implies that the Neumann trace of  $\alpha$  is well-defined. These assumptions will be used in the following basic well-posedness result on the state equation (1.1).

**Proposition 2.1.** *For every  $u \in L^2(0, T)$  and  $\rho_0 \in L^2(\Omega)$  there exists a unique solution to (1.1). If moreover  $\rho_0 \in H^1(\Omega)$  and  $\Delta \alpha \in L^\infty(\Omega)$ , then  $\rho_t \in L^2(0, T; L^2(\Omega))$ ,  $\rho \in C([0, T]; H^1(\Omega))$ ,  $\nabla \cdot (\nu \Delta \rho + \rho \nabla G)$ ,  $\Delta \rho \in L^2(0, T; L^2(\Omega))$  and  $(\nu \nabla \rho + \rho \nabla G) \cdot \mathbf{n} = 0$  in  $L^2(0, T; H^{-1/2}(\Gamma))$ .*

*Proof.* The claim can be verified by a standard Galerkin approximation technique and we therefore only give the necessary *a priori* estimates. Taking the inner products with  $\rho(t)$  in (2.1) we obtain

$$\frac{1}{2} \frac{d}{dt} |\rho(t)|^2 + \nu |\nabla \rho(t)|^2 \leq (|\nabla G|_{L^\infty} + |u(t)| |\nabla \alpha|_{L^\infty}) |\rho(t)| |\nabla \rho(t)|$$

and hence

$$\frac{d}{dt} |\rho(t)|^2 + \nu |\nabla \rho(t)|^2 \leq \frac{1}{\nu} (|\nabla G|_{L^\infty}^2 + |u(t)|^2 |\nabla \alpha|_{L^\infty}^2) |\rho(t)|^2. \tag{2.2}$$

By Gronwall's lemma we have for every  $t \geq 0$

$$|\rho(t)|^2 \leq |\rho_0|^2 \exp \left( \frac{1}{\nu} \int_0^t (|\nabla G|_{L^\infty}^2 + |u(s)|^2 |\nabla \alpha|_{L^\infty}^2) ds \right).$$

Together with (2.2) this implies that  $\rho \in L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T, L^2(\Omega))$ . To verify that  $\rho_t \in L^2(0, T; (H^1(\Omega))^*)$  we recall that  $|\phi|_{(H^1(\Omega))^*} = \sup_{|\psi|_{H^1(\Omega)} \leq 1} \langle \phi, \psi \rangle_{(H^1(\Omega))^*, H^1(\Omega)}$ , for any  $\phi \in (H^1(\Omega))^*$ . We obtain the existence of a constant  $C$  independent of  $\rho, G$ , and  $\alpha$  such that

$$\int_0^T |\rho_t(t)|_{H^1(\Omega)^*} \leq C \int_0^T (|\nabla \rho(t)|^2 + |\nabla G|_{L^\infty}^2 |\rho(t)|^2 + |\nabla \alpha|_{L^\infty}^2 |\rho|_{L^\infty(0, T; L^2(\Omega))}^2 |u(t)|^2) dt.$$

Since the right hand side is bounded we have that  $\rho_t \in L^2(0, T; H^1(\Omega))^*$ , and thus  $\rho \in W(0, T)$ .

To gain extra regularity we set  $v = e^{G/\nu} \rho_t$  in (2.1) and obtain, using  $\nabla \alpha \cdot \mathbf{n} = 0$  on  $\Gamma$ , that

$$|e^{G/2\nu} \rho_t|^2 + \nu \langle e^{-G/\nu} \nabla(e^{G/\nu} \rho), \nabla(e^{G/\nu} \rho_t) \rangle + u(t) \langle \rho(t) \nabla \alpha, \nabla(e^{G/\nu} \rho_t(t)) \rangle = 0$$

and thus

$$\begin{aligned} |e^{G/2\nu} \rho_t|^2 + \frac{\nu}{2} \frac{d}{dt} |e^{-G/2\nu} \nabla(e^{G/\nu} \rho)|^2 &\leq |u(t)| \langle \nabla(\rho(t) \nabla \alpha), e^{G/\nu} \rho_t \rangle \\ &\leq |u(t)| |e^{G/2\nu} \nabla(\rho(t) \nabla \alpha)| |e^{G/2\nu} \rho_t|. \end{aligned}$$

This implies the estimate

$$\begin{aligned} |e^{G/2\nu} \rho_t(t)|^2 + \nu \frac{d}{dt} |e^{-G/2\nu} \nabla(e^{G/\nu} \rho(t))|^2 &\leq |u(t)|^2 |e^{+G/2\nu} (\rho(t) \Delta \alpha + \nabla \rho(t) \cdot \nabla \alpha)|^2 \\ &\leq 2|u(t)|^2 (|e^{+G/2\nu} \rho(t) \Delta \alpha|^2 + |e^{+G/2\nu} \nabla \rho(t) \cdot \nabla \alpha|^2) \\ &\leq 2|u(t)|^2 (|e^{G/2\nu} \Delta \alpha|_{L^\infty}^2 |\rho(t)|^2 + 2|\nabla \alpha|_{L^\infty}^2 |e^{G/2\nu} (\nabla \rho(t) + \frac{1}{\nu} \rho(t) \nabla G)|^2 \\ &\quad + \frac{2}{\nu} |e^{G/2\nu}|_{L^\infty}^2 |\nabla \alpha|_{L^\infty}^2 |\nabla G|_{L^\infty}^2 |\rho(t)|^2). \end{aligned}$$

With  $K_1 = 2|e^{G/2\nu}|_{L^\infty}^2 (|\Delta \alpha|_{L^\infty}^2 + \frac{2}{\nu} |\nabla \alpha|_{L^\infty}^2 + |\nabla G|_{L^\infty}^2)$  and  $K_2 = 4|\nabla \alpha|_{L^\infty}^2$  we have

$$|e^{G/2\nu} \rho_t(t)|^2 + \nu \frac{d}{dt} |e^{-G/2\nu} \nabla(e^{G/\nu} \rho(t))|^2 \leq K_1 |u(t)|^2 |\rho(t)|^2 + K_2 |u(t)|^2 |e^{-G/2\nu} \nabla(e^{G/\nu} \rho(t))|^2.$$

Integration on  $(0, t)$ , with  $t \in (0, T]$  implies that

$$\begin{aligned} &\nu |e^{-G/2\nu} \nabla(e^{G/\nu} \rho(t))|^2 + \int_0^t |e^{G/2\nu} \rho_t(s)|^2 ds \\ &\leq \nu |e^{-G/2\nu} \nabla(e^{G/\nu} \rho_0)|^2 + K_1 |\rho|_{C([0, T], L^2(\Omega))}^2 |u|_{L^2(0, T)}^2 \\ &\quad + K_2 \int_0^t |u(s)|^2 |e^{-G/2\nu} \nabla(e^{G/\nu} \rho(s))|^2 ds. \end{aligned} \tag{2.3}$$

Neglecting for a moment the second term on the left hand side of the inequality and applying Gronwall's inequality implies that  $e^{-G/\nu} \nabla(e^{G/\nu} \rho) \in C([0, T], L^2(\Omega)^n)$  and hence  $\rho \in C([0, T], H^1(\Omega))$ . Using this fact in (2.3) implies that  $\rho_t \in L^2(0, T; L^2(\Omega))$  and hence  $\nabla \cdot (\nu \nabla \rho + \rho \nabla G) \in L^2(0, T; L^2(\Omega))$  from (2.1). Thus

$\nu \nabla \rho + \rho \nabla G \in L^2(0, T; L^2_{\text{div}}(\Omega))$ , and again by (2.1) we have  $(\nu \nabla \rho + \rho \nabla G) \cdot \mathbf{n} = 0$  in  $L^2(0, T; H^{-1/2}(\Gamma))$ , see (e.g. [29], p. 101) as desired. Here  $L^2_{\text{div}}(\Omega)$  denotes the space  $\{\varphi \in L^2(\Omega)^n : \nabla \cdot \varphi \in L^2(\Omega)\}$ . The properties that  $\nabla \cdot (\nu \nabla \rho + \rho \nabla G) \in L^2(0, T; L^2(\Omega))$  and  $\rho \in L^2(0, T; H^1(\Omega))$  can be exploited to obtain that  $\Delta \rho \in L^2(0, T; L^2(\Omega))$ . Since  $\nabla \cdot (\rho \nabla G) = \nabla \rho \cdot \nabla G + \rho \Delta G$  and  $\nabla \rho \cdot \nabla G \in L^2(0, T; L^2(\Omega))$ , by the fact that  $\nabla G \in L^\infty(\Omega)$ , it suffices to argue that  $\rho \Delta G \in L^2(0, T; L^2(\Omega))$ . This follows from the continuous embedding of  $H^1(\Omega)$  into  $L^{(\frac{2n}{n-2})}(\Omega)$  and the Hölder inequality with weights  $p = \frac{n}{n-2}, p' = \frac{n}{2}$ .  $\square$

The solution of the Fokker–Planck equation satisfies structural properties including preservation of probability and nonnegativity which we establish next.

**Proposition 2.2.** *Let  $u \in L^2(0, T)$  and  $\rho_0 \in L^2(\Omega)$ .*

- (i) *For every  $t \in [0, T]$  we have  $\int_\Omega \rho(t) \, dx = \int_{\Omega_0} \rho_0 \, dx$ .*
- (ii) *If  $\rho_0 \geq 0$  a.e. on  $\Omega$ , then  $\rho(x, t) \geq 0$  for all  $t > 0$  and almost all  $x \in \Omega$ .*

*Proof.* Setting  $v = 1$  in (2.1) we obtain the preservation of probability  $\int_\Omega \rho(t) \, dx = \int_\Omega \rho_0 \, dx$  for all  $t \in [0, T]$ .

Turning to the verification of (ii) let us denote by  $\rho = \rho^+ - \rho^-$  the decomposition of the state  $\rho$  into its nonnegative and its negative part, respectively. It then also holds that

$$e^{\frac{c}{\nu}} \rho = \left( e^{\frac{c}{\nu}} \rho \right)^+ - \left( e^{\frac{c}{\nu}} \rho \right)^- = e^{\frac{c}{\nu}} (\rho^+ - \rho^-),$$

and  $\rho^+, \rho^- \in L^2(0, T; H^1(\Omega))$ , see (e.g. [9], Lem. 11.2) Note that we can write  $\nu \nabla \rho + \rho \nabla G = e^{-\frac{c}{\nu}} \nabla (ve^{\frac{c}{\nu}} \rho)$ . Hence by setting  $v = -e^{\frac{c}{\nu}} \rho^-$  in (2.1) we obtain that

$$\left\langle \frac{d}{dt} (\rho^+ - \rho^-), -e^{\frac{c}{\nu}} \rho^- \right\rangle = -\nu \left\langle e^{-\frac{c}{\nu}} \nabla \left( e^{\frac{c}{\nu}} \rho \right), \nabla \left( e^{\frac{c}{\nu}} \rho^- \right) \right\rangle + u \left\langle \rho \nabla \alpha, \nabla \left( e^{\frac{c}{\nu}} \rho^- \right) \right\rangle.$$

With ([9], Lem. 11.2) it now follows that

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \left\| e^{\frac{c}{2\nu}} \rho^- \right\|^2 &= \left\langle \frac{d}{dt} \left( e^{\frac{c}{2\nu}} \rho^- \right), e^{\frac{c}{2\nu}} \rho^- \right\rangle \\ &\leq -\nu \left\| e^{-\frac{c}{2\nu}} \nabla \left( e^{\frac{c}{\nu}} \rho^- \right) \right\|^2 + |u| \left\| e^{\frac{c}{2\nu}} \rho^- \nabla \alpha \right\| \left\| e^{-\frac{c}{2\nu}} \nabla \left( e^{\frac{c}{\nu}} \rho^- \right) \right\| \\ &\leq -\nu \left\| e^{-\frac{c}{2\nu}} \nabla \left( e^{\frac{c}{\nu}} \rho^- \right) \right\|^2 + \frac{\nu}{2} \left\| e^{-\frac{c}{2\nu}} \nabla \left( e^{\frac{c}{\nu}} \rho^- \right) \right\|^2 \\ &\quad + \frac{1}{2\nu} |u|^2 \|\nabla \alpha\|_{L^\infty(\Omega)}^2 \left\| e^{\frac{c}{2\nu}} \rho^- \right\|^2 \\ &\leq \frac{1}{2\nu} |u|^2 \|\nabla \alpha\|_{L^\infty(\Omega)}^2 \left\| e^{\frac{c}{2\nu}} \rho^- \right\|^2 \end{aligned}$$

An application of Gronwall’s inequality now yields that from  $\rho^-(0) = 0$ , it follows that  $e^{\frac{c}{2\nu}} \rho^-(t) = 0$ , and hence that  $\rho^-(t) = 0$  for all  $t \geq 0$ .  $\square$

### 3. THE OPERATOR FORM OF THE FOKKER–PLANCK EQUATION

The goal of this section is to formulate (1.1) as an abstract Cauchy problem such that the linearized system can be studied by means of semigroup methods. Hence, let us consider the abstract *bilinear* control system

$$\begin{aligned} \dot{\rho}(t) &= \mathcal{A}\rho(t) + \mathcal{N}\rho(t)u(t), \\ \rho(0) &= \rho_0, \end{aligned} \tag{3.1}$$

where the operators  $\mathcal{A}$  and  $\mathcal{N}$  are defined as follows

$$\begin{aligned} \mathcal{A}: \mathcal{D}(\mathcal{A}) \subset L^2(\Omega) &\rightarrow L^2(\Omega), \\ \mathcal{D}(\mathcal{A}) &= \{ \rho \in H^2(\Omega) \mid (\nu \nabla \rho + \rho \nabla G) \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma \}, \\ \mathcal{A}\rho &= \nu \Delta \rho + \nabla \cdot (\rho \nabla G), \\ \mathcal{N}: H^1(\Omega) &\rightarrow L^2(\Omega), \quad \mathcal{N}\rho = \nabla \cdot (\rho \nabla \alpha). \end{aligned} \tag{3.2}$$

Let us recall [1] that we have the following embeddings

$$W^{2,2}(\Omega) \hookrightarrow \begin{cases} C(\Omega) & \text{if } n = 1, 2, 3, \\ L^q(\Omega), \quad q \in [1, \infty) & \text{if } n = 4, \\ L^{\frac{2n}{n-4}}(\Omega) & \text{if } n \geq 5. \end{cases}$$

Since by assumption  $\alpha, G \in W^{1,\infty}(\Omega) \cap W^{2,\max(2,n)}(\Omega)$ , a short computation involving the Hölder inequality shows that  $\mathcal{A}$  and  $\mathcal{N}$  are well-defined. Its  $L^2(\Omega)$ -adjoints are now given by

$$\begin{aligned} \mathcal{A}^*: \mathcal{D}(\mathcal{A}^*) \subset L^2(\Omega) &\rightarrow L^2(\Omega), \\ \mathcal{D}(\mathcal{A}^*) &= \{ \varphi \in H^2(\Omega) \mid (\nu \nabla \varphi) \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma \}, \\ \mathcal{A}^*\varphi &= \nu \Delta \varphi - \nabla G \cdot \nabla \varphi, \\ \mathcal{N}^*: H^1(\Omega) &\rightarrow L^2(\Omega), \quad \mathcal{N}^*\varphi = -\nabla \varphi \cdot \nabla \alpha. \end{aligned} \tag{3.3}$$

We emphasize that, due to (1.3), a solution  $\rho \in \mathcal{D}(\mathcal{A})$  of (3.1) automatically satisfies the zero flux boundary conditions of (1.1).

### 3.1. Properties of the Fokker–Planck operator

For what follows, it will be convenient to summarize some known qualitative properties of the uncontrolled Fokker–Planck equation

$$\dot{\rho}(t) = \mathcal{A}\rho(t), \quad \rho(0) = \rho_0, \tag{3.4}$$

compare ([27], Chap. 5/6). For the sake of a self-contained presentation, we also provide the proofs for the statements.

Following [27], let us introduce  $\Phi(x) = \log \nu + \frac{G(x)}{\nu}$ , such that  $e^{\frac{\Phi(x)}{2}} = \sqrt{\nu} e^{\frac{G(x)}{2\nu}}$ . Further, define the operator

$$\begin{aligned} \mathcal{A}_s: \mathcal{D}(\mathcal{A}_s) \subset L^2(\Omega) &\rightarrow L^2(\Omega), \\ \mathcal{D}(\mathcal{A}_s) &= \left\{ \varrho \in H^2(\Omega) \mid \left( \nu \nabla \varrho + \frac{1}{2} \varrho \nabla G \right) \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma \right\}, \\ \mathcal{A}_s &= e^{\frac{\Phi}{2}} \mathcal{A} e^{-\frac{\Phi}{2}}. \end{aligned} \tag{3.5}$$

A straightforward calculation using  $\nu \nabla \Phi = \nabla G$  shows that

$$\mathcal{A} \left( e^{-\frac{\Phi}{2}} \rho \right) = \nu e^{-\frac{\Phi}{2}} \left( \Delta \rho + \frac{1}{2} \rho \Delta \Phi - \frac{1}{4} \rho \nabla \Phi \cdot \nabla \Phi \right).$$

Using the previously mentioned embeddings and Hölder inequality, it can be shown that  $\mathcal{A}_s \rho = e^{\frac{\Phi}{2}} \mathcal{A} e^{-\frac{\Phi}{2}} \rho$  is indeed in  $L^2(\Omega)$  for  $\rho \in H^2(\Omega)$ . Moreover, it turns out that the spectrum of  $\mathcal{A}$  coincides with that of  $\mathcal{A}_s$  and, in particular, is discrete.

**Lemma 3.1.** *The operator  $\mathcal{A}_s$  is self-adjoint. The spectrum  $\sigma(\mathcal{A}_s)$  of  $\mathcal{A}_s$  consists of pure point spectrum contained in  $\overline{\mathbb{R}}_-$  with  $0 \in \sigma(\mathcal{A}_s)$  and only accumulation point  $-\infty$ . The eigenfunctions  $\{\psi_i\}_{i=0}^\infty$  form a complete orthogonal set. Further  $\sigma(\mathcal{A}_s) = \sigma(\mathcal{A})$  and  $\psi_i$  is an eigenfunction of  $\mathcal{A}$  if and only if  $e^{\frac{\Phi}{2}} \psi_i$  is an eigenfunction of  $\mathcal{A}_s$ . Similarly,  $\psi_i$  is an eigenfunction of  $\mathcal{A}$  if and only if  $e^{\Phi} \psi_i$  is an eigenfunction of  $\mathcal{A}^*$ . Finally,  $\rho_\infty = e^{-\Phi}$  is an eigenfunction of  $\mathcal{A}$  associated to the eigenvalue 0.*

*Proof.* Let  $\varrho_1, \varrho_2 \in \mathcal{D}(\mathcal{A}_s)$ . Since  $\nu \nabla \Phi = \nabla G$ , we conclude that  $\mathcal{A}_s \varrho_1$  is given as

$$\begin{aligned} \mathcal{A}_s \varrho_1 &= e^{\frac{\Phi}{2}} \mathcal{A} e^{-\frac{\Phi}{2}} \varrho_1 = \nu e^{\frac{\Phi}{2}} \left( \Delta \left( e^{-\frac{\Phi}{2}} \varrho_1 \right) + \nabla \cdot \left( e^{-\frac{\Phi}{2}} \varrho_1 \nabla \Phi \right) \right) \\ &= \nu e^{\frac{\Phi}{2}} \nabla \cdot \left( \nabla \left( e^{-\frac{\Phi}{2}} \varrho_1 \right) + e^{-\frac{\Phi}{2}} \varrho_1 \nabla \Phi \right) \\ &= \nu e^{\frac{\Phi}{2}} \nabla \cdot \left( e^{-\Phi} \nabla \left( e^{\frac{\Phi}{2}} \varrho_1 \right) \right). \end{aligned} \tag{3.6}$$

Similarly we obtain that

$$0 = \left( \nu \nabla \varrho_1 + \frac{1}{2} \varrho_1 \nabla G \right) \cdot \mathbf{n} = \left( \nu e^{-\frac{\Phi}{2}} \nabla \left( e^{\frac{\Phi}{2}} \varrho_1 \right) \right) \cdot \mathbf{n} \quad \text{on } \Gamma.$$

Thus, it holds that

$$\begin{aligned} \int_{\Omega} \varrho_2 \mathcal{A}_s \varrho_1 \, dx &= \int_{\Omega} \varrho_2 \left( \nu e^{\frac{\Phi}{2}} \nabla \cdot \left( e^{-\Phi} \nabla \left( e^{\frac{\Phi}{2}} \varrho_1 \right) \right) \right) \, dx \\ &= \int_{\Gamma} \left( \varrho_2 \nu e^{-\frac{\Phi}{2}} \nabla \left( e^{\frac{\Phi}{2}} \varrho_1 \right) \right) \cdot \mathbf{n} \, ds - \nu \int_{\Omega} e^{-\Phi} \nabla \left( e^{\frac{\Phi}{2}} \varrho_2 \right) \nabla \left( e^{\frac{\Phi}{2}} \varrho_1 \right) \, dx \\ &= \int_{\Gamma} \left( \varrho_2 \nu e^{-\frac{\Phi}{2}} \nabla \left( e^{\frac{\Phi}{2}} \varrho_1 \right) \right) \cdot \mathbf{n} \, ds - \nu \int_{\Gamma} \left( \varrho_1 e^{-\frac{\Phi}{2}} \nabla \left( e^{\frac{\Phi}{2}} \varrho_2 \right) \right) \cdot \mathbf{n} \, ds \\ &\quad + \nu \int_{\Omega} \varrho_1 e^{\frac{\Phi}{2}} \nabla \cdot \left( e^{-\Phi} \nabla \left( e^{\frac{\Phi}{2}} \varrho_2 \right) \right) \, dx \\ &= \int_{\Omega} \varrho_1 \mathcal{A}_s \varrho_2 \, dx. \end{aligned}$$

As a consequence we have that  $\mathcal{A}_s^* = \mathcal{A}_s$ , thus it is a self-adjoint and closed operator in  $L^2(\Omega)$ . By (3.6), we also have that for each  $\varrho \in \mathcal{D}(\mathcal{A}_s)$ :

$$\int_{\Omega} \varrho \mathcal{A}_s \varrho \, dx = -\nu \int_{\Omega} |\nabla \left( e^{\frac{\Phi}{2}} \varrho \right)|^2 e^{-\Phi} \, dx \leq 0,$$

and hence  $\mathcal{A}_s$  is a negative operator. It follows that there exists  $\beta \in \mathbb{R}_+$  which is in the resolvent set of  $-\mathcal{A}_s$ . Consequently, for each  $f \in L^2(\Omega)$

$$(-\mathcal{A}_s + \beta I)\varrho = f$$

has a unique solution  $\varrho \in \mathcal{D}(\mathcal{A}_s)$  depending continuously on  $f$ . We observe that  $\varrho$  is the solution to

$$\begin{aligned} -\nu \Delta \varrho - \frac{1}{2} e^{\frac{\Phi}{2}} \nabla \cdot \left( e^{-\frac{\Phi}{2}} \varrho \nabla G \right) + \beta \varrho + \frac{1}{2} \nabla \varrho \cdot \nabla G &= f && \text{in } \Omega, \\ \left( \nu \nabla \varrho + \frac{1}{2} \varrho \nabla G \right) \cdot \mathbf{n} &= 0 && \text{on } \Gamma. \end{aligned}$$

Testing this equation with  $\varrho$  we obtain

$$\nu \int_{\Omega} |\nabla \varrho|^2 \, dx + \frac{1}{2} \int_{\Omega} e^{-\frac{\Phi}{2}} \varrho \nabla \left( e^{\frac{\Phi}{2}} \varrho \right) \cdot \nabla G \, dx + \frac{1}{2} \int_{\Omega} \varrho \nabla \varrho \cdot \nabla G \, dx = \int_{\Omega} (f - \beta \varrho) \varrho \, dx$$

and hence

$$\begin{aligned} \nu |\nabla \varrho|_{L^2(\Omega)}^2 &\leq |\nabla G|_{L^\infty(\Omega)} |e^{-\frac{\Phi}{2}}|_{L^\infty(\Omega)} |e^{\frac{\Phi}{2}}|_{L^\infty(\Omega)} |\varrho|_{L^2(\Omega)} (|\nabla \varrho|_{L^2(\Omega)} + |\varrho|_{L^2(\Omega)}) \\ &\quad + (|f|_{L^2(\Omega)} + \beta |\varrho|_{L^2(\Omega)}) |\varrho|_{L^2(\Omega)}. \end{aligned}$$

Together with the continuous dependence of  $\varrho \in L^2(\Omega)$  on  $f$ , we deduce the existence of a constant  $K$  such that

$$|\varrho|_{H^1(\Omega)} \leq K |f|_{L^2(\Omega)}.$$

Thus  $-\mathcal{A}_s + \beta I$  has a compact resolvent as operator in  $L^2(\Omega)$ . Consequently, the spectrum of  $\mathcal{A}_s$  consists entirely of isolated eigenvalues with finite multiplicity in  $\mathbb{R}_-$ , with only accumulation point  $-\infty$ , see, (e.g., [17], Chap. 3).

The relation between the eigenfunctions of  $\mathcal{A}$  and  $\mathcal{A}_s$  follow immediately from the definition of the operator  $\mathcal{A}_s$ . Moreover, note that by (3.6) it holds that  $e^{-\frac{\phi}{2}}$  is an eigenfunction of  $\mathcal{A}_s$  associated to the eigenvalue 0. The associated eigenfunctions of  $\mathcal{A}$  and  $\mathcal{A}^*$  are  $\rho_\infty = e^{-\phi}$  and the constant function  $\mathbb{1}$  with value 1, respectively.  $\square$

Since  $\mathcal{A}_s$  is self-adjoint, it follows from

$$\int_{\Omega} \varrho \mathcal{A}_s \varrho \, dx \leq 0 \quad \text{for all } \varrho \in \mathcal{D}(\mathcal{A}_s)$$

that  $\mathcal{A}_s$  is dissipative, (see [24], Chap. 1, Def. 4.1). Together with the fact that the range of  $\beta I - \mathcal{A}_s$  is surjective, the Lumer–Phillips Theorem ([24], Chap. 1, Thm. 4.3) implies that  $\mathcal{A}_s$  generates a semigroup of contractions on  $L^2(\Omega)$ . Consequently  $\mathcal{A}$  generates a semigroup  $S(t)$  of class  $G(M, 0)$  in  $L^2(\Omega)$ , i.e.  $\|S(t)\| \leq M$  for all  $t$ . Moreover,  $S(t)$  is an analytic semigroup, see (e.g., [28], Sect. 5.4) and the mild solution to (3.4) is given by

$$\rho(t) = S(t)\rho_0. \quad (3.7)$$

### 3.2. Decoupling the Fokker–Planck equation

According to Lemma 3.1, it is clear that  $\rho_\infty = e^{-\phi}$  is a stationary solution of (1.1). From now on, let us assume that  $\rho_\infty$  is normalized such that  $\int_{\Omega} \rho_\infty \, dx = 1$ . While  $\rho_\infty$  is asymptotically stable, the convergence rate (given by the second eigenvalue) can be undesirably slow. An approximation of the convergence rate for small values of  $\nu$  is given by:  $Ce^{-\Delta_G/\nu}$ , where  $C > 0$  is a constant and where the constant  $\Delta_G$  – called energy activation – is the highest potential barrier that the particle has to overcome to reach the most stable equilibrium. This estimate is proved in ([23], p. 251) for 2-dimensional infinite potential fields. The case of a bistable double-well potential with reflecting conditions (in dimension 1) is also treated in ([27], Sect. 5.10.2).

Following similar works [26, 30], we subsequently study the applicability of a Riccati-based feedback law obtained from a suitable stabilization problem. Starting from (3.1), let us introduce the shifted state  $y := \rho - \rho_\infty$ . Using that  $\mathcal{A}\rho_\infty = 0$ , we obtain the transformed system

$$\begin{aligned} \dot{y}(t) &= \mathcal{A}y(t) + \mathcal{N}y(t)u(t) + \mathcal{B}u(t), \\ y(0) &= \rho_0 - \rho_\infty, \end{aligned} \quad (3.8)$$

with  $\mathcal{B} = \mathcal{N}\rho_\infty$ . Here, the control operator  $\mathcal{B}$  and its adjoint are defined as

$$\begin{aligned} \mathcal{B}: \mathbb{R} &\rightarrow L^2(\Omega), \quad \mathcal{B}c = c\mathcal{N}\rho_\infty, \\ \mathcal{B}^*: L^2(\Omega) &\rightarrow \mathbb{R}, \quad \mathcal{B}^*v = \langle \mathcal{N}\rho_\infty, v \rangle. \end{aligned}$$

For our feedback design, it will be convenient to work with a decoupled version of (3.8). We therefore introduce the projection  $\mathcal{P}$  onto  $\mathbb{1}^\perp$  along  $\rho_\infty$

$$\begin{aligned} \mathcal{P}: L^2(\Omega) &\rightarrow L^2(\Omega), \quad \mathcal{P}y = y - \int_{\Omega} y \, dx \, \rho_\infty, \\ \text{im}(\mathcal{P}) &= \left\{ v \in L^2(\Omega) : \int_{\Omega} v \, dx = 0 \right\}, \quad \ker(\mathcal{P}) = \text{span} \{ \rho_\infty \}. \end{aligned}$$

Hence, the complementary projection  $\mathcal{Q}$  is given as

$$\begin{aligned} \mathcal{Q}: L^2(\Omega) &\rightarrow L^2(\Omega), \quad \mathcal{Q}y = (I - \mathcal{P})y = \int_{\Omega} y \, dx \, \rho_\infty, \\ \text{im}(\mathcal{Q}) &= \ker(\mathcal{P}), \quad \ker(\mathcal{Q}) = \text{im}(\mathcal{P}). \end{aligned}$$



With these definitions, the  $L^2(\Omega)$  adjoint of  $\mathcal{P}$  is the projection  $\mathcal{P}^*$  onto  $\rho_\infty^\perp$  along  $\mathbb{1}$

$$\begin{aligned} \mathcal{P}^* : L^2(\Omega) &\rightarrow L^2(\Omega), \quad \mathcal{P}^* y = y - \int_\Omega \rho_\infty y \, dx \, \mathbb{1}, \\ \text{im}(\mathcal{P}^*) &= \left\{ v \in L^2(\Omega) : \int_\Omega \rho_\infty v \, dx = 0 \right\}, \quad \ker(\mathcal{P}^*) = \{ \mathbb{1} \}. \end{aligned}$$

Finally, the complementary projection  $\mathcal{Q}^*$  reads

$$\begin{aligned} \mathcal{Q}^* : L^2(\Omega) &\rightarrow L^2(\Omega), \quad \mathcal{Q}^* y = \int_\Omega \rho_\infty y \, dx \, \mathbb{1}, \\ \text{im}(\mathcal{Q}^*) &= \ker(\mathcal{P}^*), \quad \ker(\mathcal{Q}^*) = \text{im}(\mathcal{P}^*). \end{aligned}$$

We now can decompose our state space as follows

$$\begin{aligned} \mathcal{Y} = L^2(\Omega) &= \text{im}(\mathcal{P}) \oplus \text{im}(\mathcal{Q}) =: \mathcal{Y}_\mathcal{P} \oplus \mathcal{Y}_\mathcal{Q}, \\ y &= y_\mathcal{P} + y_\mathcal{Q} = \mathcal{P}y + \mathcal{Q}y, \quad y \in L^2(\Omega). \end{aligned} \tag{3.9}$$

This results in the following decomposition of (3.8)

$$\begin{aligned} \dot{y}_\mathcal{P} + \dot{y}_\mathcal{Q} &= \mathcal{A}(y_\mathcal{P} + y_\mathcal{Q}) + \mathcal{N}(y_\mathcal{P} + y_\mathcal{Q})u + \mathcal{B}u \\ y_\mathcal{P}(0) &= \mathcal{P}\rho_0, \quad y_\mathcal{Q}(0) = \mathcal{Q}\rho_0 - \rho_\infty. \end{aligned}$$

Applying respectively  $\mathcal{P}$  and  $\mathcal{Q}$  to this equation yields

$$\begin{pmatrix} \dot{y}_\mathcal{P} \\ \dot{y}_\mathcal{Q} \end{pmatrix} = \begin{pmatrix} \mathcal{P}\mathcal{A} & \mathcal{P}\mathcal{A} \\ \mathcal{Q}\mathcal{A} & \mathcal{Q}\mathcal{A} \end{pmatrix} \begin{pmatrix} y_\mathcal{P} \\ y_\mathcal{Q} \end{pmatrix} + \begin{pmatrix} \mathcal{P}\mathcal{N} & \mathcal{P}\mathcal{N} \\ \mathcal{Q}\mathcal{N} & \mathcal{Q}\mathcal{N} \end{pmatrix} \begin{pmatrix} y_\mathcal{P} \\ y_\mathcal{Q} \end{pmatrix} u + \begin{pmatrix} \mathcal{P}\mathcal{B} \\ \mathcal{Q}\mathcal{B} \end{pmatrix} u. \tag{3.10}$$

Let us note that  $\mathcal{A}\rho_\infty = 0$ ,  $\mathcal{A}^*\mathbb{1} = 0$  and  $\mathcal{N}^*\mathbb{1} = 0$ . For  $y_\mathcal{P} \in \text{im}(\mathcal{P}) \cap \mathcal{D}(\mathcal{A})$ ,  $y_\mathcal{Q} \in \text{im}(\mathcal{Q}) \cap \mathcal{D}(\mathcal{A})$  and  $v \in \mathcal{D}(\mathcal{A}^*)$ , observe that

$$\langle \mathcal{A}y_\mathcal{Q}, v \rangle = 0, \quad \langle \mathcal{Q}\mathcal{A}y_\mathcal{P}, v \rangle = \langle y_\mathcal{P}, \mathcal{A}^*\mathcal{Q}^*v \rangle = 0.$$

For  $y \in H^1(\Omega)$  and  $v \in L^2(\Omega)$ ,

$$\begin{aligned} \langle \mathcal{Q}\mathcal{N}y, v \rangle &= \langle y, \mathcal{N}^*\mathcal{Q}^*v \rangle = 0, \\ \langle \mathcal{Q}\mathcal{B}, v \rangle &= \langle \mathcal{Q}\mathcal{N}\rho_\infty, v \rangle = \langle \rho_\infty, \mathcal{N}^*\mathcal{Q}^*v \rangle = 0. \end{aligned}$$

Hence, we have the identities:

$$\begin{aligned} \mathcal{P}\mathcal{A} &= \mathcal{A} \text{ (on } \mathcal{D}(\mathcal{A})\text{)}, \quad \mathcal{Q}\mathcal{A} = 0 \text{ (on } \mathcal{D}(\mathcal{A})\text{)}, \quad \mathcal{P}\mathcal{N} = \mathcal{N} \text{ (on } H^1(\Omega)\text{)} \\ \mathcal{Q}\mathcal{N} &= 0 \text{ (on } H^1(\Omega)\text{)}, \quad \mathcal{P}\mathcal{B} = \mathcal{B} \text{ (on } \mathbb{R}\text{)}, \quad \mathcal{Q}\mathcal{B} = 0 \text{ (on } \mathbb{R}\text{)}. \end{aligned} \tag{3.11}$$

As a consequence, (3.10) simplifies as follows:

$$\begin{pmatrix} \dot{y}_\mathcal{P} \\ \dot{y}_\mathcal{Q} \end{pmatrix} = \begin{pmatrix} \mathcal{P}\mathcal{A} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_\mathcal{P} \\ y_\mathcal{Q} \end{pmatrix} + u \begin{pmatrix} \mathcal{P}\mathcal{N} & \mathcal{P}\mathcal{N} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_\mathcal{P} \\ y_\mathcal{Q} \end{pmatrix} + u \begin{pmatrix} \mathcal{P}\mathcal{B} \\ 0 \end{pmatrix}$$

hence,

$$\begin{aligned} \dot{y}_\mathcal{P} &= \mathcal{A}y_\mathcal{P} + \mathcal{N}y_\mathcal{P}u + \mathcal{N}(\mathcal{Q}\rho_0 - \rho_\infty)u + \mathcal{B}u, \quad y_\mathcal{P}(0) = \mathcal{P}\rho_0, \\ y_\mathcal{Q}(t) &= \mathcal{Q}\rho_0 - \rho_\infty, \quad t \geq 0. \end{aligned}$$

By definition of  $\mathcal{B}$  and the fact that  $\int_{\Omega} \rho_0 \, dx = 1$ , we finally obtain:

$$\begin{aligned} \dot{y}_{\mathcal{P}} &= \widehat{\mathcal{A}}y_{\mathcal{P}} + \widehat{\mathcal{N}}y_{\mathcal{P}}u + \widehat{\mathcal{B}}u, \quad y_{\mathcal{P}}(0) = \mathcal{P}\rho_0, \\ y_{\mathcal{Q}}(t) &= \mathcal{Q}\rho_0 - \rho_{\infty} = 0, \quad t \geq 0, \end{aligned} \tag{3.12}$$

where  $I_{\mathcal{P}}: \mathcal{Y}_{\mathcal{P}} \rightarrow \mathcal{Y}$  denotes the injection of  $\mathcal{Y}_{\mathcal{P}}$  into  $\mathcal{Y}$  and

$$\begin{aligned} \widehat{\mathcal{A}} &= \mathcal{A}I_{\mathcal{P}} \text{ with } \mathcal{D}(\widehat{\mathcal{A}}) = \mathcal{D}(\mathcal{A}) \cap \mathcal{Y}_{\mathcal{P}}, \\ \widehat{\mathcal{N}} &= \mathcal{N}I_{\mathcal{P}} \text{ with } \mathcal{D}(\widehat{\mathcal{N}}) = H^1(\Omega) \cap \mathcal{Y}_{\mathcal{P}}, \\ \widehat{\mathcal{B}} &= \mathcal{B} \end{aligned}$$

are operators considered in  $\mathcal{Y}_{\mathcal{P}}$ .

#### 4. A RICCATI-BASED FEEDBACK LAW

##### 4.1. Stabilizing the linearized system

For the linearized decoupled and shifted system

$$\dot{y}_{\mathcal{P}} = (\widehat{\mathcal{A}} + \delta I)y_{\mathcal{P}}(t) + \widehat{\mathcal{B}}u, \quad y_{\mathcal{P}}(0) = \mathcal{P}\rho_0, \tag{4.1}$$

let us focus on the cost functional

$$J(y_{\mathcal{P}}, u) = \frac{1}{2} \int_0^{\infty} \langle y_{\mathcal{P}}(t), \mathcal{M}y_{\mathcal{P}}(t) \rangle_{L^2(\Omega)} \, dt + \frac{1}{2} \int_0^{\infty} |u(t)|^2 \, dt, \tag{4.2}$$

where  $\mathcal{M} \in \mathcal{L}(\mathcal{Y}_{\mathcal{P}})$  is a self-adjoint nonnegative operator on  $\mathcal{Y}_{\mathcal{P}}$  which is such that the pair  $(\mathcal{A}, \mathcal{M})$  is detectable. We denote by  $\Theta$  the orthogonal projection on  $\mathcal{Y}_{\mathcal{P}}$ :

$$\begin{aligned} \Theta: L^2(\Omega) &\rightarrow L^2(\Omega), \quad \Theta y = y - \frac{1}{|\Omega|} \int_{\Omega} y \, dx \, \mathbb{1}, \\ \text{im}(\Theta) &= \text{im}(\mathcal{P}) = \mathcal{Y}_{\mathcal{P}}, \quad \ker(\Theta) = \{\mathbb{1}\}. \end{aligned} \tag{4.3}$$

Note that  $\Theta^* = \Theta$  and, in particular,  $\Theta = I_{\mathcal{P}}^*$ . Let us then define the operator

$$\mathcal{A}^{\sharp}: \mathcal{D}(\mathcal{A}^{\sharp}) \subset \mathcal{Y}_{\mathcal{P}} \rightarrow \mathcal{Y}_{\mathcal{P}}, \quad \mathcal{D}(\mathcal{A}^{\sharp}) = \mathcal{D}(\mathcal{A}^*) \cap \mathcal{Y}_{\mathcal{P}}, \quad \mathcal{A}^{\sharp}\phi = \Theta\mathcal{A}^*\phi.$$

**Lemma 4.1.** *The operator  $\mathcal{A}^{\sharp}$  is the adjoint operator of  $\widehat{\mathcal{A}}$ . Moreover, let  $(\lambda, \phi) \in \mathbb{R} \times \mathcal{D}(\mathcal{A}^{\sharp})$  be such that  $\mathcal{A}^{\sharp}\phi = \lambda\phi$ . Then,  $(\lambda, \mathcal{P}^*\phi)$  is an eigenpair of  $\mathcal{A}^*$ . Conversely, if  $(\lambda, \varphi) \in \mathbb{R} \times \mathcal{D}(\mathcal{A}^*)$  is an eigenpair of  $\mathcal{A}^*$ , then  $(\lambda, \Theta\varphi)$  is an eigenpair of  $\mathcal{A}^{\sharp}$ .*

*Proof.* For  $y_{\mathcal{P}} \in \mathcal{D}(\widehat{\mathcal{A}})$  and  $z_{\mathcal{P}} \in \mathcal{D}(\mathcal{A}^{\sharp})$  it now holds that

$$\langle \widehat{\mathcal{A}}y_{\mathcal{P}}, z_{\mathcal{P}} \rangle = \langle \mathcal{A}I_{\mathcal{P}}y_{\mathcal{P}}, z_{\mathcal{P}} \rangle = \langle y_{\mathcal{P}}, I_{\mathcal{P}}^*\mathcal{A}^*z_{\mathcal{P}} \rangle = \langle y_{\mathcal{P}}, \Theta\mathcal{A}^*z_{\mathcal{P}} \rangle = \langle y_{\mathcal{P}}, \mathcal{A}^{\sharp}z_{\mathcal{P}} \rangle.$$

Note also that

$$\langle \mathcal{P}\widehat{\mathcal{A}}y_{\mathcal{P}}, z_{\mathcal{P}} \rangle = \langle y_{\mathcal{P}}, I_{\mathcal{P}}^*\mathcal{A}^*(z_{\mathcal{P}} - \mathcal{Q}^*z_{\mathcal{P}}) \rangle = \langle y_{\mathcal{P}}, I_{\mathcal{P}}^*\mathcal{A}^*z_{\mathcal{P}} \rangle = \langle y_{\mathcal{P}}, \mathcal{A}^{\sharp}z_{\mathcal{P}} \rangle,$$

such that we conclude that  $\mathcal{A}^{\sharp} = (\mathcal{P}\mathcal{A}I_{\mathcal{P}})^* = \widehat{\mathcal{A}}^*$ . For what follows, let  $y \in \mathcal{D}(\mathcal{A})$  and  $z \in \mathcal{D}(\mathcal{A}^*)$  be given. Since  $I = \mathcal{P}^* + \mathcal{Q}^*$  and  $I = \Theta + (I - \Theta)$ , we then have

$$\langle y, \mathcal{A}^*z \rangle = \langle y, \mathcal{P}^*\Theta\mathcal{A}^*z \rangle + \langle y, \mathcal{P}^*(I - \Theta)\mathcal{A}^*z \rangle + \langle y, \mathcal{Q}^*\mathcal{A}^*z \rangle.$$

Using that  $\text{im}(I - \Theta) = \ker(\mathcal{P}^*)$  and  $\text{im}(\mathcal{Q}) = \{\rho_{\infty}\}$ , we obtain

$$\langle y, \mathcal{A}^*z \rangle = \langle y, \mathcal{P}^*\Theta\mathcal{A}^*z \rangle + \langle \mathcal{A}\mathcal{Q}y, z \rangle = \langle y, \mathcal{P}^*\Theta\mathcal{A}^*z \rangle. \tag{4.4}$$

This yields the following relation between the eigenfunctions of  $\mathcal{A}^*$  and those of  $\mathcal{A}^\sharp$ . Let  $(\lambda, \phi)$  be such that  $\mathcal{A}^\sharp\phi = \lambda\phi$ . It then follows by (4.4) that

$$\mathcal{P}^*(\lambda\phi) = \mathcal{P}^*(\mathcal{A}^\sharp\phi) = \mathcal{P}^*(\Theta\mathcal{A}^*)\phi = \mathcal{A}^*\phi = \mathcal{A}^*(\mathcal{P}^* + \mathcal{Q}^*)\phi = \mathcal{A}^*\mathcal{P}^*\phi.$$

Hence,  $(\lambda, \mathcal{P}^*\phi)$  is an eigenpair of  $\mathcal{A}^*$ . Analogously, assume that  $(\lambda, \varphi)$  satisfies  $\mathcal{A}^*\varphi = \lambda\varphi$ . We now obtain

$$\Theta(\lambda\varphi) = \Theta(\mathcal{A}^*\varphi) = \Theta\mathcal{A}^*(\Theta + (I - \Theta))\varphi = \Theta\mathcal{A}^*\Theta\varphi,$$

implying that  $(\lambda, \Theta\varphi)$  is an eigenpair of  $\mathcal{A}^\sharp$ . □

### 4.2. Stabilizability and the choice of $\alpha$

Let us also note that the adjoint of  $\widehat{\mathcal{B}} = \mathcal{P}\mathcal{B}$  as operator from  $\mathbb{R}$  to  $\mathcal{Y}_{\mathcal{P}}$  is given by  $\widehat{\mathcal{B}}^* = \mathcal{B}^*\mathcal{P}^* = \mathcal{B}^*I_{\mathcal{P}}$  and we drop the notation  $I_{\mathcal{P}}$  below.

Up to this point, we have assumed that  $\alpha \in W^{1,\infty}(\Omega) \cap W^{2,\max(2,n)}(\Omega)$  is such that (1.3) is fulfilled. Let us now provide further details on how to choose  $\alpha$ . It is well-known [10] that the cost functional (4.2) is naturally associated to the following operator Riccati equation

$$(\widehat{\mathcal{A}} + \delta I)^*\widehat{\Pi} + \widehat{\Pi}(\widehat{\mathcal{A}} + \delta I) - \widehat{\Pi}\widehat{\mathcal{B}}\widehat{\mathcal{B}}^*\widehat{\Pi} + \mathcal{M} = 0,$$

which is interpreted in the weak sense, *i.e.*,

$$\begin{aligned} &\langle (\widehat{\mathcal{A}} + \delta I)^*\widehat{\Pi}y_1, y_2 \rangle_{L^2(\Omega)} + \langle \widehat{\Pi}(\widehat{\mathcal{A}} + \delta I)y_1, y_2 \rangle_{L^2(\Omega)} \\ &\quad - \langle \widehat{\mathcal{B}}^*\widehat{\Pi}y_1, \widehat{\mathcal{B}}^*\widehat{\Pi}y_2 \rangle_{\mathbb{R}} + \langle \mathcal{M}y_1, y_2 \rangle_{L^2(\Omega)} = 0, \end{aligned}$$

for all  $y_1, y_2 \in \mathcal{D}(\widehat{\mathcal{A}})$ . In particular, in case the pair  $(\widehat{\mathcal{A}}, \widehat{\mathcal{B}})$  is  $\delta$ -stabilizable (see [10], Def. 5.2.1), there exists a unique nonnegative self-adjoint solution  $\widehat{\Pi} \in \mathcal{L}(\mathcal{Y}_{\mathcal{P}})$  such that

$$\widehat{\mathcal{A}}_{\Pi} := \widehat{\mathcal{A}} + \delta I - \widehat{\mathcal{B}}\widehat{\mathcal{B}}^*\widehat{\Pi}$$

generates an exponentially stable semigroup on  $\mathcal{Y}_{\mathcal{P}}$ , see (*e.g.*, [10], p. 295, [19], p. 125–127 or [6], p. 519), where it is also proved that  $\widehat{\Pi}$  enjoys extra regularity since  $\widehat{\mathcal{A}}$  is analytic. With regard to  $\delta$ -stabilizability of  $(\widehat{\mathcal{A}}, \widehat{\mathcal{B}})$ , assume that  $\varphi_i$  are eigenfunctions of  $\mathcal{A}^*$  associated to the eigenvalues

$$-\delta \leq \lambda_d \leq \dots \leq \lambda_2 < 0 = \lambda_1.$$

With the notation introduced before, consider then the elliptic equation

$$\begin{aligned} \nabla \cdot (\rho_\infty \nabla \alpha) &= \mathcal{P} \sum_{i=2}^d e^{-\Phi} \varphi_i && \text{in } \Omega, \\ (\rho_\infty \nabla \alpha) \cdot \mathbf{n} &= 0 && \text{on } \Gamma. \end{aligned} \tag{4.5}$$

From classical elliptic regularity results, see (*e.g.*, [33], Thm. 3.28/3.29), we conclude that there exists a unique solution  $\alpha \in W^{2,p}(\Omega)/\mathbb{R}$  for any  $p > 0$  to (4.5). In particular,  $\alpha \in W^{1,\infty}(\Omega)/\mathbb{R} \cap W^{2,\max(2,n)}(\Omega)/\mathbb{R}$ . As a consequence of this choice of  $\alpha$ , we obtain the desired stabilizability result.

**Lemma 4.2.** *Let  $\alpha \in W^{1,\infty}(\Omega)/\mathbb{R} \cap W^{2,\max(2,n)}(\Omega)/\mathbb{R}$  denote the unique solution to (4.5). Then the pair  $(\widehat{\mathcal{A}}, \widehat{\mathcal{B}})$  is  $\delta$ -stabilizable.*

*Proof.* We are going to verify the assertion by means of the infinite dimensional Hautus test for stabilizability (see [6], Part V, Prop. 3.3 or [10], Thm. 5.2.11). Hence, we need to show that

$$\ker(\lambda I - \mathcal{A}^\sharp) \cap \ker(\widehat{\mathcal{B}}^*) = \{0\} \quad \text{for } \lambda \in \overline{\mathbb{C}}_{-\delta} \cap \sigma(\mathcal{A}^\sharp),$$

where  $\bar{\mathbb{C}}_{-\delta} = \{\lambda \in \mathbb{C} \mid \operatorname{Re}(\lambda) \geq -\delta\}$ . Let us therefore assume that  $(\lambda_j, \phi_j)$ ,  $j \in \{2, \dots, d\}$  is an eigenpair of  $\mathcal{A}^\#$ . By Lemma 3.1, Lemma 4.1 and (4.5), it follows that

$$\widehat{\mathcal{B}}^* \phi_j = \langle \widehat{\mathcal{B}}, \phi_j \rangle = \left\langle \mathcal{P} \sum_{i=2}^d e^{-\Phi} \varphi_i, \phi_j \right\rangle = \sum_{i=2}^d \langle e^{-\Phi} \varphi_i, \varphi_j \rangle = \|e^{-\frac{\Phi}{2}} \varphi_j\|^2$$

which shows the statement.  $\square$

From now on, we assume that  $\alpha$  is such that the Hautus criterion is satisfied and therefore that  $(\widehat{\mathcal{A}}, \widehat{\mathcal{B}})$  is  $\delta$ -stabilizable.

### 4.3. The Riccati equation

With the notation introduced in (3.9), consider the following two Riccati equations:

$$(\mathcal{A}^* + \delta \mathcal{P}^*) \Pi + \Pi (\mathcal{A} + \delta \mathcal{P}) - \Pi \mathcal{B} \mathcal{B}^* \Pi + \mathcal{P}^* \mathcal{M} \mathcal{P} = 0, \quad \Pi \in \mathcal{L}(\mathcal{Y}), \quad \Pi^* = \Pi \quad (\text{R1})$$

$$(\widehat{\mathcal{A}}^* + \delta I) \widehat{\Pi} + \widehat{\Pi} (\widehat{\mathcal{A}} + \delta I) - \widehat{\Pi} \widehat{\mathcal{B}} \widehat{\mathcal{B}}^* \widehat{\Pi} + \mathcal{M} = 0, \quad \widehat{\Pi} \in \mathcal{L}(\mathcal{Y}_{\mathcal{P}}), \quad \widehat{\Pi}^* = \widehat{\Pi}. \quad (\text{R2})$$

**Lemma 4.3.** *If the operator  $\Pi \in \mathcal{L}(\mathcal{Y})$  is a solution to (R1), then  $\widehat{\Pi} := \Theta \Pi I_{\mathcal{P}} \in \mathcal{L}(\mathcal{Y}_{\mathcal{P}})$  is a solution to (R2) and there exists  $\gamma \in \mathbb{R}$  such that  $\Pi = \mathcal{P}^* \widehat{\Pi} \mathcal{P} + \gamma \mathbb{1} \mathbb{1}^*$ . Conversely, if  $\widehat{\Pi}$  is a solution to (R2), then for all  $\gamma \in \mathbb{R}$ ,  $\Pi = \mathcal{P}^* \widehat{\Pi} \mathcal{P} + \gamma \mathbb{1} \mathbb{1}^*$  is a solution to (R1).*

*Proof.* Let us define:

$$\mathcal{R}: y \in \mathcal{Y} \mapsto (\mathcal{P}y, \langle \mathbb{1}, y \rangle) \in \mathcal{Y}_{\mathcal{P}} \times \mathbb{R}.$$

The operator  $\mathcal{R}$  is a homeomorphism. Note that for all  $(z, \alpha) \in \mathcal{Y}_{\mathcal{P}} \times \mathbb{R}$  and for all  $y \in \mathcal{Y}$ ,

$$\mathcal{R}^{-1}(z, \alpha) = z + \alpha \rho_{\infty}, \quad \mathcal{R}^*(z, \alpha) = \mathcal{P}^* z + \alpha \mathbb{1}, \quad \mathcal{R}^{-*} y = (\Theta y, \langle \rho_{\infty}, y \rangle).$$

Let  $\Pi \in \mathcal{L}(\mathcal{Y})$  be a solution to (R1) and define  $\widetilde{\Pi} = \mathcal{R}^{-*} \Pi \mathcal{R}^{-1} \in \mathcal{L}(\mathcal{Y}_{\mathcal{P}} \times \mathbb{R})$ . The operator  $\widetilde{\Pi}$  is a solution to the following equation:

$$(\widetilde{\mathcal{A}} + \delta \widetilde{\mathcal{P}})^* \widetilde{\Pi} + \widetilde{\Pi} (\widetilde{\mathcal{A}} + \delta \widetilde{\mathcal{P}}) - \widetilde{\Pi} \widetilde{\mathcal{B}} \widetilde{\mathcal{B}}^* \widetilde{\Pi} + \widetilde{\mathcal{M}} = 0, \quad (\text{4.6})$$

where:

$$\widetilde{\mathcal{M}} = \mathcal{R}^{-*} \mathcal{P}^* \mathcal{M} \mathcal{P} \mathcal{R}^{-1}, \quad \widetilde{\mathcal{A}} = \mathcal{R} \mathcal{A} \mathcal{R}^{-1}, \quad \widetilde{\mathcal{P}} = \mathcal{R} \mathcal{P} \mathcal{R}^{-1}, \quad \text{and} \quad \widetilde{\mathcal{B}} = \mathcal{R} \mathcal{B}.$$

We represent any operator  $\mathcal{X} \in \mathcal{L}(\mathcal{Y}_{\mathcal{P}} \times \mathbb{R})$  as follows:  $\mathcal{X} = \begin{pmatrix} \mathcal{X}_{11} & \mathcal{X}_{12} \\ \mathcal{X}_{21} & \mathcal{X}_{22} \end{pmatrix}$ , where  $\mathcal{X}_{11} \in \mathcal{L}(\mathcal{Y}_{\mathcal{P}})$ ,  $\mathcal{X}_{12} \in \mathcal{Y}_{\mathcal{P}}$ ,  $\mathcal{X}_{21} \in \mathcal{Y}_{\mathcal{P}}^*$ , and  $\mathcal{X}_{22} \in \mathbb{R}$  are uniquely defined by the relation:

$$\mathcal{X}(y, \beta) = (\mathcal{X}_{11}y + \mathcal{X}_{12}\beta, \mathcal{X}_{21}y + \mathcal{X}_{22}\beta), \quad \forall (y, \beta) \in \mathcal{Y}_{\mathcal{P}} \times \mathbb{R}.$$

One can easily check with (3.11) that:

$$\widetilde{\mathcal{A}} = \begin{pmatrix} \widehat{\mathcal{A}} & 0 \\ 0 & 0 \end{pmatrix}, \quad \widetilde{\mathcal{P}} = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad \widetilde{\mathcal{M}} = \begin{pmatrix} \mathcal{M} & 0 \\ 0 & 0 \end{pmatrix}, \quad \widetilde{\mathcal{B}} \widetilde{\mathcal{B}}^* = \begin{pmatrix} \mathcal{B} \mathcal{B}^* & 0 \\ 0 & 0 \end{pmatrix}. \quad (\text{4.7})$$

Combining (4.6) and (4.7), we obtain that  $\widetilde{\Pi}_{11}$  is a solution to (R2). Moreover,

$$(\widehat{\mathcal{A}}^* + \delta I) \widetilde{\Pi}_{12} - \widetilde{\Pi}_{11} \mathcal{B} \mathcal{B}^* \widetilde{\Pi}_{12} = 0 \quad \text{and} \quad \widetilde{\Pi}_{12}^* \mathcal{B} \mathcal{B}^* \widetilde{\Pi}_{12} = 0.$$

Thus,  $\mathcal{B}^* \widetilde{\Pi}_{12} = 0$  and  $(\widehat{\mathcal{A}}^* + \delta I) \widetilde{\Pi}_{12} = 0$ . As a consequence of the Hautus criterion,  $\widetilde{\Pi}_{12} = 0$ . Setting  $\widehat{\Pi} = \widetilde{\Pi}_{11}$  and  $\gamma = \widetilde{\Pi}_{22}$ , one can easily check that:  $\Pi = \mathcal{R}^* \widetilde{\Pi} \mathcal{R} = \mathcal{P}^* \widehat{\Pi} \mathcal{P} + \gamma \mathbb{1} \mathbb{1}^*$ . The converse implication can be proved in a similar manner.  $\square$

**Lemma 4.4.** *There exists a unique non-negative self-adjoint operator  $\Pi$  solution to (R1) such that  $\Pi \rho_{\infty} = 0$ .*

*Proof.* Let  $\Pi$  be defined by  $\Pi = \mathcal{P}^* \widehat{\Pi} \mathcal{P}$ , where  $\widehat{\Pi}$  is the unique non-negative solution to (R2). By Lemma 4.3,  $\Pi$  is a solution to (R1) and clearly,  $\Pi$  is non-negative and  $\Pi \rho_{\infty} = 0$ . Now, let  $\Pi'$  be a non-negative self-adjoint operator, solution to (R1), and such that  $\Pi' \rho_{\infty} = 0$ . By Lemma 4.3, there exist an operator  $\widehat{\Pi}'$ , solution to (R2)

and  $\gamma \in \mathbb{R}$  such that  $\Pi' = \mathcal{P}^* \widehat{\Pi}' \mathcal{P} + \gamma \mathbb{1} \mathbb{1}^*$ . Since  $\Pi' \rho_\infty = 0$ , we have:  $0 = \mathcal{P}^* \widehat{\Pi}' \mathcal{P} \rho_\infty + \gamma \mathbb{1} \mathbb{1}^* \rho_\infty$ , and therefore,  $\gamma = 0$ , since  $\mathcal{P} \rho_\infty = 0$  and  $\mathbb{1}^* \rho_\infty \neq 0$ . Since  $\Pi'$  is non-negative, we obtain that for all  $y \in \mathcal{Y}_{\mathcal{P}}$ ,  $0 \leq \langle y, \Pi' y \rangle = \langle \mathcal{P} y, \widehat{\Pi}' \mathcal{P} y \rangle = \langle y, \widehat{\Pi}' y \rangle$ , which proves that  $\widehat{\Pi}'$  is non-negative. Therefore,  $\widehat{\Pi}' = \widehat{\Pi}$  and  $\Pi' = \mathcal{P}^* \widehat{\Pi} \mathcal{P} = \Pi$ . Finally,  $\Pi$  is the unique non-negative solution to (R1) such that  $\Pi \rho_\infty = 0$ .  $\square$

**Remark 4.5.** The Riccati equations (R1) and (R2) both provide the same feedback. Let  $\Pi$  be a solution to (R1), let  $\widehat{\Pi}$  be a solution to (R2), let  $\gamma \in \mathbb{R}$  be such that  $\Pi = \mathcal{P}^* \widehat{\Pi} \mathcal{P} + \gamma \mathbb{1} \mathbb{1}^*$ . Then, for all  $y \in \mathcal{Y}$ ,

$$-\mathcal{B}^* \Pi y = -\mathcal{B}^* (\mathcal{P}^* \widehat{\Pi} \mathcal{P} + \gamma \mathbb{1} \mathbb{1}^*) y = -\mathcal{B}^* \mathcal{P}^* \widehat{\Pi} (\mathcal{P} y) = -\widehat{\mathcal{B}}^* \widehat{\Pi} (\mathcal{P} y),$$

since  $\mathcal{B}^* \mathbb{1} = 0$  and  $\widehat{\mathcal{B}}^* = \mathcal{B}^* \mathcal{P}^*$ . The first and the last term of the above equation respectively correspond to the feedback controls associated with  $\Pi$  and  $\widehat{\Pi}$ .

#### 4.4. Local exponential stabilization of the nonlinear system

In this section, we study the effect of the static state feedback law  $u = -\widehat{\mathcal{B}}^* \widehat{\Pi} y_{\mathcal{P}}$  when applied to the nonlinear system

$$\dot{y}_{\mathcal{P}} = \widehat{\mathcal{A}} y_{\mathcal{P}} + u \widehat{\mathcal{N}} y_{\mathcal{P}} + \widehat{\mathcal{B}} u, \quad y_{\mathcal{P}}(0) = \mathcal{P} \rho_0.$$

Since we are interested in local exponential stabilization results, let us introduce the transformed state  $z_{\mathcal{P}} = e^{\delta t} y_{\mathcal{P}}$  where  $\delta$  is as in Section 4.1. We then obtain the transformed system

$$\dot{z}_{\mathcal{P}} = (\widehat{\mathcal{A}} + \delta I) z_{\mathcal{P}} + u \widehat{\mathcal{N}} z_{\mathcal{P}} + \widehat{\mathcal{B}} (e^{\delta t} u), \quad z_{\mathcal{P}}(0) = \mathcal{P} \rho_0.$$

As a consequence, our goal is a local stability result for the system

$$\dot{z}_{\mathcal{P}} = (\widehat{\mathcal{A}} + \delta I) z_{\mathcal{P}} - (\widehat{\mathcal{B}}^* \widehat{\Pi} z_{\mathcal{P}}) \widehat{\mathcal{N}}_{\delta} z_{\mathcal{P}} - \widehat{\mathcal{B}} (\widehat{\mathcal{B}}^* \widehat{\Pi} z_{\mathcal{P}}), \quad z_{\mathcal{P},\delta}(0) = \mathcal{P} \rho_0.$$

where  $\widehat{\mathcal{N}}_{\delta} = e^{-\delta t} \widehat{\mathcal{N}}$ . Using once more the notation  $\widehat{\mathcal{A}}_{\Pi} = \widehat{\mathcal{A}} + \delta I - \widehat{\mathcal{B}} \widehat{\mathcal{B}}^* \widehat{\Pi}$ , let us first consider the following nonhomogeneous system

$$\dot{z}_{\mathcal{P}} = \widehat{\mathcal{A}}_{\Pi} z_{\mathcal{P}} + f, \quad z_{\mathcal{P}}(0) = \mathcal{P} \rho_0. \quad (4.8)$$

Since  $\widehat{\mathcal{B}} \widehat{\mathcal{B}}^* \Pi \in \mathcal{L}(\mathcal{Y}_{\mathcal{P}})$  we conclude that

$$\mathcal{D}(\widehat{\mathcal{A}}_{\Pi}) = \mathcal{D}(\widehat{\mathcal{A}}) = \mathcal{D}(\mathcal{A}) \cap \mathcal{Y}_{\mathcal{P}}, \quad \mathcal{D}(\widehat{\mathcal{A}}_{\Pi}^*) = \mathcal{D}(\widehat{\mathcal{A}}^*) = \mathcal{D}(\mathcal{A}^*) \cap \mathcal{Y}_{\mathcal{P}}. \quad (4.9)$$

For the following calculus of interpolation spaces, assume that  $\lambda \in \mathbb{R}$  in the resolvent set of  $\mathcal{A}$  is chosen such that the fractional powers of  $\mathcal{A}_{\lambda} := (\lambda I - \mathcal{A})$  are well-defined. From ([31], Sect. 1.17.1), it follows that

$$\begin{aligned} [\mathcal{D}(\mathcal{A}_{\lambda}) \cap \mathcal{Y}_{\mathcal{P}}, \mathcal{Y} \cap \mathcal{Y}_{\mathcal{P}}]_{\alpha} &= [\mathcal{D}(\mathcal{A}_{\lambda}), \mathcal{Y}]_{\alpha} \cap \mathcal{Y}_{\mathcal{P}}, \\ [\mathcal{D}(\mathcal{A}_{\lambda}^*) \cap \mathcal{Y}_{\mathcal{P}}, \mathcal{Y} \cap \mathcal{Y}_{\mathcal{P}}]_{\alpha} &= [\mathcal{D}(\mathcal{A}_{\lambda}^*), \mathcal{Y}]_{\alpha} \cap \mathcal{Y}_{\mathcal{P}}. \end{aligned}$$

According to ([19], Appendix 3A), for  $\alpha = \frac{1}{2}$ , we can identify the above interpolation spaces as follows

$$[\mathcal{D}(\mathcal{A}_{\lambda}), \mathcal{Y}]_{\frac{1}{2}} \cap \mathcal{Y}_{\mathcal{P}} = H^1(\Omega) \cap \mathcal{Y}_{\mathcal{P}} = [\mathcal{D}(\mathcal{A}_{\lambda}^*), \mathcal{Y}]_{\frac{1}{2}} \cap \mathcal{Y}_{\mathcal{P}}.$$

Moreover, with ([21], Vol. I, Sect. 12) it holds that

$$\left[ [\mathcal{D}(\mathcal{A}_{\lambda}), \mathcal{Y}_{\mathcal{P}}]_{\frac{1}{2}}, [\mathcal{D}(\mathcal{A}_{\lambda}^*), \mathcal{Y}_{\mathcal{P}}]_{\frac{1}{2}}' \right]_{\frac{1}{2}} = \mathcal{Y}_{\mathcal{P}}.$$

For the following result, let us introduce the space

$$W_{\mathcal{P}}(Q_{\infty}) := L^2(0, \infty; H^1(\Omega) \cap \mathcal{Y}_{\mathcal{P}}) \cap H^1(0, \infty; [H^1(\Omega) \cap \mathcal{Y}_{\mathcal{P}}]')$$

endowed with the norm

$$\|y_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})} = \left( \int_0^{\infty} \|y_{\mathcal{P}}\|_{H^1(\Omega)}^2 + \|y'_{\mathcal{P}}\|_{[H^1(\Omega)]'}^2 dt \right)^{\frac{1}{2}}.$$

Based on known regularity results for analytic semigroups, we now have.

**Theorem 4.6.** *Let  $f \in L^2(0, \infty; [H^1(\Omega) \cap \mathcal{Y}_{\mathcal{P}}]')$  and  $\rho_0 \in L^2(\Omega)$  be given. Then there exists a unique mild solution  $z_{\mathcal{P}} \in W_{\mathcal{P}}(Q_{\infty})$  to (4.8) satisfying*

$$\|z_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})} \leq C (\|f\|_{L^2(0, \infty; [H^1(\Omega)]')} + \|\rho_0\|_{L^2(\Omega)}).$$

In particular,  $z_{\mathcal{P}} \in C_b([0, \infty); \mathcal{Y}_{\mathcal{P}})$ .

*Proof.* The result immediately follows from ([6], Chap. 3, Thm. 2.2 and [21], Vol. I, Thm. 4.2) together with the given characterizations of the involved interpolation spaces. □

The next lemma will be used in the following theorem.

**Lemma 4.7.** *Let  $y_{\mathcal{P}}, z_{\mathcal{P}} \in W_{\mathcal{P}}(Q_{\infty})$ . Then*

$$\begin{aligned} & \left\| (\widehat{\mathcal{B}}^* \widehat{\Pi} y_{\mathcal{P}}) \widehat{\mathcal{N}}_{\delta} y_{\mathcal{P}} - (\widehat{\mathcal{B}}^* \widehat{\Pi} z_{\mathcal{P}}) \widehat{\mathcal{N}}_{\delta} z_{\mathcal{P}} \right\|_{L^2(0, \infty; [H^1(\Omega)]')} \\ & \leq \widetilde{C} (\|y_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})} + \|z_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})}) \|y_{\mathcal{P}} - z_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})}. \end{aligned}$$

*Proof.* First note that we can extend the operator  $\widehat{\mathcal{N}}: H^1(\Omega) \cap \mathcal{Y}_{\mathcal{P}} \rightarrow \mathcal{Y}_{\mathcal{P}}$  to a continuous linear operator  $\widehat{\mathcal{N}}: \mathcal{Y}_{\mathcal{P}} \rightarrow [H^1(\Omega) \cap \mathcal{Y}_{\mathcal{P}}]'$ . For  $y_{\mathcal{P}}, z_{\mathcal{P}} \in W_{\mathcal{P}}(Q_{\infty})$ , we have

$$\begin{aligned} & \left\| (\widehat{\mathcal{B}}^* \widehat{\Pi} y_{\mathcal{P}}) \widehat{\mathcal{N}}_{\delta} y_{\mathcal{P}} - (\widehat{\mathcal{B}}^* \widehat{\Pi} z_{\mathcal{P}}) \widehat{\mathcal{N}}_{\delta} z_{\mathcal{P}} \right\|_{L^2(0, \infty; [H^1(\Omega)]')} \\ & \leq \widetilde{C} \left\| (\widehat{\mathcal{B}}^* \widehat{\Pi} (y_{\mathcal{P}} - z_{\mathcal{P}})) \widehat{\mathcal{N}} y_{\mathcal{P}} \right\|_{L^2(0, \infty; [H^1(\Omega)]')} + \widetilde{C} \left\| (\widehat{\mathcal{B}}^* \widehat{\Pi} z_{\mathcal{P}}) \widehat{\mathcal{N}} (y_{\mathcal{P}} - z_{\mathcal{P}}) \right\|_{L^2(0, \infty; [H^1(\Omega)]')}. \end{aligned}$$

For the first term, it holds that

$$\begin{aligned} & \left\| (\widehat{\mathcal{B}}^* \widehat{\Pi} (y_{\mathcal{P}} - z_{\mathcal{P}})) \widehat{\mathcal{N}} y_{\mathcal{P}} \right\|_{L^2(0, \infty; [H^1(\Omega)]')}^2 = \int_0^{\infty} |\langle \widehat{\Pi} \widehat{\mathcal{N}} \rho_{\infty}, y_{\mathcal{P}} - z_{\mathcal{P}} \rangle|^2 \|\widehat{\mathcal{N}} y_{\mathcal{P}}\|_{[H^1(\Omega)]'}^2 dt \\ & \leq \widetilde{C} \int_0^{\infty} \|y_{\mathcal{P}} - z_{\mathcal{P}}\|_{L^2(\Omega)}^2 \|y_{\mathcal{P}}\|_{L^2(\Omega)}^2 dt \leq \widetilde{C} \left( \sup_{t \in [0, \infty)} \|y_{\mathcal{P}}\|_{L^2(\Omega)} \right)^2 \|y_{\mathcal{P}} - z_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})}^2. \end{aligned}$$

With ([21], Vol. I, Thm. 4.2) this yields

$$\left\| (\widehat{\mathcal{B}}^* \widehat{\Pi} (y_{\mathcal{P}} - z_{\mathcal{P}})) \widehat{\mathcal{N}} y_{\mathcal{P}} \right\|_{L^2(0, \infty; [H^1(\Omega)]')} \leq \widetilde{C} \|y_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})} \|y_{\mathcal{P}} - z_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})}.$$

Similarly, we continue with

$$\begin{aligned} & \left\| (\widehat{\mathcal{B}}^* \widehat{\Pi} z_{\mathcal{P}}) \widehat{\mathcal{N}} (y_{\mathcal{P}} - z_{\mathcal{P}}) \right\|_{L^2(0, \infty; [H^1(\Omega)]')} = \int_0^{\infty} |\langle \widehat{\Pi} \widehat{\mathcal{B}}, z_{\mathcal{P}} \rangle|^2 \|\widehat{\mathcal{N}} (y_{\mathcal{P}} - z_{\mathcal{P}})\|_{[H^1(\Omega)]'}^2 dt \\ & \leq \widetilde{C} \int_0^{\infty} \|z_{\mathcal{P}}\|_{L^2(\Omega)}^2 \|y_{\mathcal{P}} - z_{\mathcal{P}}\|_{L^2(\Omega)}^2 dt \leq \widetilde{C} \left( \sup_{t \in [0, \infty)} \|z_{\mathcal{P}}\|_{L^2(\Omega)} \right)^2 \|y_{\mathcal{P}} - z_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})}^2. \end{aligned}$$

As before, this leads to

$$\left\| (\widehat{\mathcal{B}}^* \widehat{\Pi} z_{\mathcal{P}}) \widehat{\mathcal{N}} (y_{\mathcal{P}} - z_{\mathcal{P}}) \right\|_{L^2(0, \infty; [H^1(\Omega)]')} \leq \widetilde{C} \|z_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})} \|y_{\mathcal{P}} - z_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})}.$$

Combining both estimates shows the assertion. □

**Theorem 4.8.** *Let  $C$  and  $\widetilde{C}$  denote the constants from Theorem 4.6 and Lemma 4.7, respectively. If  $\|\rho_0\|_{L^2(\Omega)} \leq \frac{3}{16C^2\widetilde{C}}$ , then*

$$\dot{z}_{\mathcal{P}} = \widehat{\mathcal{A}}_{\Pi} z_{\mathcal{P}} - (\widehat{\mathcal{B}}^* \widehat{\Pi} z_{\mathcal{P}}) \widehat{\mathcal{N}}_{\delta} z_{\mathcal{P}}, \quad z_{\mathcal{P}}(0) = \mathcal{P} \rho_0,$$

admits a unique solution  $z_{\mathcal{P}} \in W_{\mathcal{P}}(Q_{\infty})$  satisfying

$$\|z_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})} \leq \frac{1}{4C\tilde{C}}.$$

*Proof.* We are going to show the assertion by a fixed point argument. For this purpose, consider the mapping  $\mathcal{F}: W_{\mathcal{P}}(Q_{\infty}) \rightarrow W_{\mathcal{P}}(Q_{\infty})$ ,  $w_{\mathcal{P}} \mapsto z_{\mathcal{P},w}$  defined by

$$\dot{z}_{\mathcal{P},w} = \widehat{\mathcal{A}}_{\Pi} z_{\mathcal{P},w} - \left(\widehat{\mathcal{B}}^* \widehat{\Pi} w_{\mathcal{P}}\right) \widehat{\mathcal{N}}_{\delta} w_{\mathcal{P}}, \quad z_{\mathcal{P},w}(0) = \mathcal{P}\rho_0.$$

Let  $w_{\mathcal{P}} \in W_{\mathcal{P}}(Q_{\infty})$  such that  $\|w_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})} \leq \frac{1}{4C\tilde{C}}$ . Lemma 4.7 then implies that

$$\left\| \left(\widehat{\mathcal{B}}^* \widehat{\Pi} w_{\mathcal{P}}\right) \widehat{\mathcal{N}}_{\delta} w_{\mathcal{P}} \right\|_{L^2(0,\infty;[H^1(\Omega)]')} \leq \frac{1}{16C^2\tilde{C}}.$$

With Theorem 4.6 we conclude that the corresponding solution satisfies

$$\|z_{\mathcal{P},w}\|_{W_{\mathcal{P}}(Q_{\infty})} \leq C \left( \frac{1}{16C^2\tilde{C}} + \frac{3}{16C^2\tilde{C}} \right) = \frac{1}{4C\tilde{C}}.$$

Similarly, for  $w_{\mathcal{P},1}, w_{\mathcal{P},2} \in W_{\mathcal{P}}(Q_{\infty})$  with  $\|w_{\mathcal{P},i}\|_{W_{\mathcal{P}}(Q_{\infty})} \leq \frac{1}{4C\tilde{C}}, i = 1, 2$ , the associated solutions  $z_{\mathcal{P},w_1}$  and  $z_{\mathcal{P},w_2}$  fulfill

$$\begin{aligned} \dot{z}_{\mathcal{P},w_1} - \dot{z}_{\mathcal{P},w_2} &= \widehat{\mathcal{A}}_{\Pi}(z_{\mathcal{P},w_1} - z_{\mathcal{P},w_2}) + (\widehat{\mathcal{B}}^* \widehat{\Pi} w_{\mathcal{P},2}) \widehat{\mathcal{N}}_{\delta} w_{\mathcal{P},2} - (\widehat{\mathcal{B}}^* \widehat{\Pi} w_{\mathcal{P},1}) \widehat{\mathcal{N}}_{\delta} w_{\mathcal{P},1} \\ z_{\mathcal{P},w_1}(0) - z_{\mathcal{P},w_2}(0) &= 0. \end{aligned}$$

Hence, Theorem 4.6 yields

$$\|z_{\mathcal{P},w_1} - z_{\mathcal{P},w_2}\|_{W_{\mathcal{P}}(Q_{\infty})} \leq C \left\| (\widehat{\mathcal{B}}^* \widehat{\Pi} w_{\mathcal{P},2}) \widehat{\mathcal{N}}_{\delta} w_{\mathcal{P},2} - (\widehat{\mathcal{B}}^* \widehat{\Pi} w_{\mathcal{P},1}) \widehat{\mathcal{N}}_{\delta} w_{\mathcal{P},1} \right\|_{L^2(0,\infty;[H^1(\Omega)]')}.$$

Moreover, with Lemma 4.6, we obtain that

$$\|z_{\mathcal{P},w_1} - z_{\mathcal{P},w_2}\|_{W_{\mathcal{P}}(Q_{\infty})} \leq 2C\tilde{C} \left( \frac{1}{4C\tilde{C}} \right)^2 = \frac{1}{2} \frac{1}{4C\tilde{C}}.$$

In other words, the mapping  $\mathcal{F}$  is a contraction in the set

$$\left\{ z_{\mathcal{P}} \in W_{\mathcal{P}}(Q_{\infty}) \left| \|z_{\mathcal{P}}\|_{W_{\mathcal{P}}(Q_{\infty})} \leq \frac{1}{4C\tilde{C}} \right. \right\}$$

and the statement is shown. □

As a consequence of Theorem 4.8, we have that  $e^{\delta t} y_{\mathcal{P}} \in W_{\mathcal{P}}(Q_{\infty})$  implying that there exists a constant  $C$  such that  $\|y_{\mathcal{P}}\|_{L^2(\Omega)} \leq C e^{-\delta t} \|\rho_0\|_{L^2(\Omega)}$ .

### 5. A LYAPUNOV BASED FEEDBACK LAW

As an alternative to the Riccati based approach, in this section, we propose a feedback law that allows to construct a *global* Lyapunov function for the nonlinear closed loop system. The idea is inspired by the observations found in [4] for hyperbolic systems.

With the previously introduced notation, assume that  $(\lambda_2, \psi_2)$  denotes the eigenpair of  $\widehat{\mathcal{A}}$  associated to the first nonzero eigenvalue. Hence,  $\lambda_2$  determines the exponential decay rate of the uncontrolled systems. Instead of using (4.5), let us determine the control shape function  $\alpha$  as a solution to the elliptic equation

$$\begin{aligned} \nabla \cdot (\rho_{\infty} \nabla \alpha) &= \psi_2 & \text{in } & \Omega, \\ (\rho_{\infty} \nabla \alpha) \cdot \mathbf{n} &= 0 & \text{on } & \Gamma. \end{aligned} \tag{5.1}$$

As a consequence, this choice of  $\alpha$  yields  $\widehat{\mathcal{B}} = \widehat{\mathcal{N}}\rho_\infty = \psi_2$ . Let further  $\mu > 0$  be chosen such that

$$\langle (\mu I - \widehat{\mathcal{A}})y_{\mathcal{P}}, y_{\mathcal{P}} \rangle_{L^2(\Omega)} \geq \langle y_{\mathcal{P}}, y_{\mathcal{P}} \rangle_{H^1(\Omega)}, \quad \text{for all } y_{\mathcal{P}} \in \mathcal{D}(\widehat{\mathcal{A}}). \tag{5.2}$$

Since  $\widehat{\mathcal{A}}$  generates an exponentially stable semigroup, it is well-known ([10], Thm. 4.1.23) that there exists a unique self-adjoint nonnegative solution  $\Upsilon$  to the Lyapunov equation for  $y_{\mathcal{P}}, z_{\mathcal{P}} \in \mathcal{D}(\widehat{\mathcal{A}})$ :

$$\langle \Upsilon y_{\mathcal{P}}, \widehat{\mathcal{A}}z_{\mathcal{P}} \rangle + \langle \widehat{\mathcal{A}}y_{\mathcal{P}}, \Upsilon z_{\mathcal{P}} \rangle = -2\mu \langle y_{\mathcal{P}}, z_{\mathcal{P}} \rangle. \tag{5.3}$$

We then obtain the following result.

**Theorem 5.1.** *Let  $\mu$  and  $\Upsilon$  be as in (5.2) and (5.3), respectively. Consider the system*

$$\dot{y}_{\mathcal{P}} = \widehat{\mathcal{A}}y_{\mathcal{P}} + u\widehat{\mathcal{N}}y_{\mathcal{P}} + \widehat{\mathcal{B}}u, \quad y_{\mathcal{P}}(0) = \mathcal{P}\rho_0. \tag{5.4}$$

where the control  $u$  is defined by the feedback law

$$u = -\langle \widehat{\mathcal{B}} + \widehat{\mathcal{N}}y_{\mathcal{P}}, \Upsilon y_{\mathcal{P}} + y_{\mathcal{P}} \rangle.$$

Then the function  $V(y_{\mathcal{P}}) := \langle y_{\mathcal{P}}, \Upsilon y_{\mathcal{P}} + y_{\mathcal{P}} \rangle$  is a global Lyapunov function for (5.4).

*Proof.* Since  $\Upsilon$  is self-adjoint and nonnegative, it obviously holds that  $V(y_{\mathcal{P}}) \geq \|y_{\mathcal{P}}\|^2$ . Moreover, we obtain that

$$\begin{aligned} \frac{d}{dt}V(y_{\mathcal{P}}) &= \langle \widehat{\mathcal{A}}y_{\mathcal{P}}, \Upsilon y_{\mathcal{P}} + y_{\mathcal{P}} \rangle + \langle y_{\mathcal{P}}, \Upsilon \widehat{\mathcal{A}}y_{\mathcal{P}} + \widehat{\mathcal{A}}y_{\mathcal{P}} \rangle \\ &\quad - \langle \widehat{\mathcal{B}} + \widehat{\mathcal{N}}y_{\mathcal{P}}, \Upsilon y_{\mathcal{P}} + y_{\mathcal{P}} \rangle \langle \widehat{\mathcal{B}} + \widehat{\mathcal{N}}y_{\mathcal{P}}, \Upsilon y_{\mathcal{P}} + y_{\mathcal{P}} \rangle \\ &\quad - \langle \widehat{\mathcal{B}} + \widehat{\mathcal{N}}y_{\mathcal{P}}, \Upsilon y_{\mathcal{P}} + y_{\mathcal{P}} \rangle \langle y_{\mathcal{P}}, \Upsilon(\widehat{\mathcal{N}}y_{\mathcal{P}} + \widehat{\mathcal{B}}) + (\widehat{\mathcal{N}}y_{\mathcal{P}} + \widehat{\mathcal{B}}) \rangle \\ &= -2\mu \langle y_{\mathcal{P}}, y_{\mathcal{P}} \rangle + \langle \widehat{\mathcal{A}}y_{\mathcal{P}}, y_{\mathcal{P}} \rangle + \langle y_{\mathcal{P}}, \widehat{\mathcal{A}}y_{\mathcal{P}} \rangle \\ &\quad - 2\langle \widehat{\mathcal{B}} + \widehat{\mathcal{N}}y_{\mathcal{P}}, \Upsilon y_{\mathcal{P}} + y_{\mathcal{P}} \rangle^2 \\ &\leq 2\langle (\widehat{\mathcal{A}} - \mu I)y_{\mathcal{P}}, y_{\mathcal{P}} \rangle \leq -2\langle y_{\mathcal{P}}, y_{\mathcal{P}} \rangle_{H^1(\Omega)} \end{aligned}$$

which shows the assertion. □

In addition to the previous result, the feedback law locally increases the exponential decay rate.

**Theorem 5.2.** *Let  $\lambda_i, i = 2, 3, \dots$  denote the eigenvalues of the operator  $\widehat{\mathcal{A}}$ . Assume that*

$$\widetilde{\lambda}_2 := \lambda_2 - \|\psi_2\|^2 + \frac{\mu}{\lambda_2} \|\psi_2\|^2 \neq \lambda_j, \quad j = 3, \dots$$

Then for the spectrum of the linearized closed loop operator it holds that

$$\sigma(\widehat{\mathcal{A}} - \widehat{\mathcal{B}}\widehat{\mathcal{B}}^*\Upsilon - \widehat{\mathcal{B}}\widehat{\mathcal{B}}^*) = \{\widetilde{\lambda}_2\} \cup \{\lambda_j\}, \quad j \geq 3.$$

*Proof.* Due to (5.3), we find that

$$\langle \Upsilon \psi_2, \widehat{\mathcal{A}}\psi_2 \rangle + \langle \widehat{\mathcal{A}}\psi_2, \Upsilon \psi_2 \rangle = -2\mu \|\psi_2\|^2.$$

Since  $\psi_2$  is an eigenfunction of  $\widehat{\mathcal{A}}$ , this implies that

$$\langle \Upsilon \psi_2, \psi_2 \rangle = -\frac{\mu}{\lambda_2} \|\psi_2\|^2.$$



Further, from our choice of  $\alpha$ , we already know that  $\widehat{\mathcal{B}} = \psi_2$ . Hence, it follows that

$$(\widehat{\mathcal{A}} - \widehat{\mathcal{B}}\widehat{\mathcal{B}}^*\Upsilon - \widehat{\mathcal{B}}\widehat{\mathcal{B}}^*)\psi_2 = \lambda_2\psi_2 - \langle \Upsilon\psi_2, \psi_2 \rangle \psi_2 - \|\psi_2\|^2\psi_2$$

which shows the first part. For  $\beta_j := \frac{\langle \Upsilon\psi_j + \psi_j, \psi_2 \rangle}{\lambda_2 - \lambda_j}, j = 3, \dots$  we further arrive at

$$\begin{aligned} (\widehat{\mathcal{A}} - \widehat{\mathcal{B}}\widehat{\mathcal{B}}^*\Upsilon - \widehat{\mathcal{B}}\widehat{\mathcal{B}}^*)(\psi_j + \beta_j\psi_2) &= \beta_j\widetilde{\lambda}_2\psi_2 + \widehat{\mathcal{A}}\psi_j - \widehat{\mathcal{B}}\widehat{\mathcal{B}}^*\Upsilon\psi_j - \widehat{\mathcal{B}}\widehat{\mathcal{B}}^*\psi_j \\ &= \beta_j\widetilde{\lambda}_2\psi_2 + \lambda_j\psi_j - \psi_2\langle \Upsilon\psi_j, \psi_2 \rangle - \psi_2\langle \psi_j, \psi_2 \rangle = \lambda_j(\psi_j + \beta_j\psi_2). \end{aligned}$$

This shows the claim. □

**Remark 5.3.** Let us emphasize that the feedback law is particularly useful in cases where  $\lambda_2$  is close to the imaginary axis and there is a gap between  $\lambda_2$  and  $\lambda_3$ . Indeed, for  $\lambda_2 \rightarrow 0$ , the term  $\frac{\mu}{\lambda_2} \rightarrow -\infty$ , such that the modified eigenvalue  $\widetilde{\lambda}_2$  is moved far away from the imaginary axis.

### 6. NUMERICAL STUDY – A TWO DIMENSIONAL DOUBLE WELL POTENTIAL

As a numerical example, we consider

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \nu \Delta \rho + \nabla \cdot (\rho \nabla G) + u \nabla \cdot (\rho \nabla \alpha) && \text{in } \Omega \times (0, \infty), \\ 0 &= (\nu \nabla \rho + \rho \nabla V) \cdot \mathbf{n} && \text{on } \Gamma \times (0, \infty), \\ \rho(x, 0) &= \rho_0(x) && \text{in } \Omega, \end{aligned} \tag{6.1}$$

on  $\Omega = (-1.5, 1.5) \times (-1, 1) \subset \mathbb{R}^2$ , with  $\nu = 1$  and a two dimensional double well potential of the form

$$G(x) = 3(x_1^2 - 1)^2 + 6x_2^2.$$

For the spatial semidiscretization, a finite difference scheme with  $k = n_{x_1} \cdot n_{x_2} = 96 \cdot 64 = 6144$  degrees of freedom was implemented. The discretization  $A \in \mathbb{R}^{k \times k}$  of the operator  $\mathcal{A}$  defined as in (3.2) was obtained by first discretizing the operator  $\mathcal{A}^*$  as given by (3.3) and then taking the transpose of the resulting matrix. The reason for this indirect approach was that the discretization of  $\mathcal{A}^*$  only required the incorporation of “standard” Neumann boundary conditions rather than the mixed boundary conditions arising for  $\mathcal{A}$ . Due to the convective terms included in  $\mathcal{A}$  and  $\mathcal{A}^*$ , a first order upwind scheme was utilized. Let us emphasize that even for the value  $\nu = 1$ , this turned out to be essential for the accuracy of the discretization. We also mention the possibility of using more advanced discretization schemes that have been proposed in the context of the Fokker–Planck equation, see, e.g., [2, 8]. However, the finite difference scheme lead to accurate approximations of the stationary distribution and the preservation of probability was ensured up to machine precision in all our numerical results. Figure 1 now shows the discretization of the double well potential as well as the corresponding (spatially discrete) stationary distribution  $\rho_\infty^k$ . For both the Riccati-based and the Lyapunov-based control strategy, the discrete control operators  $N$  and  $B = N\rho_\infty^k$  were derived based on the solutions  $\alpha(x)$  to (4.5) and (5.1). To be more precise, first, the involved elliptic equations were also discretized by a finite difference scheme which, due to the Neumann boundary conditions, lead to matrices  $C$  with a zero eigenvalue. The individual spatially discrete shape functions  $\alpha^k$  were obtained by utilizing the Moore-Penrose pseudoinverse of the  $C$  matrices. Finally, with the resulting  $\alpha^k$ , the matrices  $N$  were generated by the discretization of the operator  $\mathcal{N}$  defined in (3.2). For the Riccati-based approach, we incorporated the eigenfunctions to the first three nonzero eigenvalues into (4.5), i.e., we set  $d = 4$ . Since varying the value  $d$  lead to qualitatively similar behavior, we only report on the results for the special case  $d = 4$ . Due to the Hautus criterion, it was thus possible to solve the associated Riccati equation with  $\delta \approx 12.26$ . The corresponding control shape functions for the Riccati-based (left) and the Lyapunov-based (right) approach are given in Figure 2. In order to underline the benefit of using a “specific”  $\alpha$  rather than an “arbitrary” one, we also report on some results we obtained by rotating the control shape function  $\alpha$  (see Fig. 2 center) while still using a Riccati-based feedback law obtained from the linearized system.

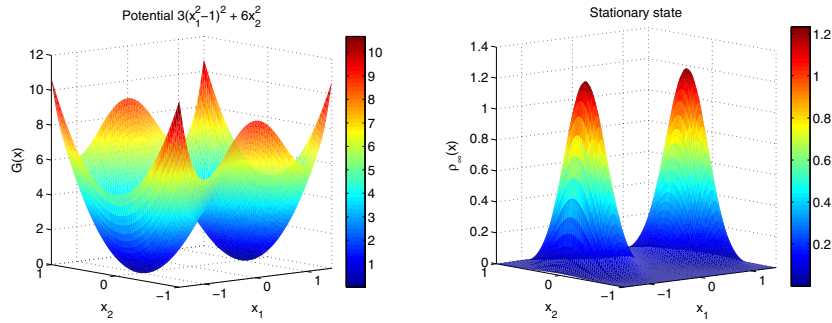


FIGURE 1. Confining double well potential (*left*) and associated stationary state (*right*).

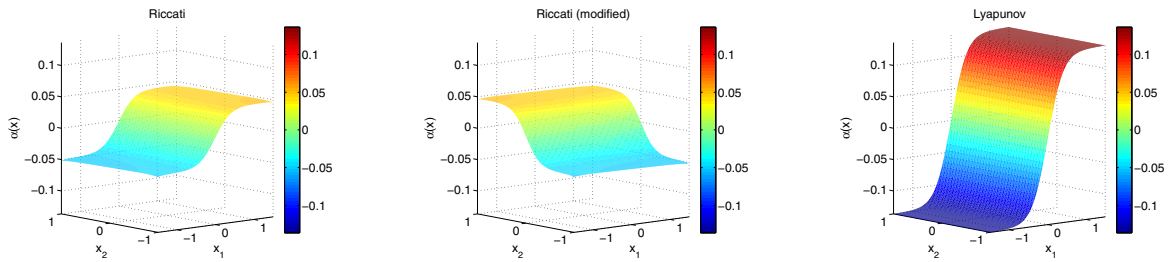


FIGURE 2. Control shape functions for different approaches.

All simulations were generated on an Intel<sup>®</sup>Xeon(R) CPU E31270 @ 3.40 GHz × 8, 16 GB RAM, Ubuntu Linux 14.04, MATLAB Version 8.0.0.783 (R2012b) 64-bit (glnxa64). The solutions of the ODE systems were always obtained by the MATLAB routine `ode23`. For solving the Riccati and Lyapunov equations, we used the MATLAB routine `care` and `lyap`, respectively, and the technique presented below.

### 6.1. Solving the Riccati equation

Based on the discretization scheme described above, let us at this point assume that  $A \in \mathbb{R}^{k \times k}$ ,  $B \in \mathbb{R}^{k \times 1}$ ,  $M \in \mathbb{R}^{k \times k}$ ,  $\rho_\infty \in \mathbb{R}^k$  are given and satisfy:

$$A\rho_\infty = 0, \quad A^\top \mathbb{1} = 0, \quad B^\top \mathbb{1} = 0, \quad M^\top = M, \quad \langle \mathbb{1}, \rho_\infty \rangle = 1,$$

where  $\mathbb{1} = h_{x_1} \cdot h_{x_2} (1, \dots, 1)^\top$  and  $h_x, h_y$  denote the mesh size. We denote by  $P$  the projection on  $\mathbb{1}^\perp$  along  $\mathbb{R}\rho_\infty$ :  $P = I_k - \rho_\infty \mathbb{1}^\top$ . We denote by  $(e_i)_{i=1, \dots, k}$  the vectors of the canonical basis. We aim at solving the following discretized Riccati equation:

$$(A^\top + \delta P^\top)\Pi + \Pi(A + \delta P) - \Pi B B^\top \Pi + P^\top M P = 0, \quad \Pi \rho_\infty = 0, \quad \Pi^\top = \Pi. \tag{6.2}$$

Let  $R \in \mathbb{R}^{k \times k}$  be a regular matrix satisfying:

$$R e_k = \rho_\infty \quad \text{and} \quad R^\top \mathbb{1} = e_k.$$

Note that the condition  $R^\top \mathbb{1} = e_k$  is equivalent to:  $\forall i = 1, \dots, k - 1, R e_i \in \mathbb{1}^\perp$ . An example of matrix  $R$  is given by:

$$R = \begin{pmatrix} 1 & & & \rho_{\infty,1} \\ & \ddots & & \vdots \\ & & 1 & \rho_{\infty,k-1} \\ -1 & \dots & -1 & \rho_{\infty,k} \end{pmatrix}.$$

Note that:

$$R^{-1} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \vdots \\ & & 1 & 0 \\ 1 & \dots & 1 & 1 \end{pmatrix} - \begin{pmatrix} \rho_{\infty,1} & \dots & \rho_{\infty,1} \\ \vdots & \vdots & \vdots \\ \rho_{\infty,k-1} & \dots & \rho_{\infty,k-1} \\ 0 & \dots & 0 \end{pmatrix}.$$

We also introduce:  $Q = \begin{pmatrix} I_{k-1} \\ 0 \end{pmatrix}$ . Consider the reduced and discretized Riccati equation (in  $\mathbb{R}^{(k-1) \times (k-1)}$ ):

$$(\widehat{A}^\top + \delta I_{k-1})\widehat{\Pi} + \widehat{\Pi}(\widehat{A} + \delta I_{k-1}) - \widehat{\Pi}\widehat{B}\widehat{B}^\top\widehat{\Pi} + \widehat{M} = 0, \quad \widehat{\Pi}^\top = \widehat{\Pi}, \tag{6.3}$$

where  $\widehat{A} = Q^\top R^{-1}ARQ$ ,  $\widehat{B} = Q^\top R^{-1}B$ ,  $\widehat{M} = Q^\top R^\top P^\top MPRQ$ .

**Lemma 6.1.** *Let  $\Pi \in \mathbb{R}^{k \times k}$ . The matrix  $\Pi$  is a solution to (6.2) if and only if there exists a solution  $\widehat{\Pi}$  to (6.3) such that  $\Pi = R^{-\top} \begin{pmatrix} \widehat{\Pi} & 0 \\ 0 & 0 \end{pmatrix} R^{-1}$ .*

*Proof.* The proof is similar to the proof of Lemma 4.3. Observe that  $\Pi$  is a solution to (6.2) if and only if  $\widetilde{\Pi} = R^\top \Pi R$  is a solution to

$$(\widetilde{A}^\top + \delta \widetilde{P}^\top)\widetilde{\Pi} + \widetilde{\Pi}(\widetilde{A} + \delta \widetilde{P}) - \widetilde{\Pi}(\widetilde{B}\widetilde{B}^\top)\widetilde{\Pi} + \widetilde{M} = 0, \quad \widetilde{\Pi}e_N = 0, \quad \widetilde{\Pi}^\top = \widetilde{\Pi} \tag{6.4}$$

where:  $\widetilde{A} = R^{-1}AR$ ,  $\widetilde{P} = R^{-1}PR$ ,  $\widetilde{B} = R^{-1}B$ ,  $\widetilde{M} = R^\top P^\top MPR$ . One can easily check that the last row and the last column of the following matrices are null:  $\widetilde{A}$ ,  $\widetilde{P}$ ,  $\widetilde{B}\widetilde{B}^\top$ ,  $\widetilde{M}$ . Moreover, the upper left block of  $\widetilde{P}$  is  $I_{k-1}$ . The equivalence follows directly from a block decomposition of equation (6.4).  $\square$

**Remark 6.2.** Let us emphasize that computing the solution  $\widehat{\Pi}$  to (6.3) is a challenging task already in the case when  $\Omega \subset \mathbb{R}^n$  with  $n = 2, 3$ , respectively, in particular because the matrices defining the reduced Riccati equation (6.3) are dense. On the other hand, according to Lemma 3.1 the only accumulation point of the spectrum of  $\mathcal{A}$  is  $-\infty$ . Thus, as a perspective for future developments geared at considering control of the Fokker–Planck equation in higher dimensions, it is of interest to only  $\delta$ -stabilize the part of the spectrum that is closest to the imaginary axis. This way, the resolution of a Riccati equation of large dimension can be avoided at almost no loss of performance. The idea goes back (at least) to [32] and is also studied in [25] and the references therein. A detailed discussion together with an implementation tailored to the special structure of the Fokker–Planck equation is currently being investigated. As an alternative way for reducing the complexity we also mention specific model reduction approaches as considered in [13, 14].

### 6.2. A random initial state

The first test case is concerned with the evolution of the uncontrolled and controlled systems for a random initial state  $\rho_0^k$  ( $\text{rand}(k)$ ). The temporal evolution of the deviation of the state  $\rho(t)$  from the stationary distribution  $\rho_\infty^k$  with respect to the  $L^2(\Omega)$ -norm is shown in Figure 3. In addition to the dynamics of the systems, we also visualized the exponential decay rate  $\delta$  that one would expect from solving the Riccati equation discussed in Section 6.1. Some comments are in order. It can be seen that in the beginning, the uncontrolled system approaches the stationary distribution as fast as the controlled systems. After some time, however, the convergence rate becomes significantly slower. For the controlled solutions, let us point out that there is almost no visible difference between the Lyapunov-based approach and the Riccati-based approach. On the other hand, with the rotated control shape function  $\alpha$ , the performance is clearly worse. In fact, in this case, the controlled dynamics converge slower than for the uncontrolled case. This phenomenon is understood better when considering snapshots of the solution for different time steps. In Figure 4 the results are shown for  $t = 0.01$  and  $t = 0.15$ . Except for the case of the rotated  $\alpha$ , all solutions have approximately approached the stationary state

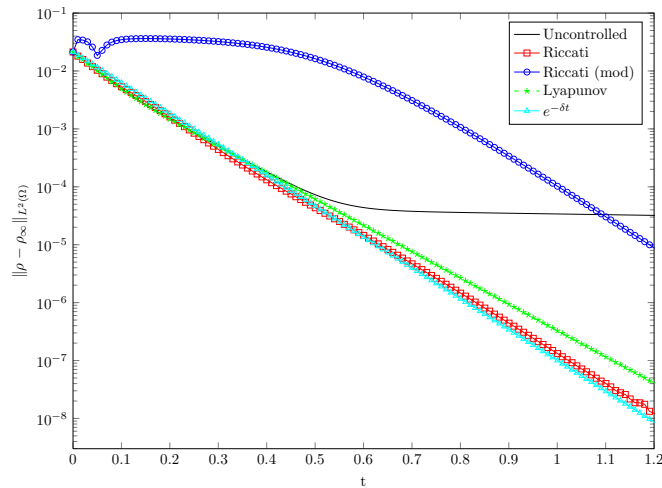


FIGURE 3. Comparison of  $L^2(\Omega)$ -norm evolution.

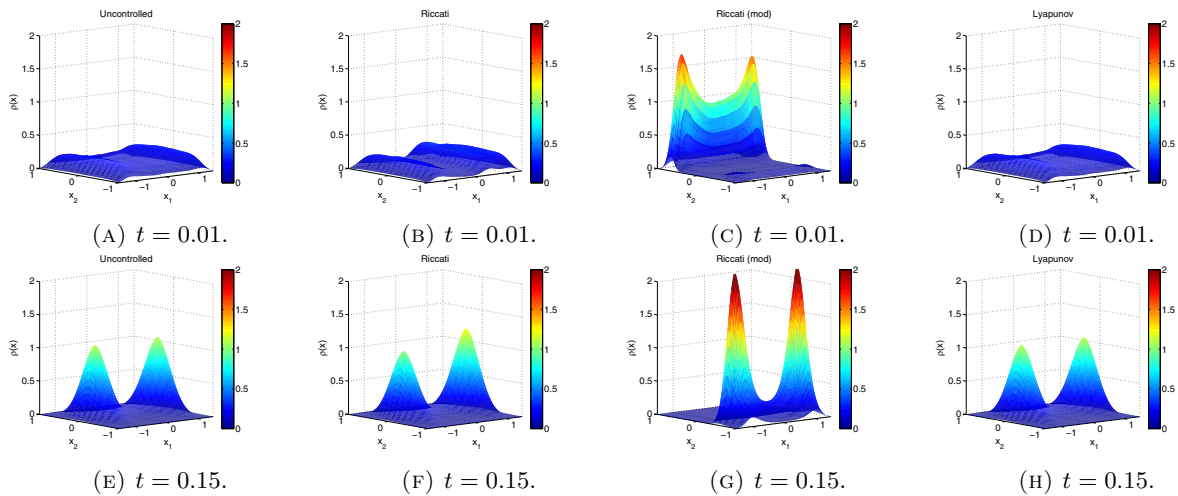


FIGURE 4. Temporal evolution of the state  $\rho$ .

at time  $t = 0.15$  already. Taking into account the shape of the stationary distribution, the shape of  $\alpha$  for the Riccati-based and the Lyapunov-based approach are intuitive. In both cases, the control allows to lower the potential around the left well and to raise it around the right well. Obviously, since  $u$  is allowed to be positive as well as negative, this effect can be reversed such that the right well is given preference. On the other hand, when the shape function is subject to a rotation as done in the experiments, both wells are equally important and no direct transition between them is possible. This is exactly what happens in the simulation. The control law pushes the particle first to the upper boundary before it is moved back to the lower boundary, see Figure 4. Figure 5 visualizes the influence of the different control strategies on the potential  $G(x)$ . Again, the effect of the modified Riccati approach is the lowering of the potential on the bottom and top boundary instead of the left and right boundary, respectively. It is further worthwhile to note that the Lyapunov-based feedback law influences the potential only moderately.

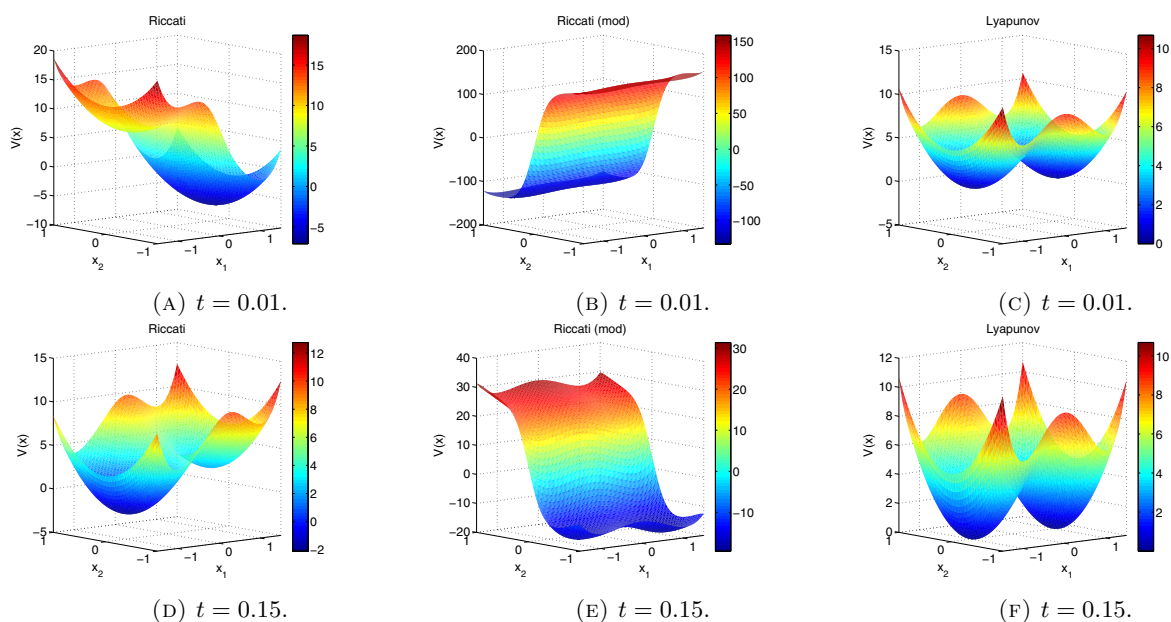


FIGURE 5. Temporal evolution of the potential  $V(x)$ .

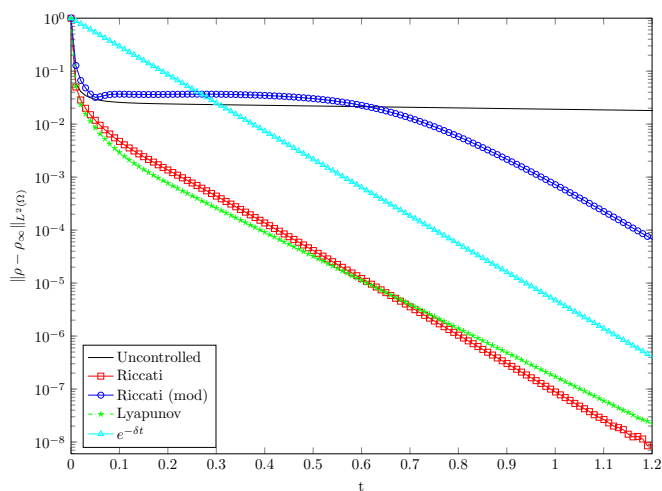
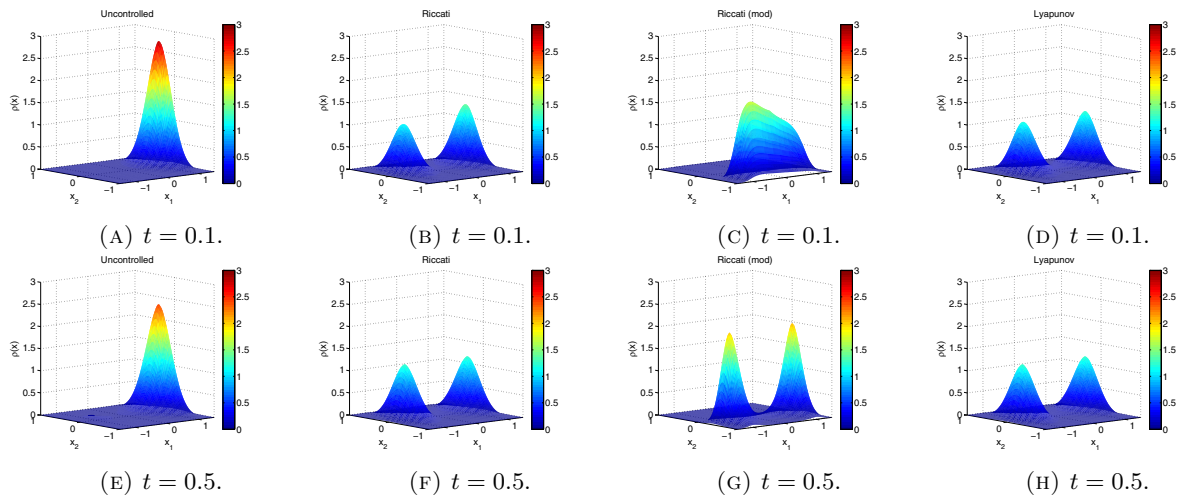
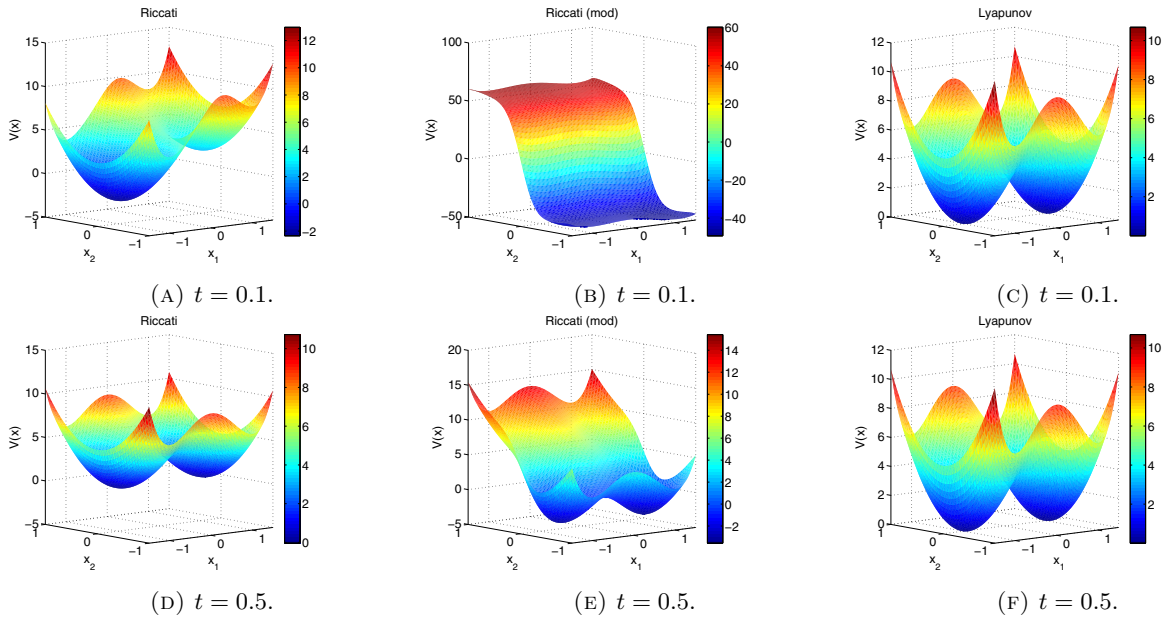


FIGURE 6. Comparison of  $L^2(\Omega)$ -norm evolution.

### 6.3. The particle located in one well

For the second test case, we assume the particle is initially located in the center of the right potential well, *i.e.*, the initial state reflects a numerical point mass at  $x_1 = 1, x_2 = 0$ . As is shown in Figure 6, in this case the convergence rate of the uncontrolled system is undesirably slow. We already mentioned that this is mainly

FIGURE 7. Temporal evolution of the state  $\rho$ .FIGURE 8. Temporal evolution of the potential  $V(x)$ .

reflected by the fact that the particle has to overcome the “energy barrier” between the potential wells. Here, the feedback laws act by lowering this barrier, hence allowing the particle to “jump” into the left potential well. As in the previous case, Figures 7 and 8 show the temporal evolution of the state of the systems as well as the influence on the potential. Again, the modified Riccati approach acts on the dynamics by first attracting the particle at the lower boundary from where it is slowly moved to the center of the wells.

*Acknowledgements.* This work was supported in part by the ERC advanced grant 668998 (OCLOC) under the EU’s H2020 research program.

## REFERENCES

- [1] R. Adams, Sobolev Spaces. New York Academic Press (1975).
- [2] M. Annunziato and A. Borzi, A Fokker–Planck control framework for multidimensional stochastic processes. *J. Comput. Appl. Math.* (2013) 487–507.
- [3] M. Badra and T. Takahashi, Feedback stabilization of a fluid-rigid body interaction system. *Adv. Differ. Equ.* **19** (2014) 1137–1184.
- [4] J. Ball and M. Slemrod, Feedback stabilization of distributed semilinear control systems. *Appl. Math. Optimiz.* **5** (1979) 169–179.
- [5] V. Barbu, I. Lasiecka and R. Triggiani, Tangential boundary stabilization of Navier–Stokes equations. *Memoirs Amer. Math. Soc.* **181** (2006) 1–128.
- [6] A. Bensoussan, G. Da Prato, M. Delfour and S. Mitter, Representation and Control of Infinite Dimensional Systems. Birkhäuser Boston Basel Berlin (2007).
- [7] V.I. Bogachev, G. Da Prato and M. Röckner, Fokker–Planck equations and maximal dissipativity for Kolmogorov operators with time dependent singular drifts in Hilbert spaces. *J. Functional Anal.* **256** (2009) 1269–1298.
- [8] J. Chang and G. Cooper, A practical scheme for Fokker–Planck equations. *J. Comput. Phys.* **6** (1970) 1–16.
- [9] M. Chipot, Elements of Nonlinear Analysis. Birkhäuser (2000).
- [10] R. Curtain and H. Zwart, An Introduction to Infinite-Dimensional Linear Systems Theory. Springer Verlag (2005).
- [11] C.W. Gardiner, Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences, vol. 13 of Springer Series in Synergetics. Springer Verlag, Berlin, 3rd edition (2004).
- [12] J.J. Gorman, A. Balijepalli and T.W. LeBrun, Feedback Control of MEMS to Atoms. Springer US, Boston, MA, ch. Feedback Control of Optically Trapped Particles (2012) 141–177.
- [13] C. Hartmann, Balanced model reduction of partially-observed Langevin equations: an averaging principle. *Math. Comput. Model. Dynamical Syst.* **17** (2011) 463–490.
- [14] C. Hartmann, B. Schäfer–Bund and A. Thöns-Zueva, Balanced averaging of bilinear systems with applications to stochastic control. *SIAM J. Control Optimiz.* **51** (2013) 2356–2378.
- [15] W. Huang, M. Ji, Z. Liu and Y. Yi, Steady states of Fokker–Planck equations: I. existence. *J. Dynamics Differ. Equ.* **27** (2015) 721–742.
- [16] P. H. Jones, M.M. Onofrio and G. Volpe, Optical Tweezers: Principles and Applications. Cambridge University Press (2015).
- [17] T. Kato, Perturbation Theory for Linear Operators. Springer Verlag, Berlin/Heidelberg, Germany (1980).
- [18] A. Khapalov, Controllability of partial differential equations governed by multiplicative controls, vol. 1995 of *Lect. Notes Math.*, Springer Verlag, Berlin (2010).
- [19] I. Lasiecka and R. Triggiani, Control Theory for Partial Differential Equations: Volume 1, Abstract Parabolic Systems: Continuous and Approximation Theories, vol. 1, Cambridge University Press (2000).
- [20] C. Le Bris and P.-L. Lions, Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients. *Commun. Partial Differ. Equ.* **33** (2008) 1272–1317.
- [21] J. Lions and E. Magenes, Non-homogeneous Boundary Value Problems and Applications. Vol. I/II. *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen*, Springer Verlag, Berlin (1972).
- [22] P.-L. Lions and A.-S. Sznitman, Stochastic differential equations with reflecting boundary conditions. *Commun. Pure Appl. Math.* **37** (1984) 511–537.
- [23] B. J. Matkowsky and Z. Schuss, Eigenvalues of the Fokker–Planck operator and the approach to equilibrium for diffusions in potential fields. *SIAM J. Appl. Math.* **40** (1981) 242–254.
- [24] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations. Springer New York (1983).
- [25] J. Raymond and L. Thevenet, Boundary feedback stabilization of the two dimensional Navier–Stokes equations with finite dimensional controllers. *Discrete Contin. Dyn. Syst.* **27** (2010) 1159–1187.
- [26] J.-P. Raymond, Feedback boundary stabilization of the two-dimensional Navier–Stokes equations. *SIAM J. Control Optimiz.* **45** (2006) 790–828.
- [27] R. Risken, The Fokker–Planck Equation: Methods of Solutions and Applications. Springer Verlag Berlin (1996).
- [28] H. Tanabe, Equations of evolution, vol. 6 of *Monographs and Studies in Mathematics*, Pitman (Advanced Publishing Program), Boston, Mass.-London (1979). Translated from the Japanese by N. Mugibayashi and H. Haneda.
- [29] L. Tartar, An Introduction to Sobolev Spaces and Interpolation Spaces. Springer Berlin Heidelberg, Berlin, Heidelberg (2007) 99–101.
- [30] L. Thevenet, J.-M. Buchot and J.-P. Raymond, Nonlinear feedback stabilization of a two-dimensional Burgers equation. *ESAIM: COCV* **16** (2010) 929–955.
- [31] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*. North-Holland Publishin Company (1978).
- [32] R. Triggiani, On the stabilizability problem in Banach space, *J. Math. Anal. Appl.* **52** (1975) 383–403.
- [33] G. Troianiello, Elliptic Differential Equations and Obstacle Problems. *The University Series in Mathematics*, Plenum Press, New York (1987).