INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus
Mécanique

Karl Kunisch and Daniel Walter

Optimal feedback control of dynamical systems via value-function approximation

Published online: 20 July 2023

https://doi.org/10.5802/crmeca.199

Part of Special Issue: The scientific legacy of Roland Glowinski
Guest editors: Gregoire Allaire (CMAP, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France), Jean-Michel Coron (Laboratoire Jacques-Louis Lions, Sorbonne Université) and Vivette Girault (Laboratoire Jacques-Louis Lions, Sorbonne Université)

Les Comptes Rendus. Mécanique *sont membres du Centre Mersenne pour l'édition scientifique ouverte*
www.centre-mersenne.org
e-ISSN : 1873-7234

Comptes Rendus
Mécanique
Published online: 20 July 2023
https://doi.org/10.5802/crmeca.199

The scientific legacy of Roland Glowinski / *L'héritage scientifique de Roland Glowinski*

Optimal feedback control of dynamical systems via value-function approximation

Karl Kunisch [*, a, b] **and Daniel Walter** [c]

[a] University of Graz, Institute of Mathematics and Scientific Computing, Heinrichstr. 36, A-8010 Graz, Austria

[b] Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenberger Straße 69, 4040 Linz, Austria

[c] Institut für Mathematik, Humboldt-Universität zu Berlin, Rudower Chaussee 25, 10117 Berlin, Germany

E-mails: karl.kunisch@uni-graz.at (K. Kunisch), daniel.walter@hu-berlin.de (D. Walter)

Abstract. A self-learning approach for optimal feedback gains for finite-horizon nonlinear continuous time control systems is proposed and analysed. It relies on parameter dependent approximations to the optimal value function obtained from a family of universal approximators. The cost functional for the training of an approximate optimal feedback law incorporates two main features. First, it contains the average over the objective functional values of the parametrized feedback control for an ensemble of initial values. Second, it is adapted to exploit the relationship between the maximum principle and dynamic programming. Based on universal approximation properties, existence, convergence and first order optimality conditions for optimal neural network feedback controllers are proved.

Keywords. optimal feedback control, neural networks, Hamilton–Jacobi–Bellman equation, self-learning, reinforcement learning.

2020 Mathematics Subject Classification. 49J15, 49N35, 68Q32, 93B52, 93D15.

Published online: 20 July 2023

1. Introduction

In this paper we focus on optimal feedback control for problems of the form

$$\left\{\begin{array}{l} \inf_{y,u} J(y,u) \coloneqq \frac{1}{2}\int_0^T \left(|Q_1(y(t)-y_d(t))|^2+\beta|u(t)|^2\right)\,\mathrm{d}t+\frac{1}{2}\left|Q_2\left(y(T)-y_d^T\right)\right|^2 \\ s.t. \quad \dot{y}=f(y)+g(y)u, \quad y(0)=y_0, \text{ and } u\in L^2(0,T;\mathbb{R}^m), \end{array}\right. \tag{P}$$

with nonlinear dynamics described by $f:[0,T]\times\mathbb{R}^n\to\mathbb{R}^n$. The system can be influenced by choosing a control input u which enters through a control operator $g:\mathbb{R}^n\to\mathbb{R}^{n\times m}$. We assess the perfomance of a given control by its objective functional value which comprises the (weighted)

[*] Corresponding author.

distance between the associated state trajectory y and a given desired state y_d as well as the norm of the control for some cost parameter $\beta > 0$. The weighting matrices Q_i, for $i = 1, 2$, are assumed to be symmetric positive semi-definite. Searching for an optimal control u^* in feedback form requires to find a function $F^* : [0, T] \times \mathbb{R}^n \to \mathbb{R}^m$ such that

$$u^*(t) = F^*\left(t, y^*(t)\right), \text{ for } t \in (0, T).$$

Here (u^*, y^*) denotes an optimal control-trajectory pair associated to (P). Under appropriate conditions, see e.g. [1], the feedback mapping can be expressed as

$$F^*(t, y) = -\frac{1}{\beta} g^\top(t, y)\partial_y V^*(t, y), \tag{1}$$

where V^* stands for the value function associate to (P), i.e. for $(T_0, y_0) \in [0, T] \times \mathbb{R}^n$:

$$V^*(T_0, y_0) = \min_{y,u} J_{T_0}(y, u), \text{ subject to } \dot{y} = f(y) + g(y)u, \quad y(T_0) = y_0,$$

and

$$J_{T_0}(y, u) = \frac{1}{2}\int_{T_0}^{T} \left(\left|Q_1\left(y(t) - y_d(t)\right)\right|^2 + \beta|u(t)|^2\right) \mathrm{d}t + \frac{1}{2}\left|Q_2\left(y(T) - y_d^T\right)\right|^2.$$

The value function V^* satisfies a Hamilton–Jacobi–Bellman (HJB) equation which is a time-dependent first order hyperbolic equation of spatial dimension n. Numerical realisations, therefore, are plagued by the curse of dimensionality. Indeed a direct solution of the HJB equation already becomes computationally prohibitive for moderate dimensions n.

Therefore, for practical realization, the interest in alternative techniques arises. In many situations of practical relevance researches have relied on linear approximations to the nonlinear dynamical system and have treated the resulting linear-quadratic problem by Riccati techniques. Much research has concentrated on validating this approach locally around a reference trajectory. Globally such a strategy may fail, see for instance [2, 3].

In this paper we follow an approach, possibly first proposed in [3], circumventing the construction of the value function on the basis of solving the HJB equation. Rather the feedback mapping is constructed by an unsupervised self-learning technique. In practice, this requires the approximation of V^* by a family of functions V_θ which are parametrized by a finite dimensional vector θ and satisfy a uniform approximation property. Possible families of universal approximators include, e.g., neural networks or piecewise polynomial approximations. Subsequently, in view of (1), we introduce the corresponding feedback law

$$F_\theta(t, y) = -\frac{1}{\beta} g^\top(y)\partial_y V_\theta(t, y), \text{ for } (t, y) \in [0, \infty) \times \mathbb{R}^n, \tag{2}$$

as approximation to F^*. An "optimal" parametrized feedback law is then determined by a variant of the following self-learning, structure preserving, variational problem:

$$\begin{aligned} \min_\theta\ & J(y, \mathscr{F}_\theta(y)) + \frac{1}{2}\int_0^T \gamma_1\left|V_\theta(t, y(t)) - J_t\left(y, F_\theta(\cdot, y)\right)\right|^2 + \gamma_2\left|\partial_y V_\theta(t, y(t)) - p(t)\right|^2 \mathrm{d}t + \frac{\gamma_\varepsilon}{2}|\theta|^2 \\ & \text{s.t.} \quad \dot{y} = f(y) + g(y)F_\theta(y), \quad y(0) = y_0, \quad p(T) = Q_2^\top Q_2\left(y(T) - y_d^T\right) \\ & \qquad -\dot{p} = f(y)^\top p + \left|Dg(y)^\top F_\theta(y)\right| p + Q_1^\top Q_1(y - y_d). \end{aligned} \tag{3}$$

In this problem, minimization with respect to u is replaced by minimizing with respect to the parameters θ which characterize V_θ and F_θ. The cost functional of problem (3) consists of four parts: The first term represents the objective functional of (P) where the control u is replaced by the closed loop expression $F_\theta(y)$. The next two terms realize the fact that V_θ is constructed as approximation to the value function associated to (P) and exploit the well-known property that, under certain conditions, the gradient of the value function coincides with the solution of a suitable adjoint equation, see e.g. [1, p. 21]. The final term penalizes the norm of the structural parameters. We point out that V_θ and F_θ are learned along the orbit $\mathcal{O} = \{y(t; y_0) : t \in (0,\infty)\}$ within the state space $\mathbb{R}^n$. To accommodate the case that one trajectory does not provide enough information, we propose to involve an ensemble of orbits departing from a set Y_0 of initial conditions, and to reformulate problem (3) accordingly. This will be done in Section 4 below. While we focus on linear-quadratic objective functionals, the derived results can be readily generalized if suitable coercivity and differentiability properties are assumed for this functional.

In our earlier work on learning a feedback function [3], we considered infinite horizon optimal control problems. In that case, the time-dependent HJB equation results in a stationary one. There we had not yet incorporated the structure preserving terms involving V_θ and $\partial_y V_\theta$ into the cost. Moreover we directly constructed an approximation F_θ to the vector valued function F^*, rather than approximating the scalar valued function V^* and subsequently using (2). In the present paper we provide the theoretical foundations for the learning based technique that we propose to construct an approximation to the optimal feedback function for (P). Recently in [4] a variant of the approach as in [3] was used for interesting numerical investigations to construct optimal feedback functions for finite horizon multi-agent optimal control problems.

Let us very briefly mention some of the vast literature on solving the HJB equations. Semi-Lagrangian schemes and finite difference methods have been deeply investigated to directly solve HJB equations directly, see e.g. [5–7]. Significant progress was made in solving high dimensional HJB equations by the of use policy iterations combined with tensor calculus techniques, [2, 8, 9]. The use of Hopf formulas was proposed in e.g. [10, 11]. Interpolation techniques, utilizing ensembles of open loop solutions have been analyzed in the works of [12, 13], for example. Finally we mention that optimal feedback control is intimately related to reinforcement learning, see e.g. the monograph [14], and also the survey articles [15–17].

The manuscript is structured as follows. Some pertinent notation is gathered in Section 2. In Section 3 concepts of optimal feedback control, semi-global with respect to the initial condition y_0, are gathered. Section 4 is devoted to describing the learning technique that we propose to approximate the optimal feedback function. In Section 5 the required assumptions on approximating subspaces are checked for a class of neural networks and a class of piecewise polynomials. Existence of solutions to the approximating learning problems is proved in Section 6. Their convergence is analyzed in Section 7. The case of learning from finitely many orbits is the focus of Section 8. Section 9 provides an example illustrating the numerical feasibility of the proposed method. We do not aim for sophistication in this respect. The Appendix 9.2 details the proofs of several necessary technical results.

2. Notation

For $I := (0, T)$, with $T > 0$, we define $W_T = \{y \in L^2(I;\mathbb{R}^n) \mid \dot{y} \in L^2(I;\mathbb{R}^n)\}$, where the temporal derivative is understood in the distributional sense. We equip W_T with the norm induced by the inner product

$$(y_1, y_2)_{W_T} = (\dot{y}_1, \dot{y}_2)_{L^2(I;\mathbb{R}^n)} + (y_1, y_2)_{L^2(I;\mathbb{R}^n)} \quad \text{for } y_1, y_2 \in W_T,$$

making it a Hilbert space. We recall that W_T embeds continuously into $C(\bar{I};\mathbb{R}^n)$. For a compact metric space X we denote the space of continuous functions between X and Y by $\mathscr{C}(X;Y)$ which we endow with $\|\varphi\|_{\mathscr{C}(X;Y)} = \max_{x\in X} \|\varphi(x)\|_Y$ as norm. By Y_0 we denote a compact set of initial conditions in $\mathbb{R}^n$. When arising as index, the space $\mathscr{C}(Y_0; W_T)$ will frequently be abbreviated by $\mathscr{C}$. The space $\mathscr{C}^1(X;Y)$ of continuously differentiable functions is defined analogously. Open balls of radius ε in a Banach space X with center x will be denoted by $B_\varepsilon(x)$. The space of bounded linear operators between Banach spaces X and Y, endowed with the canonical norm, is denoted by $\mathscr{B}(X, Y)$. We further abbreviate $\mathscr{B}(X) := \mathscr{B}(X, X)$.

3. Semi-global optimal feedback control

Consider the controlled nonlinear dynamical system of the form

$$\dot{y} = \mathbf{f}(y) + \mathbf{g}(y)u \quad \text{in } L^2\left(I;\mathbb{R}^n\right), \quad y(0) = y_0, \tag{4}$$

described by Nemitsky operators

$$\begin{aligned} &\mathbf{f}\colon W_T \to L^2\left(I;\mathbb{R}^n\right), && \mathbf{f}(y)(t) = f(t,y(t)) \\ &\mathbf{g}\colon W_T \to \mathscr{L}\left(L^2\left(I;\mathbb{R}^m\right);L^2\left(I;\mathbb{R}^n\right)\right), && \mathbf{g}(y)(t) = g(t,y(t)) \end{aligned} \tag{5}$$

for a.e. $t \in I$, $f\colon I\times\mathbb{R}^n \to \mathbb{R}^m$ and $g\colon I\times\mathbb{R}^n \to \mathbb{R}^{n\times m}$. The smoothness requirements on f and g will be detailed in Assumption 1 below. Our aim is to choose a control input $u^* \in L^2(I;\mathbb{R}^m)$ which keeps the associated solution $y^* \in W_T$ close to a known reference trajectory y_d, while keeping the control effort small. This is formulated as the constrained minimization problem

$$\begin{cases} \inf\limits_{y\in W_T,\, u\in L^2(I;\mathbb{R}^m)} J(y,u) \\ s.t. \quad \dot{y} = \mathbf{f}(y) + \mathbf{g}(y)u, \quad y(0) = y_0, \end{cases} \tag{P_{y_0}}$$

where

$$J(y,u) = \frac{1}{2}\int_I \left(\left|Q_1\left(y(t)-y_d(t)\right)\right|^2 + \beta|u(t)|^2\right)\,\mathrm{d}t + \frac{1}{2}\left|Q_2\left(y(T)-y_d^T\right)\right|^2,$$

which incorporates the weighted misfit between the trajectory y within the time horizon $I=(0,T)$ and at the terminal time to desired states $y_d \in L^2(I;\mathbb{R}^n)$ and $y_d^T \in \mathbb{R}^n$, as well as the norm of the control u. While this *open loop* optimal control problem captures well the objective formulated above, it comes with several disadvantages. First, its solution is a function of time only, and does not include the current state $y(t)$. This makes the open loop approach susceptible to possible perturbations in the dynamical system. Second, determining the control action for a new initial condition requires to solve (P_{y_0}) from the start.

The aforementioned limitations of open loop optimal controls motivate the study of *semi-global optimal feedback control* approaches to (P_{y_0}). More precisely, given a compact set $Y_0 \subset \mathbb{R}^n$, we look for a feedback function $F^*\colon I\times\mathbb{R}^n \to \mathbb{R}^m$ which induces a Nemitsky operator

$$\mathscr{F}^*\colon W_T \to L^2\left(I;\mathbb{R}^m\right), \quad \mathscr{F}^*(y)(t) = F^*(t,y(t)) \quad \text{for a.e. } t\in I,$$

such that for every $y_0 \in Y_0$ the *closed loop system*

$$\dot{y} = \mathbf{f}(y) + \mathbf{g}(y)\mathscr{F}^*(y), \quad y(0) = y_0, \tag{6}$$

admits a unique solution $y^*(y_0) \in W_T$ and $(y^*(y_0), \mathscr{F}^*(y^*(y_0)))$ is a minimizing pair of (P_{y_0}).

The determination of an optimal feedback function usually rests on the computation of the value function to(P_{y_0}) which is defined as

$$V^*(T_0,y_0) := \min_{\substack{y\in H^1(T_0,T;\mathbb{R}^n),\\ u\in L^2(T_0,T;\mathbb{R}^m)}} J_{T_0}(y,u) \quad s.t. \quad \dot{y} = \mathbf{f}(y) + \mathbf{g}(y)u, \quad y(t_0) = y_0, \tag{7}$$

where $(T_0,y_0) \in I\times\mathbb{R}^n$, and $J_{T_0}(y,u)$ is defined as

$$J_{T_0}(y,u) = \frac{1}{2}\int_{T_0}^T \left(\left|Q_1(y(t)-y_d(t))\right|^2 + \beta|u(t)|^2\right)\,\mathrm{d}t + \frac{1}{2}\left|Q_2(y(T)-y_d(T))\right|^2.$$

By construction V^* satisfies the final time boundary condition

$$V^*(T,y_0) = \frac{1}{2}\left|Q_2(y_0-y_d(T))\right|^2 \quad \forall\, y_0\in\mathbb{R}^n.$$

If V^* is continuously differentiable in a neighborhood of some $(t, y_0) \in I \times \mathbb{R}^n$ then it solves the instationary *Hamilton–Jacobi–Bellman (HJB) equation*

$$\partial_t V^*(t, y_0) + \left(f(y_0), \partial_y V^*(t, y_0)\right)_{\mathbb{R}^n} - \frac{1}{2\beta}\left|g(t, y_0)^\top \partial_y V^*(t, y_0)\right|^2 + \frac{1}{2}\left|Q_1(y_0 - y_d(t))\right|^2 = 0 \tag{8}$$

in the classical sense there, see e.g. [1, 18]. Here $\partial_t V^*$ denotes the partial derivative of the value function with respect to t and $\partial_y V^*$ is the gradient of V^* with respect to the y-variable. An optimal control for (P_{y_0}) in feedback form is then given by $u^* = -\frac{1}{\beta}\mathbf{g}(y^*)^\top \partial_y \mathcal{V}^*(y^*)$ where $\partial_y \mathcal{V}^*(y^*)(t) = \partial_y V^*(t, y^*(t))$ for every $t \in I$, and $y^* = y^*(y_0) \in W_T$ solves the closed loop system

$$\dot{y} = \mathbf{f}(y) - \frac{1}{\beta}\mathbf{g}(y)\mathbf{g}(y)^\top \partial_y \mathcal{V}^*(y), \quad y(0) = y_0.$$

Thus

$$\left(y^*(y_0), -\frac{1}{\beta}\mathbf{g}\left(y^*(y_0)\right)^\top \partial_y \mathcal{V}^*\left(y^*(y_0)\right)\right) \in \operatorname{argmin}\ (P_{y_0})$$

and the function

$$F^*(\cdot,\cdot) = -\frac{1}{\beta} g(\cdot,\cdot)^\top \partial_y V^*(\cdot,\cdot)$$

is an optimal feedback law.

Realizing the optimal feedback in this way requires a solution to (8) which is a partial differential equation on $\mathbb{R}^n$. This can be extremely challenging or even impossible depending on the dimension n and the computational facilities at hand. Similarly to our previous manuscript [3], we take a different approach by formulating minimization problem over a suitable set of feedback functions involving the closed loop system as a constraint. This relates to a learning problem, within which the feedback functions are trained to achieve optimal stabilization. This makes the problem computationally amenable.

The procedure just described will be formalized in the following section. Here we first summarize the assumptions on the nonlinear dynamical system that we refer to throughout the paper.

Assumption 1.

(**A.1**) *The functions $f\colon I \times \mathbb{R}^n \to \mathbb{R}^n$ and $g\colon I \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are twice continuously differentiable. Their Jacobians and Hessians with respect to the second variable, denoted by $D_y f, D_{yy} f$, and $D_y g, D_{yy} g$, respectively, are Lipschitz continuous on compact sets, uniformly for $t \in I$.*

(**A.2**) *There exists a constant $M_{Y_0} > 0$ such that the value function $V(\cdot,\cdot)\colon I \times \mathbb{R}^n \to \mathbb{R}$ for (P_{y_0}) is twice continuously differentiable on $I \times \bar{B}_{2\widehat{M}}(0)$ with Lipschitz continuous gradient and Hessian (w.r.t. y uniformly in $t \in I$) where*

$$\widehat{M} = M_{Y_0} \|\imath\|_{\mathscr{B}(W_T, \mathscr{C}(I;\mathbb{R}^n))}, \tag{9}$$

and $\imath$ denotes the embedding of W_T into $\mathscr{C}(I;\mathbb{R}^n)$).

As a consequence of (**A.1**), the Nemitsky operators $\mathbf{f}$, $\mathbf{g}$ are at least two times continuously differentiable with domains and ranges as defined in (5). Their derivatives, denoted by $D\mathbf{f}(y) \in \mathscr{B}(W_T, L^2(I;\mathbb{R}^n))$, and $D\mathbf{g}(y) \in \mathscr{B}(W_T; \mathscr{B}(L^2(I;\mathbb{R}^m); L^2(I;\mathbb{R}^n)))$, are the Nemitsky operators induced by $D_y f$ and $D_y g$. Moreover $\mathbf{f}, D\mathbf{f}, \mathbf{g}, D\mathbf{g}$ are Lipschitz continuous and bounded, on bounded subsets of $L^\infty(I;\mathbb{R}^n)$, and thus in particular on $\mathscr{Y}_{ad} \subset W_T$, where

$$\mathscr{Y}_{ad} := \left\{ y \in W_T \mid \|y\|_{W_T} \le 2M_{Y_0} \right\}. \tag{10}$$

Finally $D\mathbf{f}^\top(y) \in \mathscr{B}(W_T, L^2(I;\mathbb{R}^n))$ denotes the Nemitsky operator associated to $D_y f^\top$.

Analogously, due to (**A.2**), V^* induces a twice Lipschitz continuously Fréchet differentiable Nemitsky operator $\mathcal{V}^* : \mathcal{Y}_{ad} \subset W_T \to L^2(I)$. Moreover $\mathcal{V}^*$ and its first derivative $D\mathcal{V}^*$ are weak-to-strong continuous. Define the Nemitsky operator

$$\mathcal{F}^* : \mathcal{Y}_{ad} \to L^2\left(I;\mathbb{R}^n\right), \quad \mathcal{F}^*(y) = -\frac{1}{\beta}\mathbf{g}(y)^\top \partial_y \mathcal{V}^*(y), \tag{11}$$

where $\partial_y \mathcal{V}^*$ is the Nemitsky operator induced by the gradient $\partial_y V^* = D_y V(\cdot,\cdot)^\top$. Note also that $\mathcal{F}^* \in C^1(W_T;(L^2(I;\mathbb{R}^m);L^2(I;\mathbb{R}^n)))$. We further assume the following:

(**A.3**) For every $y_0 \in Y_0$ there exists a unique function $y = \mathbf{y}^*(y_0) \in W_T$ satisfying

$$\dot{y} = \mathbf{f}(y) + \mathbf{g}(y)\mathcal{F}^*(y), \quad y(0) = y_0, \quad \|y\|_{W_T} \le M_{Y_0}.$$

Moreover we have

$$\left(y^*(y_0), \mathcal{F}^*\left(y^*(y_0)\right)\right) \in \operatorname{argmin}\ (P_{y_0}) \quad \forall\, y_0 \in Y_0.$$

When referring to Assumption A we mean (**A.1**)-(**A.3**). We emphasize that the constant M appearing in (**A.2**) and (**A.3**) is assumed to be same. Note further that as a consequence of (**A.3**) problem (P_{y_0}) admits a solution for each $y_0 \in Y_0$, with the optimal control given by $u^* = \mathcal{F}^*(y^*(y_0)))$.

Remark 2. Using (**A.1**), (**A.3**) as well as the implicit function theorem it can be readily be verified that the mapping $\mathbf{y}^* \colon Y_0 \to W_T$ from (**A.3**) is continuously differentiable. Given $\delta y_0 \in \mathbb{R}^n$ the directional derivative $\delta y := \partial \mathbf{y}^*(y_0)(\delta y_0)$ of $\mathbf{y}^*$ at $y_0 \in Y_0$ in direction δy_0 satisfies the linearized ODE system

$$\dot{\delta y} = D\mathbf{f}\left(\mathbf{y}^*(y_0)\right)\delta y + \left[D\mathbf{g}\left(\mathbf{y}^*(y_0)\right)\delta y\right]\mathcal{F}^*\left(\mathbf{y}^*(y_0)\right) + \mathbf{g}\left(\mathbf{y}^*(y_0)\right)D\mathcal{F}^*\left(\mathbf{y}^*(y_0)\right)\delta y,\, \delta y(0) = \delta y_0.$$

Here $D\mathbf{g}$ is induced by $D_y g$ which is given by

$$\left[D_y g(t,y)\delta y\right]_{ij} = \left(\sum_{k=1}^n \partial_k g_{ij}(t,y)\delta y_k\right) \quad \forall\, \delta y \in \mathbb{R}^n,$$

where $g(y) = (g_{ij})$ and "∂_k" denotes the partial derivative w.r.t to the k^{th} component of y. The transposed $D\mathbf{g}(y)^\top$, which will arise in the adjoint equation below, is induced by the tensor $D_y g(t,\cdot)^\top = (D_y g(t,\cdot)_{kji}) \in \mathbb{R}^{n\times n\times m}$, with $t \in I$. In particular, we readily verify that $D\mathbf{g}(\cdot)^\top \in \mathcal{B}(L^2(I;\mathbb{R}^m);\mathcal{B}(W_T;L^2(I;\mathbb{R}^n)))$.

To end this section we collect structural information on the relation between the adjoined state, denoted by p below, the optima value function V^*, and the induced optimal feedback law $\mathcal{F}^*$.

Proposition 3. *Let Assumption 1 hold. Then there exists a unique continuous mapping* $\mathbf{p}^* \colon Y_0 \to W_T$ *such that for each* $y_0 \in Y_0$ *the tuple* $(y,p) = (\mathbf{y}^*(y_0), \mathbf{p}^*(y_0))$ *satisfies*

$$\frac{d}{dt}y = \mathbf{f}(y) + \mathbf{g}(y)\mathcal{F}^*(y),\ y(0) = y_0, \tag{12}$$

$$-\frac{d}{dt}p = D\mathbf{f}(y)^\top p + \left[D\mathbf{g}(y)^\top \mathcal{F}^*(y)\right]p + Q_1^\top Q_1(y - y_d),\ p(T) = Q_2^\top Q_2\left(y(T) - y_d^T\right), \tag{13}$$

$$\mathcal{F}^*(y) = -\frac{1}{\beta}\mathbf{g}(y)^\top p. \tag{14}$$

Moreover we have

$$V^*(t,y(t)) = J_t\left(y(t), F^*(t,y(t))\right),\ p(t) = \partial_y V(t,y(t)) \quad \forall\, t \in [0,T]. \tag{15}$$

Proof of Proposition 3. By (**A.3**) problem (P_{y_0}) admits a solution for each $y_0 \in Y_0$. Then (**A.1**)-(**A.2**) guarantee that (13), with $y = y(y_0) \in W_T$ the state component of a solution to (P_{y_0}), admits a unique solution p in W_T which continuously depends on $y \in W_T$. Moreover (12) - (14) represent the first order necessary optimality condition for (P_{y_0}) with the optimal control $u(t) = \mathcal{F}^*(y(t))$.

Since $\mathbf{y}^*: Y_0 \to W_T$ is continuous as mentioned in Remark 2 and the solution to (13) depends continuously on $y \in W_T$, the claimed continuity $\mathbf{p}^*: Y_0 \to W_T$ follows. Equation (15) is a direct consequence of the dynamic programming principle, and (**A.3**). □

4. Optimal feedback control by value function approximation

This section is devoted to introducing a family of computationally tractable minimization problems from which we will "learn" approximations of optimal feedback laws. Our approach rests on two main pillars. First, given $\varepsilon > 0$, we consider a family of functions $V^\varepsilon_\theta \in \mathscr{C}(I \times \mathbb{R}^n)$ which are finitely parametrized by $\theta \in \mathscr{R}_\varepsilon \simeq \mathbb{R}^{N_\varepsilon}$, $N_\varepsilon \in \mathbb{N}$. These serve as "discrete" approximations of the optimal value function V^*. The following a priori estimate is assumed, for some fixed $\varepsilon_0 > 0$:

Assumption 4. *For every $0 < \varepsilon \le \varepsilon_0$ there holds $V^\varepsilon_\cdot \in \mathscr{C}^4(\mathscr{R}_\varepsilon \times \mathbb{R} \times \mathbb{R}^n)$ and $V^\varepsilon_\theta(T, y_0) = \frac{1}{2}|Q_2(y_0 - y_d(T))|^2$ for every $y_0 \in \mathbb{R}^n$ and $\theta \in \mathscr{R}_\varepsilon$. Moreover there exists $\theta_\varepsilon \in \mathscr{R}_\varepsilon$ with*

$$\max_{\substack{t \in I,\\ |y| \le 2\widehat{M}}} \left|V^\varepsilon_{\theta_\varepsilon}(t,y) - V^*(t,y)\right| + \left|\partial_y\left(V^\varepsilon_{\theta_\varepsilon}(t,y) - V^*(t,y)\right)\right| + \left\|\partial_{yy}\left(V^\varepsilon_{\theta_\varepsilon}(t,y) - V^*(t,y)\right)\right\| \le c\varepsilon \tag{16}$$

for some $c > 0$ independent of $\varepsilon \in (0, \varepsilon_0]$.

Now recall from (11) that the optimal feedback law $\mathscr{F}^*$ is the superposition operator induced by $F^*(t,y) = -(1/\beta) g(t,y)^\top \partial_y V^*(t,y)$. With the aim of preserving the dependence of the feedback law on the value function in our approximation, we define a set of parametrized feedback laws $\mathscr{F}^\varepsilon_\theta$ associated to V^ε_θ, $\theta \in \mathscr{R}_\varepsilon$, by

$$\mathscr{F}^\varepsilon_\theta(y)(t) = F^\varepsilon_\theta(t, y(t)) = -\frac{1}{\beta} g(t, y(t)) \partial_y V^\varepsilon_\theta(t, y(t))$$

for all $y \in W_T$, $t \in \bar{I}$ and $\theta \in \mathscr{R}_\varepsilon$. A first approach to obtain an optimal feedback law in the form $\mathscr{F}^\varepsilon_\theta$ can then be found by replacing the open loop control u in (P_{y_0}) by the closed loop expression $\mathscr{F}^\varepsilon_\theta(y)$ and minimizing for $\theta \in \mathscr{R}_\varepsilon$:

$$\min_{y \in W_T, \theta \in \mathscr{R}_\varepsilon} J\left(y, F^\varepsilon_\theta(y)\right) + \frac{\gamma_\varepsilon}{2}\|\theta\|^2_{\mathscr{R}_\varepsilon} \quad s.t. \quad \dot{y} = \mathbf{f}(y) + \mathbf{g}(y)\mathscr{F}^\varepsilon_\theta(y), y(0) = y_0, \tag{17}$$

where $\|\cdot\|_{\mathscr{R}_\varepsilon}$ denotes a Hilbert space norm on $\mathscr{R}_\varepsilon$, $\gamma_\varepsilon > 0$ and $y_0 \in Y_0$ is fixed. This represents the goal of finding a feedback law $\mathscr{F}^\varepsilon_\theta$ together with a trajectory $y \in W_T$ which satisfy $(y, \mathscr{F}^\varepsilon_\theta(y)) \in \operatorname{argmin}\ (P_{y_0})$. However, this approach falls short in several aspects. First, we cannot hope to recover a solution of the semiglobal optimal feedback control problem for all $y_0 \in Y_0$, since the minimization in (17) is associate to a single initial condition only. Secondly it misses to impose properties that would guide $\mathscr{F}^\varepsilon_\theta(y)$ to be close to $\mathcal{V}^*$, and it does not exploit the relation between the adjoint state p, see (13), and the gradient of the value function $\partial_y \mathcal{V}^*$. Incorporating this information into the problem can, potentially, lead to improved learning results and improved parameterized feedback laws which behave similarly to $\mathscr{F}^*$. These considerations lead to the second pillar of our approach, namely a succinct choice of the cost for the learning problem. For this purpose we use all of Y_0 as "learning set" for initial conditions. It is endowed with the normalized Lebesgue measure $\mathscr{L}$. Moreover we define the augmented objective

$$J_\varepsilon(y,p,\theta) = J\left(y, \mathscr{F}^\varepsilon_\theta(y)\right) + \int_0^T \frac{\gamma_1}{2}\left|V^\varepsilon_\theta(t,y(t)) - J_t\left(y, \mathscr{F}^\varepsilon_\theta(y)\right)\right|^2 + \frac{\gamma_2}{2}\left|\partial_y V^\varepsilon_\theta(t,y(t)) - p(t)\right|^2 \,\mathrm{d}t \tag{18}$$

for penalty parameters $\gamma_1, \gamma_2 \ge 0$. The arguments in J_t are the restriction of the solution y to the equation in (17) and the feedback $\mathscr{F}^\varepsilon_\theta(y)$ to $[t,T]$. The additional terms in this new objective functional penalize the violation of the cost and its gradient by means of the approximation based on V^ε_θ, i.e. they penalize the differences between $J_t(y, \mathscr{F}^\varepsilon_\theta(y))$ and $V^\varepsilon_\theta(t, y(t)$, as well as $p(t)$ and $\partial_y V^\varepsilon_\theta(t, y(t)$.

Given a strictly positive weight function $\omega \in L^\infty(Y_0);\ 0 < c \leq \omega$ a.e., we thus propose to find a feedback law $\mathscr{F}^\varepsilon_\theta$ by solving the ensemble control problem

$$\min_{\substack{\mathbf{y}\in \mathbf{y}_{ad},\\ \mathscr{F}^\varepsilon_\theta(\mathbf{y})\in \mathbf{U}_{ad},\\ \mathbf{p}\in\mathscr{C}(Y_0;W_T)\\ \theta\in\mathscr{R}_\varepsilon}} \mathscr{J}_\varepsilon(\mathbf{y},\mathbf{p},\theta) := \int_{Y_0} \omega(y_0)\, J_\varepsilon\left(\mathbf{y}(y_0),\mathbf{p}(y_0),\theta\right)\, \mathrm{d}\mathscr{L}(y_0) + \frac{\gamma_\varepsilon}{2}\|\theta\|^2_{\mathscr{R}_\varepsilon} \tag{$\mathscr{P}_\varepsilon$}$$

subject to the system of closed loop state *and* adjoint equations

$$\dot{\mathbf{y}}(y_0) = \mathbf{f}(\mathbf{y}(y_0)) + \mathbf{g}(\mathbf{y}(y_0))\mathscr{F}^\varepsilon_\theta(\mathbf{y}(y_0)) \tag{19}$$

$$-\dot{\mathbf{p}}(y_0) = D\mathbf{f}(\mathbf{y}(y_0))^\top \mathbf{p}(y_0) + \left[D\mathbf{g}(\mathbf{y}(y_0))^\top \mathscr{F}^\varepsilon_\theta(\mathbf{y}(y_0))\right]\mathbf{p}(y_0) + \mathbf{Q}_1^\top \mathbf{Q}_1\left(\mathbf{y}(y_0) - y_d\right) \tag{20}$$

$$\mathbf{y}(y_0)(0) = y_0, \mathbf{p}(y_0)(T) = Q_2^\top Q_2\left(\mathbf{y}(y_0)(T) - y_d^T\right), \mathbf{y}(y_0) \in \mathscr{Y}_{ad} \tag{21}$$

for $\mathscr{L}$-a.e. $y_0 \in Y_0$. Above $\mathbf{y}_{ad} \subset \mathscr{C}(Y_0;W_T)$ and $\mathbf{U}_{ad} \subset L^2(Y_0;L^2(I;\mathbb{R}^m))$ denote the admissible sets of ensemble state trajectories and admissible controls. They will be specified in section 6.

5. Examples

In this section we discuss two particular examples for the parameterized mappings V^ε: deep residual networks and piecewise polynomial functions of sufficiently high degree.

5.1. *Residual networks*

To explain the approximation of the value function by residual neural networks, we first fix some notation. Let $L_\varepsilon \in \mathbb{N}$, $L_\varepsilon \geq 2$, as well as $N^\varepsilon_i \in \mathbb{N}$, $i = 1,\ldots,L_\varepsilon - 1$ be given. We set $N^\varepsilon_0 = n+1$ and $N^\varepsilon_L = 1$. Furthermore define

$$\mathscr{R}_\varepsilon = \bigtimes_{i=1}^{L_\varepsilon-1}\left(\mathbb{R}^{N^\varepsilon_i\times N^\varepsilon_{i-1}} \times \mathbb{R}^{N^\varepsilon_i\times N^\varepsilon_{i-1}} \times \mathbb{R}^{N^\varepsilon_i}\right) \times \mathbb{R}^{N^\varepsilon_L\times N^\varepsilon_{L-1}}.$$

The space $\mathscr{R}_\varepsilon$ is uniquely determined by its *architecture*

$$\operatorname{arch}(\mathscr{R}_\varepsilon) = \left(N^\varepsilon_0, N^\varepsilon_1,\ldots,N^\varepsilon_L\right) \in \mathbb{N}^{L_\varepsilon+1}.$$

A set of parameters $\theta \in \mathscr{R}_\varepsilon$ given by

$$\theta = \left(W_{11}, W_{12}, b_1, \ldots, W_{L_\varepsilon}\right)$$

is called a *neural network* with L_ε *layers*. Moreover let $\sigma \in \mathscr{C}^4(\mathbb{R})$ be given and assume that σ is not a polynomial. The function

$$V^\varepsilon_\theta(t,y) = \frac{1}{2}\left|Q_2\left(y - y_d(T)\right)\right|^2 + f^\sigma_{L_\varepsilon,\theta} \circ f^\sigma_{L_\varepsilon-1,\theta} \circ\cdots\circ f^\sigma_{1,\theta}((t,y)) - f^\sigma_{L_\varepsilon,\theta} \circ f^\sigma_{L_\varepsilon-1,\theta} \circ\cdots\circ f^\sigma_{1,\theta}((T,y)) \tag{22}$$

for $(t,y) \in \mathbb{R}\times\mathbb{R}^n$ where

$$f^\sigma_{L_\varepsilon,\theta}(x) = W_{L_\varepsilon} x \quad \forall\, x \in \mathbb{R}^{N^\varepsilon_{L-1}}$$

as well as

$$f^\sigma_{i,\theta}(x) = \sigma\left(W_{i1}x + b_i\right) + W_{i2}x \quad \forall\, x \in \mathbb{R}^{N^\varepsilon_{i-1}},\ i = 1,\ldots,L_\varepsilon - 1$$

is called the *realization* of θ with *activation function* σ. Here the application of σ is defined to act componentwise i.e. given an index $i \in \{1,\ldots,L_\varepsilon - 1\}$ and $x \in \mathbb{R}^{N^\varepsilon_i}$ we set

$$\sigma(x) = \left(\sigma(x_1),\ldots,\sigma\left(x_{N^\varepsilon_i}\right)\right)^\top.$$

By construction, V_θ^ε satisfies the terminal condition

$$V_\theta^\varepsilon(T,y) = \frac{1}{2}\left|Q_2(y - y_d(T))\right|^2 \quad \forall\, y \in \mathbb{R}^n.$$

Moreover Assumption 4 is fulfilled as confirmed by the following result.

Theorem 5. *For every $\varepsilon > 0$ there exists architectures $\mathcal{R}_\varepsilon$ and $\theta_\varepsilon \in \mathcal{R}_\varepsilon$ such that $V^\varepsilon \in \mathscr{C}^4(\mathcal{R}_\varepsilon \times \mathbb{R} \times \mathbb{R}^n)$ and $V_{\theta_\varepsilon}^\varepsilon$ satisfies* (16).

Proof. Let us set $h(t,y) = V^*(t,y)$ for $(t,y) \in I \times \bar{B}_{2\widehat{M}}(0)$. Then h is twice continuously differentiable on $I \times \bar{B}_{2\widehat{M}}(0)$ and $h(T,y) = \frac{1}{2}|Q_2(y - y_d^T)|^2$. A consequence of the universal approximation theorem implies that for all $\varepsilon > 0$ there exists $\widetilde{h}_\varepsilon \in \mathscr{M}_{\text{net}}$ such that

$$\left\| h - \widetilde{h}_\varepsilon \right\|_{C^2\left(I \times \bar{B}_{2\widehat{M}}(0)\right)} \le \frac{\varepsilon}{2}, \tag{23}$$

where $\mathscr{M}_{\text{net}} = \operatorname{span}\{\sigma(\widetilde{w} \cdot x + \widetilde{b}) : \widetilde{w} \in \mathbb{R}^{n+1}, \widetilde{b} \in \mathbb{R}\}$, see eg [19, Theorem 4.1], [20]. Let us observe that $\widetilde{h}_\varepsilon$ can be expressed as a residual network. Indeed, since

$$\widetilde{h}_\varepsilon = \sum_{i=1}^{M} \widetilde{c}_i \sigma\left(\widetilde{w}_i \cdot x + \widetilde{b}_i\right)$$

for some $M \in \mathbb{N}$, $\widetilde{w}_i \in \mathbb{R}^{n+1}, \widetilde{b}_i, \widetilde{c}_i \in \mathbb{R}$, choosing $L_\epsilon = 2, W_{11} \in \mathbb{R}^{M \times (n+1)}$ with rows $\{\widetilde{w}_i\}_{i=1}^M$,

$$b_1 = \operatorname{col}\left(\widetilde{b}_1, \ldots, \widetilde{b}_M\right), W_2 = (\widetilde{c}_1, \ldots, \widetilde{c}_M), W_{12} = 0,$$

we have $\widetilde{h}_\varepsilon = f_{2,\theta}^\sigma \circ f_{1,\theta}^\sigma$. Moreover, $\widetilde{h}_\varepsilon \in C^4(I \times \bar{B}_{2\widehat{M}})$. Following (22) we define

$$V_{\theta_\varepsilon}^\varepsilon(t,y) = \frac{1}{2}\left|Q_2\left(y - y_d^T\right)\right|^2 + \widetilde{h}_\varepsilon(t,y) - \widetilde{h}_\varepsilon(T,y) \in C^4\left(I \times \bar{B}_{2\widehat{M}}\right).$$

and estimate

$$\begin{aligned}\left\| V_{\theta_\epsilon}^\varepsilon(t,y) - V^*(t,y) \right\|_{C^2} &= \left\| \widetilde{h}_\varepsilon(t,y) - \widetilde{h}_\varepsilon(T,y) + V^*(T,y) - V^*(t,y) \right\|_{C^2} \\ &\le 2\left\| h - \widetilde{h}_\varepsilon \right\|_{C^2} \le \varepsilon,\end{aligned}$$

where all norms are taken over $I \times \bar{B}_{2\widehat{M}}(0)$. This ends the proof. □

5.2. *Piecewise polynomials*

Fix $\varepsilon_0 > 0$, and let $\varepsilon \in (0, \varepsilon_0]$ be arbitrarily fixed. Throughout this subsection we assume (**A.2**) and in particular we shall make use of the global Lipschitz continuity of $D^2 V^*$ on $\bar{K} = \bar{I} \times \bar{B}_{2\widehat{M}}(0)$. Since $\bar{K}$ is compact and hence totally bounded, there exist $n_\varepsilon \in \mathbb{N}$ and $\{(\bar{t}_i, \bar{y}_0^i)\}_{i=1}^{n_\varepsilon} \in \mathbb{R}^{n+1}$ such that

$$\bar{K} \subset \bigcup_{i=1}^{n_\varepsilon} K_i \quad \text{where} \quad K_i = B_\varepsilon\left(\left(t_i, \bar{y}_0^i\right)\right).$$

Note that we do not highlight the dependence of $(t_i, \bar{y}_0^i)$ and K_i on ε. For each i define the parametrized polynomial

$$V_i^\varepsilon(A,b,c,t,y) = \left(t - \bar{t}_i, y - \bar{y}_0^i\right)^\top A\left(t - \bar{t}_i, y - \bar{y}_0^i\right) + b^\top\left(t - \bar{t}_i, y - \bar{y}_0^i\right) + c$$

with

$$(A,b,c,t,y) \in \operatorname{Sym}(n+1) \times \mathbb{R}^{n+1} \times \mathbb{R} \times \mathbb{R}^{n+1},$$

where $Sym(n)$ denotes the space of real symmetric $n \times n$ matrices. Note that V_i^ε is infinitely many times differentiable in all of its arguments.

For each $\varepsilon \in (0,\varepsilon_0]$ we define a special partition of unity $\{\varphi_i\}_{i=1}^{n_\varepsilon}$ subordinate to K_i with $\varphi_i\colon \mathbb{R}\times\mathbb{R}^n \to [0,1]$, satisfying $\mathscr{C}^4$ and

$$\begin{cases} \operatorname{supp}\varphi_i = \bar{K}_i,\ \sum_{i=1}^{n_\varepsilon}\varphi_i(t,y) = 1,\ \forall\,(t,y)\in\bar{K}, \\ \left\|D^j\varphi_i\right\|_{C(\bar{K}_i\cap\bar{K})} = \bar{\mu}\varepsilon^{-j}, \forall\ i=1,\dots,n_\varepsilon, \text{ and } j\in\{1,2\}, \\ \operatorname{card}\left\{i : \varphi_i(t,y)\neq 0\right\} \le \mathfrak{m} \quad \forall\,(t,y)\in\bar{K}, \varepsilon\in(0,\varepsilon_0], \end{cases} \tag{24}$$

with $\bar{\mu}$ and $\mathfrak{m}$ positive constants independent of $i,(t,y)\in\bar{K},\varepsilon\in(0,\varepsilon_0]$. Finally we define

$$\mathscr{R}_\varepsilon = \bigtimes_{i=1}^{n_\varepsilon}\left(\operatorname{Sym}(n+1)\times\mathbb{R}^{n+1}\times\mathbb{R}\right),$$

and introduce the family of parameterized functions on $\mathbb{R}^{n+1}$ by

$$V_\theta^\varepsilon(t,y) = \frac{1}{2}\left|Q_2(y-y_d(T))\right|^2 + \sum_{i=1}^{n_\varepsilon}\varphi_i(t,y)\left(V_i^\varepsilon(A_i,b_i,c_i,t,y) - V_i^\varepsilon(A_i,b_i,c_i,T,y)\right) \tag{25}$$

for $\theta = (A_1,b_1,c_1,\dots,A_{n_\varepsilon},b_{n_\varepsilon},c_{n_\varepsilon})\in\mathscr{R}\varepsilon$. Obviously we have $V_\cdot^\varepsilon(\cdot)\in\mathscr{C}^4(\mathscr{R}_\varepsilon\times\mathbb{R}\times\mathbb{R}^n)$ and

$$V_\theta^\varepsilon(T,y) = \frac{1}{2}\left|Q_2\left(y-y_d^T\right)\right|^2 \quad \forall\ y\in\mathbb{R}^n.$$

Thus the final time condition in the HJB equation is fulfilled. Next we show that V_θ^ε satisfies the approximation property in Assumption 4 for the particular choice of

$$\begin{aligned}\theta_\varepsilon = \big(&\partial_{yy}V^*\left(\bar{t}_1,\bar{y}_0^1\right),\partial_y V^*\left(\bar{t}_1,\bar{y}_0^1\right),V^*\left(\bar{t}_1,\bar{y}_0^1\right),\\ &\dots,\partial_{yy}V^*\left(\bar{t}_{n_\varepsilon},\bar{y}_0^{n_\varepsilon}\right),\partial_y V^*\left(\bar{t}_{n_\varepsilon},\bar{y}_0^{n_\varepsilon}\right),V^*\left(\bar{t}_{n_\varepsilon},\bar{y}_0^{n_\varepsilon}\right)\big),\end{aligned} \tag{26}$$

i.e. V_i^ε in (25) are chosen with

$$\left(\bar{A}_i,\bar{b}_i,\bar{c}_i\right) = \left(\partial_{yy}V^*\left(\bar{t}_i,\bar{y}_0^i\right),\partial_y V^*\left(\bar{t}_i,\bar{y}_0^i\right),V^*\left(\bar{t}_i,\bar{y}_0^i\right)\right) \quad i=1,\dots,n_\varepsilon. \tag{27}$$

Theorem 6. *Let V^ε and θ_ε be chosen according to* (25) *and* (26)*, respectively, and suppose that* (**A.2**) *and* (24) *are satisfied. Then Assumption 4 holds.*

Proof. We already argued that V_θ^ε has the desired regularity. It remains to prove the required approximation capabilities. For abbreviation set $V_i^\varepsilon(t,y) = V_i^\varepsilon(\bar{A}_i,\bar{b}_i,\bar{c}_i,t,y)$, with $(\bar{A}_i,\bar{b}_i,\bar{c}_i)$ as in (27).

Since V_i^ε is the second order Taylor expansion of V at $(\bar{t}_i,\bar{y}_0^i)$ we conclude that

$$\left\|V^* - V_i^\varepsilon\right\|_{C^{2-j}(\bar{K}_i\cap\bar{K})} \le \bar{c}\varepsilon^{j+1}, \quad \text{for } j\in 0,1,2, \tag{28}$$

for some $\bar{c}>0$ depending on the global Lipschitz constant of V on $\bar{K}$, and independent of $\varepsilon\in(0,\varepsilon_0]$ and i. Still recall that the sets K_i depend on ε. To estimate $V^*(t,y)-V_\theta^\varepsilon(t,y)$ we recall that $V^*(T,y) = \frac{1}{2}|Q_2(y-y_d^T)|^2$, and express $V^*(t,y)$ as $V^*(t,y) = V^*(T,y)+V^*(t,y)-V^*(T,y)$. This leads to

$$\begin{aligned}&V^*(t,y)-V_\theta^\varepsilon(t,y)\\ &\qquad= \sum_{i\in\{1,\dots,n_\varepsilon\}}\varphi_i(t,y)\left(V^*(t,y)-V_i^\varepsilon(t,y)\right) + \sum_{i\in\{1,\dots,n_\varepsilon\}}\varphi_i(T,y)\left(V^*(T,y)-V_i^\varepsilon(T,y)\right),\end{aligned}$$

for $(t,y)\in\bar{K}$. From (24) and (24) we deduce that $\|V(t,y)-V_\theta^\varepsilon(t,y)\|_{C(\bar{K})}\le 2\bar{c}\varepsilon^3$.

For the gradient with respect to y we proceed similarly. Fixing $(t,y)\in\bar{K}$ we estimate

$$\left|\partial_y V^*(t,y) - \partial_y V_{\theta_\varepsilon}^\varepsilon(t,y)\right| \le D_1 + D_2$$

where

$$D_1 = \sum_{i\in\{1,\dots,n_\varepsilon\}} \left[\varphi_i(t,y)\left|\partial_y V^*(t,y)-\partial_y V_i^\varepsilon(t,y)\right| + \left|V_i^\varepsilon(t,y)-V^*(t,y)\right|\left|\partial_y\varphi_i(t,y)\right|\right]$$
$$D_2 = \sum_{i\in\{1,\dots,n_\varepsilon\}} \left[\varphi_i(T,y)\left|\partial_y V^*(T,y)-\partial_y V_i^\varepsilon(T,y)\right| + \left|V_i^\varepsilon(T,y)-V^*(T,y)\right|\left|\partial_y\varphi_i(t,y)\right|\right].$$

By (28) with $j=1$ the first terms in D_1 and D_2 can be estimated by $\bar{c}\varepsilon^2$. Using (24) and (28) the second terms in D_1 and D_2 can be bounded by $\mathfrak{m}\bar{\mu}\varepsilon^2$. Combining these estimate we arrive at

$$\left\|\partial_y V^*(t,y)-\partial_y V_\theta^\varepsilon(t,y)\right\|_{C(\bar{K})} \le 2\varepsilon^2(\bar{c}+\mathfrak{m}\bar{\mu}).$$

In an analogous manner one can obtain a bound of the order $O(\varepsilon)$ on the difference of the Hessians of V and V_θ^ε. This finishes the proof of Theorem 6. □

In Appendix A it is shown how standard mollifiers can be used so that (24) is satisfied. This requires some extra attention due to the required bounds on the derivatives of φ_i.

6. Existence of minimizers to $(\mathscr{P}_\varepsilon)$

This section is devoted to proving the existence of minimizing triples to $(\mathscr{P}_\varepsilon)$. Throughout this section c will denote a generic constant independent of $\varepsilon>0$ and $y_0\in Y_0$.

6.1. *Existence of admissible points*

Recall from Assumption 1 and Remark 2 that the optimal ensemble state $\mathbf{y}^*\in\mathscr{C}(Y_0;W_T)$ satisfies $\|\mathbf{y}^*\|_{\mathscr{C}}\le M_{Y_0}$. Accordingly we define the set of admissible states and admissible controls as

$$\mathbf{y}_{ad}=\left\{\mathbf{y}\in\mathscr{C}(Y_0;W_T)\;\middle|\;\|\mathbf{y}\|_{\mathscr{C}}\le 2M_{Y_0}\right\},\quad \mathbf{U}_{ad}:=L^2\left(Y_0;L^2\left(I;\mathbb{R}^m\right)\right).$$

We also recall the definition $\mathscr{Y}_{ad}$ in (10).

To prove the existence of minimizers to $(\mathscr{P}_\varepsilon)$ we first argue that the admissible set

$$\mathscr{N}_{ad}^\varepsilon=\left\{(\mathbf{y},\mathbf{p},\theta)\in\mathbf{y}_{ad}\times\mathscr{C}(Y_0;W_T)\times\mathscr{R}_\varepsilon\,\middle|\,(\mathbf{y},\mathbf{p},\theta)\text{ satisfies (19)–(21)},\mathscr{F}_\theta^\varepsilon(\mathbf{y})\in\mathbf{U}_{ad}\right\} \tag{29}$$

is nonempty for ε small enough. For this purpose consider the family $\theta_\varepsilon\in\mathscr{R}_\varepsilon$, $0<\varepsilon\le\varepsilon_0$, from Assumption 4 as well as the associated closed loop system of state and adjoint equations

$$\dot{y}_\varepsilon=\mathbf{f}(y_\varepsilon)+\mathbf{g}(y_\varepsilon)\mathscr{F}_{\theta_\varepsilon}^\varepsilon(y_\varepsilon), \tag{30}$$
$$-\dot{p}_\varepsilon=D\mathbf{f}(y_\varepsilon)^\top p_\varepsilon+\left[D\mathbf{g}(y_\varepsilon)^\top\mathscr{F}_{\theta_\varepsilon}^\varepsilon(y_\varepsilon)\right]p_\varepsilon+\mathbf{Q}_1^\top\mathbf{Q}_1(y_\varepsilon-y_d), \tag{31}$$

subject to the following initial and terminal conditions

$$y_\varepsilon(0)=y_0,\;p_\varepsilon(T)=Q_2^\top Q_2\left(y_\varepsilon(T)-y_d^T\right),$$

for every $y_0\in Y_0$. We first prove the following approximation result.

Theorem 7. *Let Assumptions 1 and 4 hold. There exists a constant c such that for all $\varepsilon>0$ small enough and for all $y_0\in Y_0$ the system* (30) *and* (31) *admits unique solutions $y_\varepsilon=\mathbf{y}_\varepsilon(y_0)\in\mathscr{Y}_{ad}$ and $p_\varepsilon=\mathbf{p}_\varepsilon(y_0)\in W_T$. Furthermore $\mathbf{y}_\varepsilon\in\mathscr{C}^1(Y_0;W_T)$, $\mathbf{p}_\varepsilon\in\mathscr{C}(Y_0;W_T)$, and $\mathscr{F}^*(\mathbf{y}^*)\in\mathscr{C}(Y_0;L^2(I;\mathbb{R}^m))$ hold and*

$$\left\|\mathbf{y}_\varepsilon-\mathbf{y}^*\right\|_{\mathscr{C}^1(Y_0;W_T)}+\left\|\mathbf{p}_\varepsilon-\mathbf{p}^*\right\|_{\mathscr{C}(Y_0;W_T)}+\left\|\mathscr{F}_{\theta_\varepsilon}^\varepsilon(\mathbf{y}_\varepsilon)-\mathscr{F}^*(\mathbf{y}^*)\right\|_{\mathscr{C}(Y_0;L^2(I;\mathbb{R}^m))}\le c\varepsilon.$$

In particular, $(\mathbf{y}_\varepsilon,\mathbf{p}_\varepsilon,\theta_\varepsilon)\in\mathscr{N}_{ad}^\varepsilon$ for all $\varepsilon>0$ small enough.

In order to prove this we require several auxiliary results.

Lemma 8. *There exists a constant c such that for all ε small enough there holds*

$$\left\| \left(\mathscr{F}^*(y_1) - \mathscr{F}^\varepsilon_{\theta_\varepsilon}(y_1)\right) - \left(\mathscr{F}^*(y_2) - \mathscr{F}^\varepsilon_{\theta_\varepsilon}(y_2)\right) \right\|_{L^2(I;\mathbb{R}^m)} \leq c\varepsilon \|y_1 - y_2\|_{W_T}, \quad \forall\, y_1, y_2 \in \mathscr{Y}_{ad}.$$

Proof. According to the definition of $\mathscr{F}^*$ and $\mathscr{F}^\varepsilon_{\theta_\varepsilon}$ we split

$$\left\| \left(\mathscr{F}^*(y_1) - \mathscr{F}^\varepsilon_{\theta_\varepsilon}(y_1)\right) - \left(\mathscr{F}^*(y_2) - \mathscr{F}^\varepsilon_{\theta_\varepsilon}(y_2)\right) \right\|_{L^2(I;\mathbb{R}^m)} \leq D_1 + D_2$$

with

$$\begin{aligned} D_1 &= 1/\beta \, \|\mathbf{g}(y_1)^\top\|_{\mathscr{B}(L^2(I;\mathbb{R}^n),L^2(I;\mathbb{R}^m))} \left\| \partial_y \left(\left(\mathscr{V}^*(y_1) - \mathscr{V}^\varepsilon_{\theta_\varepsilon}(y_1)\right) - \left(\mathscr{V}^*(y_2) - \mathscr{V}^\varepsilon_{\theta_\varepsilon}(y_2)\right)\right) \right\|_{L^2(I;\mathbb{R}^n)} \\ D_2 &= 1/\beta \, \|\mathbf{g}(y_1)^\top - \mathbf{g}(y_2)^\top\|_{\mathscr{B}(L^2(I;\mathbb{R}^n),L^2(I;\mathbb{R}^m))} \left\| \partial_y \left(\mathscr{V}^*(y_2) - \mathscr{V}^\varepsilon_{\theta_\varepsilon}(y_2)\right) \right\|_{L^2(I;\mathbb{R}^n)}. \end{aligned}$$

Applying the integral mean value theorem yields

$$\begin{aligned} &\left\| \partial_y \left(\left(\mathscr{V}^*(y_1) - \mathscr{V}^\varepsilon_{\theta_\varepsilon}(y_1)\right) - \left(\mathscr{V}^*(y_2) - \mathscr{V}^\varepsilon_{\theta_\varepsilon}(y_2)\right)\right) \right\|_{L^2(I;\mathbb{R}^n)} \\ &\qquad \leq \sup_{s\in[0,1]} \left\| \partial_{yy} \left(\mathscr{V}^*(y_1+sh) - \mathscr{V}^\varepsilon_{\theta_\varepsilon}(y_1+sh)\right) \right\|_{\mathscr{B}(W_T,L^2(I;\mathbb{R}^n))} \|h\|_{W_T} \end{aligned}$$

with $h = y_2 - y_1 \in W_T$. Note that $y_1 + sh \in \mathscr{Y}_{ad}$ for all $s \in [0,1]$. Thus we can use Assumption 4 for every $s \in [0,1]$ and $\delta y \in W_\infty$ and estimate

$$\begin{aligned} &\left\| \partial_{yy} \left(\mathscr{V}^*(y_1+sh) - \mathscr{V}^\varepsilon_{\theta_\varepsilon}(y_1+sh)\right) \delta y \right\|_{L^2(I;\mathbb{R}^n)} \\ &\qquad \leq \sqrt{\int_0^T \left| \partial_{yy} \left(V^*\left(t, y_1(t)+sh(t)\right) - V^\varepsilon_{\theta_\varepsilon}\left(t, y_1(t)+sh(t)\right)\right)\right|^2_{\mathbb{R}^{n\times n}} |\delta y(t)|^2 \mathrm{d}t} \\ &\qquad \leq c\varepsilon \|\delta y\|_{L^2(I;\mathbb{R}^n)} \leq \varepsilon c \|\delta y\|_{W_T}. \end{aligned}$$

Similarly we obtain

$$\left\| \partial_y \left(\mathscr{V}^*(y_2) - \mathscr{V}^\varepsilon_{\theta_\varepsilon}(y_2)\right) \right\|_{L^2(I;\mathbb{R}^n)} = \sqrt{\int_0^T \left| \partial_y \left(V^*(t, y_2(t)) - V^\varepsilon_{\theta_\varepsilon}(t, y_2(t))\right)\right|^2 \mathrm{d}t} \leq \sqrt{T} c\varepsilon.$$

Last recall that $\mathbf{g}$ is Lipschitz continuous and uniformly bounded on $\mathscr{Y}_{ad}$. Combining these facts yields the desired statement. □

With the same arguments the following a priori estimate can be obtained. For the sake of brevity its proof is omitted.

Corollary 9. *There exists a constant c such that for all ε small enough there holds*

$$\left\| \mathscr{F}^*(y) - \mathscr{F}^\varepsilon_{\theta_\varepsilon}(y) \right\|_{L^2(I;\mathbb{R}^m)} \leq c\varepsilon \|y\|_{W_T}, \quad \forall\, y \in \mathscr{Y}_{ad}.$$

Next we establish existence of a unique solution to (30) as well as a first approximation result.

Proposition 10. *Let Assumptions 1 and 4 hold. Then for all $\varepsilon > 0$ small enough there is a unique $\mathbf{y}_\varepsilon \in \mathscr{C}^1(Y_0; W_T)$ such that $y_\varepsilon := \mathbf{y}_\varepsilon(y_0) \in \mathscr{Y}_{ad}$ satisfies* (30) *for all $y_0 \in Y_0$. Moreover there exists a constant c independent of ε such that*

$$\|\mathbf{y}^* - \mathbf{y}_\varepsilon\|_{\mathscr{C}(Y_0;W_T)} + \left\| \mathscr{F}^*(\mathbf{y}^*) - \mathscr{F}^\varepsilon_{\theta_\varepsilon}(\mathbf{y}_\varepsilon) \right\|_{\mathscr{C}(Y_0;L^2(I;\mathbb{R}^m))} \leq c\varepsilon.$$

In particular we have $\|\mathbf{y}_\varepsilon\|_{\mathscr{C}(Y_0;W_T)} \leq 2M_{Y_0}$ for all sufficiently small ε.

Proof. The proof is based on a fixed-point argument. Let $y_0 \in Y_0$ be arbitrary but fixed. Define the set

$$\mathscr{M} = \left\{ y \in W_T \mid \|y\|_{W_T} \leq \frac{3}{2} M_{Y_0} \right\} \subset \mathscr{Y}_{ad}.$$

On $\mathcal{M}$ we consider the mapping $\mathcal{Z}\colon \mathcal{M} \to W_T$, where $z = \mathcal{Z}(y) \in \mathcal{Y}_{ad}$ is the unique solution of

$$\dot{z} = \mathbf{f}(z) + \mathbf{g}(z)\mathcal{F}^*(z) + \mathbf{g}(y)\mathcal{F}^{\varepsilon}_{\theta_\epsilon}(y) - \mathbf{g}(y)\mathcal{F}^*(y), \quad z(0) = y_0. \tag{32}$$

It is well-defined since the perturbation function $v = \mathbf{g}(y)\mathcal{F}^{\varepsilon}_{\theta_\epsilon}(y) - \mathbf{g}(y)\mathcal{F}^*(y) \in L^2(I;\mathbb{R}^n)$ satisfies

$$\|v\|_{L^2} \leq \|\mathbf{g}(y)\|_{\mathcal{B}(L^2(I;\mathbb{R}^m),L^2(I;\mathbb{R}^n))} \left\|\mathcal{F}^*(y) - \mathcal{F}^{\varepsilon}_{\theta_\epsilon}(y)\right\|_{L^2} \leq \frac{3}{2} c\varepsilon M_{Y_0} \|\mathbf{g}(y)\|_{\mathcal{B}(L^2(I;\mathbb{R}^m),L^2(I;\mathbb{R}^n))}$$

where we use Corollary 9 and the definition of $\mathcal{M}$. Hence $\|v\|_{L^2} \leq c\varepsilon$. Here and below c denotes a generic constant which is independent of $y_0 \in Y_0$ and all $\epsilon > 0$ sufficiently small. We may invoke Proposition 30 and Corollary 31 from the Appendix, to assert the existence of a unique solution $z \in \mathcal{Y}_{ad}$ to (32) with

$$\|z\|_{W_T} \leq M_{Y_0} + c\|v\|_{L^2} \leq \frac{3}{2} M_{Y_0}, \quad \forall\, y_0 \in Y_0,$$

if $\varepsilon > 0$ is chosen small enough. From this we particularly conclude $\mathcal{Z}(\mathcal{M}) \subset \mathcal{M}$ for all $y_0 \in Y_0$ and $\varepsilon > 0$ small. It remains to prove that $\mathcal{Z}$ is a contraction. To this end let $y_1,\ y_2 \in \mathcal{M}$ be given. Applying Corollary 31 yields the first inequality in

$$\left\|\mathcal{Z}(y_1) - \mathcal{Z}(y_2)\right\|_{W_T} \leq c \left\|\mathcal{F}^*(y_1) - \mathcal{F}^{\varepsilon}_{\theta_\epsilon}(y_1) - \mathcal{F}^*(y_2) + \mathcal{F}^{\varepsilon}_{\theta_\epsilon}(y_2)\right\|_{L^2} \leq c\varepsilon \|y_1 - y_2\|_{W_T}$$

with a constant $c > 0$ independent of $y_1, y_2 \in \mathcal{M}$ as well as of $y_0 \in Y_0$, and ε sufficiently small. The last inequality follows from Lemma 8. Choosing $\varepsilon > 0$ small enough we conclude that $\mathcal{Z}$ admits a unique fixed point $y_\varepsilon = \mathcal{Z}(y_\varepsilon) \in W_T$ on $\mathcal{M}$. Clearly, the function $\mathbf{y}_\varepsilon(y_0) := y_\varepsilon$ satisfies (30), $y_\varepsilon \in \mathcal{M} \subset \mathcal{Y}_{ad}$ as well as

$$\left\|\mathbf{y}_\varepsilon(y_0) - \mathbf{y}^*(y_0)\right\|_{W_T} = \left\|\mathcal{Z}(\mathbf{y}_\varepsilon(y_0)) - \mathcal{Z}(0)\right\|_{W_T} \leq c\varepsilon \|y_\varepsilon\|_{W_T} \leq c\varepsilon \frac{3}{2} M_{Y_0},$$

and by Corollary 9

$$\begin{aligned}
&\left\|\mathcal{F}^*\left(\mathbf{y}^*(y_0)\right) - \mathcal{F}^{\varepsilon}_{\theta_\epsilon}(\mathbf{y}_\varepsilon(y_0))\right\|_{L^2} \\
&\qquad \leq \|\mathcal{F}^*\left(\mathbf{y}^*(y_0)\right) - \mathcal{F}^*\left(\mathbf{y}_\varepsilon(y_0)\right)\|_{L^2} + \left\|\mathcal{F}^*\left(\mathbf{y}_\varepsilon(y_0)\right) - \mathcal{F}^{\varepsilon}_{\theta_\epsilon}(\mathbf{y}_\varepsilon(y_0))\right\|_{L^2} \\
&\qquad \leq c\left\|\mathbf{y}^*(y_0) - \mathbf{y}_\varepsilon(y_0)\right\|_{W_T} + c\varepsilon\|\mathbf{y}_\varepsilon(y_0)\|_{W_T} \leq c\varepsilon.
\end{aligned}$$

Finally according to Proposition 30 the solution $\mathbf{y}_\varepsilon(y_0)$ is unique and the mapping $\mathbf{y}_\varepsilon$ is at least of class $\mathcal{C}^1$. □

Next we estimate the $W^{1,2}$ difference between $\mathbf{y}_\varepsilon$ and $\mathbf{y}^*$.

Proposition 11. *The mapping $\mathbf{y}_\varepsilon \in \mathcal{C}^1(Y_0; W_T)$ from Theorem 10 satisfies*

$$\left\|\mathbf{y}_\varepsilon - \mathbf{y}^*\right\|_{\mathcal{C}^1(Y_0;W_T)} \leq c\varepsilon$$

for $c > 0$ independently of ε small enough.

Proof. By the previous proposition the estimate is already known for $\mathcal{C}^1(Y_0; W_T)$ replaced by $\mathcal{C}(Y_0; W_T)$. Now fix $y_0 \in Y_0$ and $i \in \{1, \ldots, n\}$. By the inverse mapping theorem the partial derivatives of $\mathbf{y}^*$ and $\mathbf{y}_\varepsilon$ at y_0 are given by $\partial_i \mathbf{y}^*(y_0) = T_*(y_0)^{-1}(0, e_i)$, $\partial_i \mathbf{y}_\varepsilon(y_0) = T_\varepsilon(y_0)^{-1}(0, e_i)$. Here, e_i denotes the i$^{\text{th}}$ canonical basis vector in $\mathbb{R}^n$ and

$$T_*(y_0)^{-1}, T_\varepsilon(y_0)^{-1}\colon L^2\left(I;\mathbb{R}^n\right) \times \mathbb{R}^n \to W_T$$

denote the linear continuous inverses of

$$T_*(y_0)\delta y = \begin{pmatrix} \dot{\delta y} - D\mathbf{f}\left(\mathbf{y}^*(y_0)\right)\delta y - \left[D\mathbf{g}\left(\mathbf{y}^*(y_0)\right)\delta y\right]\mathcal{F}^*\left(\mathbf{y}^*(y_0)\right) - \mathbf{g}\left(\mathbf{y}^*(y_0)\right) D\mathcal{F}^*\left(\mathbf{y}^*(y_0)\right)\delta y \\ \delta y(0) \end{pmatrix}$$

and

$$T_\varepsilon(y_0)\delta y = \begin{pmatrix} \dot{\delta y} - D\mathbf{f}\big(\mathbf{y}_\varepsilon(y_0)\big)\delta y - \big[D\mathbf{g}\big(\mathbf{y}_\varepsilon(y_0)\big)\delta y\big]\mathscr{F}^\varepsilon_{\theta_\varepsilon}(\mathbf{y}_\varepsilon(y_0)) - \mathbf{g}(\mathbf{y}_\varepsilon(y_0))D\mathscr{F}^\varepsilon_{\theta_\varepsilon}(\mathbf{y}_\varepsilon(y_0))\delta y \\ \delta y(0) \end{pmatrix}.$$

Using Gronwall's inequality, we readily verify that

$$\max\left\{\left\| T_\varepsilon(y_0)^{-1}(\delta v,\delta y_0)\right\|_{W_T}, \left\| T_*(y_0)^{-1}(\delta v,\delta y_0)\right\|_{W_T}\right\} \leq C\left(\|\delta v\|_{L^2(I;\mathbb{R}^n)} + |\delta y_0|_{\mathbb{R}^n}\right) \tag{33}$$

for all $\delta v \in L^2(I,\mathbb{R}^n)$, $\delta y_0 \in \mathbb{R}^n$, $y_0 \in Y_0$ and some $C>0$ independent of $y_0, \delta v, \delta y_0$. Now we recall that $\mathbf{y}_\varepsilon(y_0), \mathbf{y}^*(y_0) \in \mathscr{Y}_{ad}$ and that $D\mathbf{f}, D\mathbf{g}, \mathbf{g}$ are Lipschitz continuous, and thus in particular bounded, on $\mathscr{Y}_{ad}$, see Assumption (**A.1**). Together with boundedness of $\{\|\mathscr{F}^*(\mathbf{y}(y_0))\|_{L^2} : y_0 \in Y_0\}$, Corollary 9 and Theorem 10 we conclude

$$\left\|\big(T_*(y_0) - T_\varepsilon(y_0)\big)\delta y\right\|_{L^2(I;\mathbb{R}^n)\times\mathbb{R}^n} \leq c\varepsilon\|\delta y\|_{W_T} \quad \forall\, \delta y \in W_T \tag{34}$$

for some $c>0$ again independent of $y_0 \in Y_0$. Recalling that $B^{-1} - A^{-1} = A^{-1}(A-B)B^{-1}$ for invertible bounded linear operators A and B, we obtain

$$\begin{aligned}\left\|\partial_i \mathbf{y}_\varepsilon(y_0) - \partial_i \mathbf{y}^*(y_0)\right\|_{W_T} &= \left\| T_\varepsilon(y_0)^{-1}(0,e_i) - T_*(y_0)^{-1}(0,e_i)\right\|_{W_T} \\ &\leq C^2 \sup_{\|\delta y\|_{W_T}\leq 1} \left\|\big(T_*(y_0) - T_\varepsilon(y_0)\big)\delta y\right\|_{L^2(I;\mathbb{R}^n)\times\mathbb{R}^n} \leq c\varepsilon,\end{aligned}$$

where $C>0$ is the constant from (33). Since all involved constants are independent of $y_0 \in Y_0$ we obtain the desired estimate $\left\|\partial_i \mathbf{y}_\varepsilon - \partial_i \mathbf{y}^*\right\|_{\mathscr{C}} \leq c\varepsilon$. □

Next we address the solvability of the adjoint equation (20).

Proposition 12. *There exists a constant c such that for all ε small enough there exists $\mathbf{p}_\varepsilon \in \mathscr{C}(Y_0;W_T)$ such that $p_\varepsilon := \mathbf{p}_\varepsilon(y_0) \in W_T$ satisfies* (31) *for all $y_0 \in Y_0$ and*

$$\|\mathbf{p}_\varepsilon - \mathbf{p}^*\|_{\mathscr{C}} \leq c\varepsilon.$$

Proof. Given $y \in \mathscr{Y}_{ad}$ consider the linear ordinary differential equation

$$-\dot{p} = D\mathbf{f}(y)p + \left[D\mathbf{g}(y)^\top \mathscr{F}^\varepsilon_{\theta_\varepsilon}(y)\right]p + \mathbf{Q}_1^\top\mathbf{Q}_1(y - y_d),\ p(T) = Q_2^\top Q_2\left(y(T) - y_d^T\right).$$

It admits a unique solution $p = P(y) \in W_T$ which is bounded independently of $y \in \mathscr{Y}_{ad}$. Moreover the mapping $P\colon W_T \to W_T$ is continuous on $\mathscr{Y}_{ad}$ in virtue of the Gronwall lemma and Assumption 1. The existence of a mapping $\mathbf{p}_\varepsilon$ which satisfies (31) then follows by setting $\mathbf{p}_\varepsilon = P \circ \mathbf{y}_\varepsilon$.

It remains to prove the estimate for the difference between $\mathbf{p}_\varepsilon$ satisfying (31) and $\mathbf{p}^*$ satisfying (13). For this purpose we can use the same technique as in the proof of Proposition 11 and therefore we only give the main estimates. Recall that $D\mathbf{f}(\cdot)^\top, D\mathbf{g}(\cdot)^\top$ are Lipschitz continuous on $\mathscr{Y}_{ad}$. The the most involved term in the estimate analogous to (34) is

$$\begin{aligned}&\left\|\left[D\mathbf{g}\big(\mathbf{y}_\varepsilon(y_0)\big)^\top \mathscr{F}^\varepsilon_{\theta_\varepsilon}\big(\mathbf{y}_\varepsilon(y_0)\big) - D\mathbf{g}\big(\mathbf{y}^*(y_0)\big)^\top \mathscr{F}^*\big(\mathbf{y}^*(y_0)\big)\right]\delta p\right\|_{L^2} \\ &\qquad\leq c\left(\left\|\mathbf{y}_\varepsilon(y_0) - \mathbf{y}^*(y_0)\right\|_{W_T} + \left\|\mathscr{F}^\varepsilon_{\theta_\varepsilon}(\mathbf{y}_\varepsilon(y_0)) - \mathscr{F}^*\big(\mathbf{y}^*(y_0)\big)\right\|_{L^2}\right)\|\delta p\|_{W_T}\end{aligned}$$

with $c>0$ independent of $\varepsilon>0$ and $\delta p \in W_T$. Now a perturbation argument as in the proof of Propostion 11 provides us with

$$\begin{aligned}&\left\|\mathbf{p}_\varepsilon(y_0) - \mathbf{p}^*(y_0)\right\|_{W_T} \\ &\quad\leq c\left(\left\|\mathbf{y}_\varepsilon(y_0) - \mathbf{y}^*(y_0)\right\|_{W_T} + \left\|\mathscr{F}^\varepsilon_{\theta_\varepsilon}\big(\mathbf{y}_\varepsilon(y_0)\big) - \mathscr{F}^*\big(\mathbf{y}^*(y_0)\big)\right\|_{L^2} + \left|\mathbf{y}_\varepsilon(y_0)(T) - \mathbf{y}^*(y_0)(T)\right|\right) \\ &\quad\leq c\left(\left\|\mathbf{y}_\varepsilon(y_0) - \mathbf{y}^*(y_0)\right\|_{W_T} + \left\|\mathscr{F}^\varepsilon_{\theta_\varepsilon}\big(\mathbf{y}_\varepsilon(y_0)\big) - \mathscr{F}^*\big(\mathbf{y}^*(y_0)\big)\right\|_{L^2}\right) \leq c\varepsilon\end{aligned}$$

where $W_T \hookrightarrow \mathscr{C}(\bar{I};\mathbb{R}^n)$ is used in the second inequality, and Proposition 11 and Corollary 9 are utilized in the final one. Since all involved constants are again independent of $y_0 \in Y_0$, this finishes the proof. □

Summarizing all previous observations we arrive at the proof of Theorem 7.

Proof of Theorem 7. This follows directly by combining Proposition 10, Proposition 11, and Proposition 12. □

6.2. *Closedness of $\mathcal{N}^{\varepsilon}_{ad}$*

As a last prerequisite for proving existence to ($\mathscr{P}_\varepsilon$) we argue that the admissible set $\mathcal{N}^{\varepsilon}_{ad}$ is closed. The existence of at least one minimizing triple to ($\mathscr{P}_\varepsilon$) then follows by variational arguments. From here on we always assume that $\mathcal{N}^{\varepsilon}_{ad}$ from (29) is nonempty, i.e. that ε is sufficiently small.

Proposition 13. *Let $(\mathbf{y}_k, \mathbf{p}_k, \theta_k)_{k\in\mathbb{N}} \subset \mathcal{N}^{\varepsilon}_{ad}$ be a sequence with weak limit $(\mathbf{y}, \mathbf{p}, \theta)$ in $L^2(Y_0; W_T)^2 \times \mathscr{R}_\varepsilon$. Then $(\mathbf{y}, \mathbf{p}, \theta) \in \mathcal{N}^{\varepsilon}_{ad}$ and we have*

$$(\mathbf{y}, \mathbf{p}) \in \mathscr{C}(Y_0; W_T)^2, \ \lim_{k\to\infty} \mathbf{y}_k(y_0) = \mathbf{y}(y_0) \text{ and } \lim_{k\to\infty} \mathbf{p}_k(y_0) = \mathbf{p}(y_0) \text{ in } W_T, \quad \forall\ y_0 \in Y_0.$$

The proof builds upon the following two lemmas.

Lemma 14. *Let the sequence $(\mathbf{y}_k, \mathbf{p}_k, \theta_k)_{k\in\mathbb{N}} \subset \mathcal{N}^{\varepsilon}_{ad}$ satisfy the prerequisites of Proposition 13. Then $\mathbf{y} \in \mathbf{y}_{ad}$, $\mathbf{y}_k(y_0) \to \mathbf{y}(y_0)$ in W_T, $\mathscr{F}^{\varepsilon}_{\theta_k}(\mathbf{y}_k(y_0)) \to \mathscr{F}^{\varepsilon}_{\theta}(\mathbf{y}(y_0))$ in $L^\infty(I;\mathbb{R}^m)$, and*

$$\dot{\mathbf{y}}(y_0) = \mathbf{f}\big(\mathbf{y}(y_0)\big) + \mathbf{g}\big(\mathbf{y}(y_0)\big)\mathscr{F}^{\varepsilon}_{\theta}(\mathbf{y}(y_0)),\ \mathbf{y}(y_0)(0) = y_0, \tag{35}$$

for all $y_0 \in Y_0$.

Proof. By assumption we have $\mathbf{y}_k \in \mathbf{y}_{ad}$, and hence $\|\mathbf{y}_k(y_0)\|_{W_T} \le 2M_{Y_0}$ for all $k \in \mathbb{N}$ and $y_0 \in Y_0$, and $\mathbf{y}_k \in \mathscr{C}^1(Y_0; W_T)$ for all $k \in \mathbb{N}$, see Proposition 10. Let us fix an arbitrary $y_0 \in Y_0$. and set $y_k := \mathbf{y}_k(y_0)$ for abbreviation. Then there exists a subsequence, denoted by the same index, and $\tilde{y} \in W_T$ such that $y_k \rightharpoonup \tilde{y}$ in W_T. Since $W_T \hookrightarrow_c \mathscr{C}(\bar{I};\mathbb{R}^n) \hookrightarrow L^p(I;\mathbb{R}^n)$, $1 \le p \le +\infty$, we immediately get

$$y_k(0) \to \tilde{y}(0) \text{ in } \mathbb{R}^n,\ \mathbf{f}(y_k) \to \mathbf{f}(\tilde{y}) \text{ in } L^2(I;R^n),\ \mathbf{g}(y_k) \to \mathbf{g}(\tilde{y}) \text{ in } \mathscr{B}\left(L^2\left(I;\mathbb{R}^m\right), L^2\left(I;\mathbb{R}^n\right)\right)$$

as well as $\mathscr{F}^{\varepsilon}_{\theta}(y_k) \to \mathscr{F}^{\varepsilon}_{\theta}(\tilde{y})$ in $L^\infty(I;\mathbb{R}^m)$. Moreover by Assumption 4 for every $\delta > 0$ there exits $K_\delta \in \mathbb{N}$ such that

$$\left|\partial_y V^{\varepsilon}_{\theta_k}(t,y) - \partial_y V^{\varepsilon}_{\theta}(t,y)\right| \le \delta \quad \forall\ (t,y) \in \bar{I} \times \bar{B}_{2\widehat{M}}(0) \tag{36}$$

for all $k \ge K_\delta$. Here $\widehat{M}$ denotes the constant from (**A.2**). For all such k we get utilizing (36) for a constant c independent of k

$$\begin{aligned}\left\|\mathscr{F}^{\varepsilon}_{\theta_k}(y_k) - \mathscr{F}^{\varepsilon}_{\theta}(\tilde{y})\right\|_{L^\infty} &\le c\left\|\mathscr{F}^{\varepsilon}_{\theta_k}(y_k) - \mathscr{F}^{\varepsilon}_{\theta}(y_k)\right\|_{L^\infty} + \left\|\mathscr{F}^{\varepsilon}_{\theta}(y_k) - \mathscr{F}^{\varepsilon}_{\theta}(\tilde{y})\right\|_{L^\infty} \\ &\le c\delta + \left\|\mathscr{F}^{\varepsilon}_{\theta}(y_k) - \mathscr{F}^{\varepsilon}_{\theta}(\tilde{y})\right\|_{L^\infty}.\end{aligned}$$

This implies that $\lim_{k\to\infty} \mathscr{F}^{\varepsilon}_{\theta_k}(y_k) = \mathscr{F}^{\varepsilon}_{\theta}(\tilde{y})$ in $L^\infty(I;\mathbb{R}^m)$. These observations imply

$$\dot{y}_k = \mathbf{f}(y_k) + \mathbf{g}(y_k)\mathscr{F}^{\varepsilon}_{\theta_k}(y_k) \to \mathbf{f}(\tilde{y}) + \mathbf{g}(\tilde{y})\mathscr{F}^{\varepsilon}_{\theta}(\tilde{y}).$$

Together with $y_k \rightharpoonup \tilde{y}$ in W_T this implies that $\mathbf{y}_k(y_0) = y_k \to \tilde{y}$ in W_T and

$$\dot{\tilde{y}} = \mathbf{f}(\tilde{y}) + \mathbf{g}(\tilde{y})\mathscr{F}^{\varepsilon}_{\theta}(\tilde{y}), \tilde{y}(0) = y_0. \tag{37}$$

Since the solution to this equation is unique, every weak accumulation point of y_k satisfies (37) and we have $\mathbf{y}_k(y_0) \to \tilde{y}$ in W_T for the whole sequence. We repeat this construction for all $y_0 \in Y_0$. This defines a function $\tilde{\mathbf{y}} \colon Y_0 \to W_T$ such that $\mathbf{y}_k(y_0) \to \tilde{\mathbf{y}}(y_0)$ in W_T and such that (37) is satisfied with $\tilde{y} = \tilde{\mathbf{y}}(y_0)$ for each $y_0 \in Y_0$. By Proposition 10 it is the unique solution to (35).

Lebesgue's dominated convergence theorem for Bochner integrals [21, p. 45] implies that $\mathbf{y}_k \to \tilde{\mathbf{y}}$ in $L^1(Y_0; W_T)$, and by boundedness of $\{\|\mathbf{y}_k\|_{\mathscr{C}}\}_{k=1}^{\infty}$ also in $L^2(Y_0; W_T)$. By assumption $\mathbf{y}_k$ converges weakly in $L^2(Y_0; W_T)$ to $\mathbf{y}$. Thus we have $\mathbf{y} = \tilde{\mathbf{y}}$. Moreover $\|\mathbf{y}\|_{\mathscr{C}} \le 2M_{Y_0}$ and hence $\mathbf{y} \in \mathbf{y}_{ad}$. □

Next we consider the behavior of the adjoint states $\mathbf{p}_k$.

Lemma 15. *Let $(\mathbf{y}_k,\mathbf{p}_k,\theta_k)_{k\in\mathbb{N}} \subset \mathcal{N}^\varepsilon_{ad}$ be a sequence with weak limit $(\mathbf{y},\mathbf{p},\theta)$ satisfying the prerequisites of Proposition 13. Then $\|\mathbf{p}_k\|_{\mathscr{C}} \le C$ for some $C>0$ and all $k\in\mathbb{N}$ large enough, and $\mathbf{p}\in\mathscr{C}(Y_0;W_T)$. Moreover $\mathbf{p}_k(y_0)\to\mathbf{p}(y_0)$ in W_T, and*

$$
\begin{aligned}
-\dot{\mathbf{p}}(y_0) &= D\mathbf{f}(\mathbf{y}(y_0))^\top \mathbf{p}(y_0) + \left[D\mathbf{g}(\mathbf{y}(y_0))^\top \mathscr{F}^\varepsilon_\theta(\mathbf{y}(y_0))\right]\mathbf{p}(y_0) + \mathbf{Q}_1^\top\mathbf{Q}_1(\mathbf{y}(y_0)-y_d),\\
\mathbf{p}(y_0))(T) &= Q_2^\top Q_2\left(\mathbf{y}(T)(y_0)-y_d^T\right),
\end{aligned}
\tag{38}
$$

for all $y_0\in Y_0$.

Proof. From Lemma 14 recall that for the sequences $y_k := \mathbf{y}_k(y_0)\in\mathscr{Y}_{ad}$ and $y:=\mathbf{y}(y_0)$ we have for each $y_0\in Y_0$

$$
y_k \to y \text{ in } W_T,\ \mathscr{F}^\varepsilon_{\theta_k}(y_k)\to\mathscr{F}^\varepsilon_\theta(y) \text{ in } L^\infty(I;\mathbb{R}^m).
$$

Further for each $k\in\mathbb{N}$ and $y_0\in Y_0$, the element $p_k := \mathbf{p}_k(y_0)\in W_T$ satisfies

$$
-\dot{p}_k = D\mathbf{f}(y_k)^\top p_k + \left[D\mathbf{g}(y_k)^\top \mathscr{F}^\varepsilon_{\theta_k}(y_k)\right] p_k + \mathbf{Q}_1^\top\mathbf{Q}_1(y_k-y_d),\ p_k(T) = Q_2^\top Q_2\left(y_k(T)-y_d^T\right). \tag{39}
$$

Recall from Assumption 4 that $\partial_y V^\varepsilon_\cdot$ is uniformly continuous on compact sets. Thus for every $\delta>0$ there is $K_\delta\in\mathbb{N}$ such that

$$
\left|F^\varepsilon_{\theta_k}(t,x)\right| \le \left|F^\varepsilon_{\theta_k}(t,x)-F^\varepsilon_\theta(t,x)\right| + \left|F^\varepsilon_\theta(t,x)\right| \le \delta + \max_{(t,x)\in I\times\bar{B}_{2\widehat{M}}(0)}\left|F^\varepsilon_\theta(t,x)\right| < \infty
$$

for all $(t,x)\in I\times\bar{B}_{2\widehat{M}}(0)$ and $k\ge K_\delta$. Consequently we obtain

$$
\sup_{k\ge K_\delta}\max_{(t,x)\in I\times\bar{B}_{2\widehat{M}}(0)} \|A_k(t,x)\|_{\mathbb{R}^{n\times n}} < \infty, \text{ where } A_k(t,x) = Df(t,x)^\top + Dg(t,x)^\top F^\varepsilon_{\theta_k}(t,x).
$$

Applying Proposition 29 to the time-reversed equation (39) implies that

$$
\|p_k\|_{W_T} \le c\left(\left\|\mathbf{Q}_1^\top\mathbf{Q}_1(y_k-y_d)\right\|_{L^2} + \left|y_k(T)-y_d^T\right|\right)
$$

for some $c>0$ independent of $y_0\in Y_0$ and all sufficiently large k. Since $\|\mathbf{y}_k\|_{\mathscr{C}}\le 2M_{Y_0}$ we finally conclude $\|\mathbf{p}_k\|_{\mathscr{C}}\le C$ for some $C>0$ independent of k sufficiently large. We are now prepared to pass to the limit in (39). For this purpose we proceed as in the proof of Lemma 14 and use

$$
D\mathbf{f}(y_k) + D\mathbf{g}(y_k)^\top\mathscr{F}^\varepsilon_{\theta_k}(y_k) \to D\mathbf{f}(y) + D\mathbf{g}(y)^\top\mathscr{F}^\varepsilon_\theta(y) \text{ in } \mathscr{B}\left(L^2(Y;\mathbb{R}^n)\right),
$$

as well as

$$
\mathbf{Q}_1^\top\mathbf{Q}_1(y_k-y_d) \to \mathbf{Q}_1^\top\mathbf{Q}_1(y-y_d) \text{ in } L^2\left(I;\mathbb{R}^n\right),
$$

and

$$
Q_2^\top Q_2(y_k(T)-y_d^T) \to Q_2^\top Q_2\left(y(T)-y_d^T\right) \text{ in } \mathbb{R}^n
$$

to show that every weak accumulation point $\widetilde{p}\in W_T$ of p_k is in fact a strong accumulation point and satisfies the differential equation in (38). Since the solution to this equation is unique we get $p_k\to\widetilde{p}$ in W_T for the whole sequence. Finally utilizing $\|\mathbf{p}_k\|_{\mathscr{C}}\le C$ and Lebesgue's dominated convergence theorem we conclude $\widetilde{p}=\mathbf{p}(y_0)$ for all $y_0\in Y_0$. □

Proof of Proposition 13. This is a direct consequence of Lemma 14 and Lemma 15. □

6.3. *Existence of minimizers*

Finally we prove the existence of at least one minimizing triplet to $(\mathscr{P}_\varepsilon)$.

Theorem 16. *Let Assumption 1 and 4 hold. Then for all $\varepsilon>0$ small enough, Problem $(\mathscr{P}_\varepsilon)$ admits at least one minimizing triplet $(\mathbf{y}^*_\varepsilon,\mathbf{p}^*_\varepsilon,\theta^*_\varepsilon)\in\mathscr{C}(Y_0;W_T)^2\times\mathscr{R}_\varepsilon$.*

Proof. According to Theorem 7, the admissible set $\mathscr{N}^\varepsilon_{ad}$ is nonempty for $\varepsilon>0$ small enough. Fix such a $\varepsilon>0$ and let $(\mathbf{y}_k,\mathbf{p}_k,\theta_k)\in\mathscr{N}^\varepsilon_{ad}$ denote a minimizing sequence for $\mathscr{J}_\varepsilon$ i.e.

$$\mathscr{J}_\varepsilon(\mathbf{y}_k,\mathbf{p}_k,\theta_k)\to\inf_{(\mathbf{y},\mathbf{p},\theta)\in\mathscr{N}^\varepsilon_{ad}}\mathscr{J}_\varepsilon(\mathbf{y},\mathbf{p},\theta).$$

Since $\mathbf{y}_k\in\mathbf{y}_{ad}$ and $\frac{\gamma_\varepsilon}{2}\|\theta_k\|^2_{\mathscr{R}_\varepsilon}\leq\mathscr{J}_\varepsilon(\mathbf{y}_k,\mathbf{p}_k,\theta_k)$, for all $k\in\mathbb{N}$, the sequence $\{(\mathbf{y}_k,\theta_k)\}\in L^2(Y_0;W_T)\times\mathscr{R}_\varepsilon$ is bounded. Thus it admits at least one subsequence, denoted by the same index, with

$$(\mathbf{y}_k,\theta_k)\rightharpoonup(\mathbf{y}^*_\varepsilon,\theta^*_\varepsilon)\text{ in }L^2(Y_0;W_T)\times\mathscr{R}_\varepsilon$$

for some $(\mathbf{y}^*_\varepsilon,\theta^*_\varepsilon)$. As in the proof of Lemma 15 we verify that $\|\mathbf{y}_k\|_{\mathscr{C}}\leq C$ and $\|\mathbf{p}_k\|_{\mathscr{C}}\leq C$ for some $C>0$ independent of $k\in\mathbb{N}$. Consequently, by possibly taking another subsequence we arrive at

$$(\mathbf{y}_k,\mathbf{p}_k,\theta_k)\rightharpoonup\left(\mathbf{y}^*_\varepsilon,\mathbf{p}^*_\varepsilon,\theta^*_\varepsilon\right)\text{ in }L^2(Y_0;W_T)^2\times\mathscr{R}_\varepsilon$$

for some $(\mathbf{y}^*_\varepsilon,\mathbf{p}^*_\varepsilon,\theta^*_\varepsilon)\in\mathscr{N}^\varepsilon_{ad}$. For the following estimates it will be convenient to recall the augmented functional J_ε, see (18), which arises in the running cost of $(\mathscr{P}_\varepsilon)$ in compact form:

$$J_\varepsilon(y,p,\theta)=J\left(y,\mathscr{F}^\varepsilon_\theta(y)\right)+\frac{\gamma_1}{2}\left\|\mathscr{V}(y)-J_\bullet\left(y,\mathscr{F}^\varepsilon_{\theta^*_\varepsilon}(y)\right)\right\|^2_{L^2(I;\mathbb{R})}+\frac{\gamma_2}{2}\left\|p-\partial_y\mathscr{V}(y)\right\|^2_{L^2(I;\mathbb{R}^n)},\tag{40}$$

where J_t was defined below (7). Now fix an arbitrary $y_0\in Y_0$ and set

$$y_k:=\mathbf{y}_k(y_0),\ p_k:=\mathbf{p}_k(y_0),\ y^*:=\mathbf{y}^*_\varepsilon(y_0),\ p:=\mathbf{p}^*_\varepsilon(y_0).$$

From Lemma 14 and Lemma 15 we get

$$y_k\to\widetilde{y},\ p_k\to p\text{ in }W_T,\ \mathscr{F}^\varepsilon_{\theta_k}(y_k)\to\mathscr{F}^\varepsilon_{\theta^*_\varepsilon}(\widetilde{y})\text{ in }L^2\left(I;\mathbb{R}^n\right)$$

and, again using the uniform continuity of $V^\varepsilon_\bullet$ and $\partial_y V^\varepsilon_\bullet$, we conclude

$$\mathscr{V}^\varepsilon_{\theta_k}(y_k)\to\mathscr{V}^\varepsilon_{\theta^*_\varepsilon}(\widetilde{y})\text{ in }L^2(I),\ \partial_y\mathscr{V}^\varepsilon_{\theta_k}(y_k)\to\partial_y\mathscr{V}^\varepsilon_{\theta^*_\varepsilon}(\widetilde{y})\text{ in }L^2\left(I;\mathbb{R}^n\right),$$

as well as the uniform boundedness of $\mathscr{V}^\varepsilon_{\theta_k}(\mathbf{y}_k)$ and $\partial_y\mathscr{V}^\varepsilon_{\theta_k}(\mathbf{y}_k)$ in $\mathscr{C}(Y_0;L^2(I))$ and $\mathscr{C}(Y_0;L^2(I;\mathbb{R}^n))$, respectively. Moreover we readily verify that

$$\left|J_t(y_k,\mathscr{F}^\varepsilon_{\theta_k}(y_k))-J_t\left(\widetilde{y},\mathscr{F}^\varepsilon_{\theta^*_\varepsilon}(\widetilde{y})\right)\right|\leq c\left(\|y_k-\widetilde{y}\|_{L^2}+\left\|\mathscr{F}^\varepsilon_{\theta_k}(y_k)-\mathscr{F}^\varepsilon_{\theta^*_\varepsilon}(\widetilde{y})\right\|_{L^2}+\left|y_k(T)-\widetilde{y}(T)\right|\right),$$

for some $c>0$ independent of $y_0\in Y_0$, $t\in(0,T)$, and $k\in\mathbb{N}$. Thus we arrive at

$$J_\bullet\left(y_k,\mathscr{F}^\varepsilon_{\theta_k}(y_k)\right)\to J_\bullet\left(y,\mathscr{F}^\varepsilon_{\theta^*_\varepsilon}(y)\right)\text{ in }L^\infty(I).$$

Summarizing the previous findings there holds

$$\left\|\mathscr{V}(y_k)-J_\bullet\left(y_k,\mathscr{F}^\varepsilon_{\theta_k}(y_k)\right)\right\|^2_{L^2}+\|p_k-\partial_y\mathscr{V}(y_k)\|^2_{L^2}\to\left\|\mathscr{V}(y)-J_\cdot\left(y,\mathscr{F}^\varepsilon_{\theta^*_\varepsilon}(y)\right)\right\|^2_{L^2}+\|p-\partial_y\mathscr{V}(y)\|^2_{L^2}$$
$$J\left(y_k,\mathscr{F}^\varepsilon_{\theta_k}(y_k)\right)\to J\left(y,\mathscr{F}^\varepsilon_{\theta^*_\varepsilon}(y)\right).$$

Using these expressions in J_ε as given in (40), and the boundedness of $\|y_k\|_{L^2},|y_k(0)|$, $\|p_k\|_{L^2}$ $\|\mathscr{F}^\varepsilon_{\theta_k}(y_k)\|_{L^2}$, $\|\mathscr{V}^\varepsilon_{\theta_k}(y_k)\|_{L^2}$ independent of $k\in\mathbb{N}$ and $y_0\in Y_0$ we finally get by using Lebesgue's dominated convergence theorem

$$\mathscr{J}_\varepsilon\left(\mathbf{y}_k,\mathbf{p}_k,\theta_k\right)\to\mathscr{J}_\varepsilon\left(\mathbf{y}^*_\varepsilon,\mathbf{p}^*_\varepsilon,\theta^*_\varepsilon\right)=\inf_{(\mathbf{y},\mathbf{p},\theta)\in\mathscr{N}^\varepsilon_{ad}}\mathscr{J}_\varepsilon(\mathbf{y},\mathbf{p},\theta).$$ □

7. Convergence towards optimal controls

In Proposition 10 and 12 it was established that the ensemble triple $(\mathbf{y}^*, \mathscr{F}(\mathbf{y}^*), \mathbf{p}^*)$ can be approximated by ensemble triples $(\mathbf{y}_\varepsilon, \mathscr{F}^\varepsilon_{\theta_\varepsilon}(\mathbf{y}_\varepsilon), \mathbf{p}_\varepsilon)$ in the order $O(\varepsilon)$. In this section, the convergence of solutions to $(\mathscr{P}_\varepsilon)$ as $\varepsilon \to 0$ is addressed. We first consider the terms in the definition $\mathscr{J}_\varepsilon$, see (18). To obtain the desired asymptotic behavior a smallness condition on the regularisation parameter γ_ϵ, in relation to the norm of the parameters θ_ε describing the approximation quality, is required.

Theorem 17. *Let Assumptions 1 and 4 hold the latter with $\theta_\varepsilon \in \mathscr{R}_\varepsilon$, and let $(\mathbf{y}^*_\varepsilon, \mathbf{p}^*_\varepsilon, \theta^*_\varepsilon)$, denote an optimal triple to $(\mathscr{P}_\varepsilon)$ for all $\varepsilon > 0$ small enough. If additionally $\gamma_\varepsilon \|\theta_\varepsilon\|^2_{\mathscr{R}_\varepsilon} = O(\varepsilon)$, then*

$$0 \le \int_{Y_0} \omega(y_0) \left[J\left(\mathbf{y}^*_\varepsilon(y_0)\right), \mathscr{F}^\varepsilon_{\theta^*_\varepsilon}\left(\mathbf{y}^*_\varepsilon(y_0)\right) - V^*(0, y_0) \right] \, \mathrm{d}\mathscr{L}(y_0) \le c\,\varepsilon$$

holds and, if $\gamma_1, \gamma_2 > 0$, we also have

$$\int_{Y_0} \omega(y_0) \left(\left\| V^\varepsilon_{\theta^*_\varepsilon}\left(t, \mathbf{y}^*_\varepsilon(y_0)\right) - J_\bullet\left(\mathbf{y}^*_\varepsilon(y_0), \mathscr{F}^\varepsilon\left(\mathbf{y}^*_\varepsilon(y_0)\right)\right) \right\|^2_{L^2} + \left\| \partial_y V^\varepsilon_{\theta^*_\varepsilon}\left(t, \mathbf{y}^*_\varepsilon(y_0)\right) - \mathbf{p}^*_\varepsilon(y_0) \right\|^2_{L^2} \right) \mathrm{d}\mathscr{L}(y_0) \le c\,\varepsilon$$

for some $c > 0$ independent of ε.

Proof. Let $\mathbf{y}_\varepsilon, \mathbf{p}_\varepsilon$ denote the ensembles of state and adjoint trajectories associated to θ_ε, see Theorem 7, for $\varepsilon > 0$ small enough. Then we have

$$\begin{aligned} &\left| J\left(\mathbf{y}_\varepsilon(y_0)\right), \mathscr{F}^\varepsilon_{\theta_\varepsilon}\left(\mathbf{y}_\varepsilon(y_0)\right) - V^*(0, y_0) \right| \\ &\qquad \le C\left(\left\|\mathbf{y}_\varepsilon(y_0) - \mathbf{y}^*(y_0)\right\|_{W_T} + \left\| \mathscr{F}^\varepsilon_{\theta_\varepsilon}\left(\mathbf{y}_\varepsilon(y_0)\right) - \mathscr{F}^*\left(\mathbf{y}^*(y_0)\right) \right\|_{L^2(I;\mathbb{R}^m)} \right) \le c\varepsilon \end{aligned}$$

for some $C > 0$ independent of ε. Here we have used $V^*(0, y_0) = J(\mathbf{y}^*(y_0), \mathscr{F}^*(\mathbf{y}^*(y_0)))$ for all $y_0 \in Y_0$, the embedding $W_T \hookrightarrow \mathscr{C}(\bar{I}; \mathbb{R}^n)$ as well as the a priori estimates of Proposition 10. Next we utilize $\mathbf{p}^*(y_0) = \partial \mathcal{V}^*(\mathbf{y}^*(y_0))$, $y_0 \in Y_0$, to estimate

$$\begin{aligned} &\left\| \partial_y \mathcal{V}^\varepsilon_{\theta_\varepsilon}\left(\mathbf{y}_\varepsilon(y_0)\right) - \mathbf{p}_\varepsilon(y_0) \right\|^2_{L^2(I;\mathbb{R}^n)} \\ &\qquad \le 2\left(\left\| \partial_y \mathcal{V}^\varepsilon_{\theta_\varepsilon}\left(\mathbf{y}_\varepsilon(y_0)\right) - \partial_y \mathcal{V}^*\left(\mathbf{y}^*(y_0)\right) \right\|^2_{L^2(I;\mathbb{R}^n)} + \left\|\mathbf{p}_\varepsilon(y_0) - \mathbf{p}^*(y_0)\right\|^2_{L^2(I;\mathbb{R}^n)} \right) \\ &\qquad \le c\varepsilon^2, \end{aligned}$$

where the last inequality is deduced from Proposition 10 and Proposition 12. Proceeding analogously and using $V^*(t, \mathbf{y}^*(y_0)(t)) = J_t(\mathbf{y}^*(y_0), \mathscr{F}^*(\mathbf{y}^*(y_0)))$ for all $y_0 \in Y_0$, $t \in I$, we obtain

$$\int_{Y_0} \omega(y_0) \int_0^T \left| V^\varepsilon_{\theta_\varepsilon}\left(t, \mathbf{y}_\varepsilon(y_0)(t)\right) - J_t\left(\mathbf{y}_\varepsilon(y_0)\right), \mathscr{F}^\varepsilon_{\theta_\varepsilon}\left(\mathbf{y}_\varepsilon(y_0)\right) \right|^2 \, \mathrm{d}t\,\mathrm{d}\mathscr{L}(y_0) \le D_1 + D_2,$$

where, using Assumption 4 and again Proposition 10

$$\begin{aligned} D_1 &:= \int_{Y_0} \omega(y_0) \left\| \mathcal{V}^\varepsilon_{\theta_\varepsilon}\left(\mathbf{y}_\varepsilon(y_0)\right) - \mathcal{V}^*\left(\mathbf{y}^*(y_0)\right) \right\|^2_{L^2(I)} \, \mathrm{d}\mathscr{L}(y_0) \le c\varepsilon^2, \\ D_2 &:= \int_{Y_0} \omega(y_0) \int_0^T \left| J_t\left(\mathbf{y}^*(y_0)\right), \mathscr{F}^*\left(\mathbf{y}^*(y_0)\right) - J_t\left(\mathbf{y}_\varepsilon(y_0)\right), \mathscr{F}^\varepsilon_{\theta_\varepsilon}\left(\mathbf{y}_\varepsilon(y_0)\right) \right|^2 \, \mathrm{d}t\mathrm{d}\mathscr{L}(y_0) \le c\varepsilon^2. \end{aligned}$$

Combining the previous estimates with the optimality of $(\mathbf{y}^*, \mathbf{p}^*, \theta^*_\varepsilon)$, and the assumption on the asymptotic behavior of γ_ϵ we deduce that

$$\begin{aligned} 0 &\le \int_{Y_0} \omega(y_0) \left[J_\varepsilon\left(\mathbf{y}^*_\varepsilon(y_0), \mathbf{p}^*_\varepsilon(y_0), \theta^*_\varepsilon\right) - V^*(0, y_0) \right] \, \mathrm{d}\mathscr{L}(y_0) + \frac{\gamma_\varepsilon}{2} \left\|\theta^*_\varepsilon\right\|^2_{\mathscr{R}_\varepsilon} \\ &\le \int_{Y_0} \omega(y_0) \left[J_\varepsilon\left(\mathbf{y}_\varepsilon(y_0), \mathbf{p}_\varepsilon(y_0), \theta_\varepsilon\right) - V^*(0, y_0) \right] \, \mathrm{d}\mathscr{L}(y_0) + \frac{\gamma_\varepsilon}{2} \|\theta_\varepsilon\|^2_{\mathscr{R}_\varepsilon} \le c\,\varepsilon. \end{aligned}$$

Recalling the definition of J_ε, this yields all claimed estimates and finishes the proof. □

Next the convergence of the ensemble trajectories $(\mathbf{y}^*_\varepsilon, \mathbf{p}^*_\varepsilon)$, the feedback controls $\mathscr{F}^\varepsilon_{\theta^*_\varepsilon}(\mathbf{y}^*_\varepsilon)$ as well as the approximate value function $\mathcal{V}^\varepsilon_{\theta^*_\varepsilon}$ are analyzed. For this purpose we make use of the additional regularity of ensemble solutions to the closed loop system, see Proposition 11, and introduce further constraints to $(\mathscr{P}_\varepsilon)$. Without changing the notation we henceforth set

$$\widehat{\mathbf{Y}}_{ad} = \left\{ \mathbf{y} \in \mathscr{C}^1(Y_0; W_T) \;\middle|\; \|\mathbf{y}\|_{\mathscr{C}} \leq 2M_{Y_0},\ \|\mathbf{y}\|_{W^{1,2}} \leq 2M_{W^{1,2}} \right\}, \tag{41}$$

where $M_{W^{1,2}} > 0$ is a constant with $\|\mathbf{y}^*\|_{W^{1,2}} \leq M_{W^{1,2}}$, the function y^* was introduced in (**A.3**), and $W^{1,2} = \{\mathbf{y} \in L^2(I; W_T) : \partial_i \mathbf{y} \in L^2(Y_0; W_T), i \in \{1, \ldots, n\}\}$ endowed with the natural norm. Next we note that

$$\frac{\beta}{2} \left\| \mathscr{F}^* \left(\mathbf{y}^*(y_0) \right) \right\|^2_{L^2} \leq J\left(\mathbf{y}^*(y_0), \mathscr{F}^*\left(\mathbf{y}^*(y_0) \right) \right) = V^*(0, y_0)$$

for all $y_0 \in Y_0$. Thus, due to the continuity of the value function V^*, see (**A.2**), there is $M_U > 0$ with $\|\mathscr{F}^*(\mathbf{y}^*)\|_{L^\infty} \leq M_U$. Correspondingly we set

$$\widehat{\mathbf{U}}_{ad} = \left\{ \mathbf{u} \in L^\infty\left(Y_0; L^2\left(I; \mathbb{R}^m \right) \right) \mid \|\mathbf{u}\|_{L^\infty} \leq 2M_U \right\}. \tag{42}$$

We point out that Theorem 16 remains valid despite the additional restriction of the set of admissible states and controls. Problem $(\mathscr{P}_\varepsilon)$ with $\mathbf{Y}_{ad}, \mathbf{U}_{ad}$ replaced by $\widehat{\mathbf{Y}}_{ad}, \widehat{\mathbf{U}}_{ad}$ will be denoted by $(\widehat{\mathscr{P}}_\epsilon)$.

Proposition 18. *Let Assumption 1 and 4 hold. Then for all $\varepsilon > 0$ small enough, Problem $(\widehat{\mathscr{P}}_\epsilon)$ admits at least one minimizing triple.*

Proof. Let $(\mathbf{y}_\varepsilon, \mathbf{p}_\varepsilon, \theta_\varepsilon)$ be defined as in Theorem 7. Then we have $\mathbf{y}_\varepsilon \in \widehat{\mathbf{Y}}_{ad}$, see Proposition 10 and Proposition 11, as well as $\mathscr{F}^\varepsilon_{\theta_\varepsilon}(\mathbf{y}_\varepsilon) \in \widehat{\mathbf{U}}_{ad}$, according to Proposition 10, for all $\varepsilon > 0$ small enough. Hence the admissible set of $(\widehat{\mathscr{P}}_\epsilon)$ is not empty. The existence of a minimizing triple then follows by repeating the arguments of the proof of Theorem 16 noting that the admissible set

$$\left\{ (\mathbf{y}, \mathbf{p}, \theta) \in \widehat{\mathbf{Y}}_{ad} \times \mathscr{C}(Y_0; W_T) \times \mathscr{R}_\varepsilon \;\middle|\; (\mathbf{y}, \mathbf{p}, \theta) \text{ satisfies } (19)-(21), \mathscr{F}^\varepsilon_\theta(\mathbf{y}) \in \widehat{\mathbf{U}}_{ad} \right\}$$

is closed w.r.t to the weak topology on $L^2(Y_0; W_T)^2 \times \mathscr{R}_\varepsilon$. □

Let us next address the convergence of the optimal ensemble states $\mathbf{y}^*_\varepsilon$, adjoint states $\mathbf{p}_\varepsilon$ and the associated feedback controls $\mathscr{F}^\varepsilon_{\theta^*_\varepsilon}(\mathbf{y}^*_\varepsilon)$ as ε tends to 0.

Theorem 19. *Let the prerequisites of Theorem 17 hold, and let $\varepsilon_k > 0$ be a strictly decreasing null sequence such that $(\widehat{\mathscr{P}}_{\varepsilon_k})$ admits a minimizing triple $(\mathbf{y}^*_k, \mathbf{p}^*_k, \theta^*_k)$. Then $(\mathbf{y}^*_k, \mathbf{p}^*_k, \mathscr{F}^{\varepsilon_k}_{\theta^*_{\varepsilon_k}}(\mathbf{y}^*_{\varepsilon_k}))$ contains at least one accumulation point $(\bar{\mathbf{y}}, \bar{\mathbf{p}}, \bar{\mathbf{u}}) \in L^\infty(Y_0; W_T)^2 \times L^\infty(Y_0; L^2(I; \mathbb{R}^m))$ w.r.t the strong topology on $L^2(Y_0; W_T)^2 \times L^2(Y_0; L^2(I; \mathbb{R}^m))$. For each accumulation point and $\mathscr{L}$-a.e. $y_0 \in Y_0$ we have that $(\bar{y}, \bar{p}, \bar{u}) := (\bar{\mathbf{y}}(y_0), \bar{\mathbf{p}}(y_0), \bar{\mathbf{u}}(y_0))$ satisfies*

$$(\bar{y}, \bar{u}) \in \min\ (P_{y_0})$$

as well as

$$\begin{aligned}
\dot{\bar{y}} &= \mathbf{f}(\bar{y}) + \mathbf{g}(\bar{y})\bar{u},\ \bar{y}(0) = y_0, \\
-\dot{\bar{p}} &= D\mathbf{f}(\bar{y})^\top \bar{p} + \left[D\mathbf{g}(\bar{y})^\top \bar{u} \right] \bar{p} + \mathbf{Q}_1^\top \mathbf{Q}_1 (\bar{y} - y_d),\ \bar{p}(T) = Q_2^\top Q_2 \left(\bar{y}(T) - y_d^T \right).
\end{aligned}$$

Proof. By choice of the admissible sets $\widehat{\mathbf{Y}}_{ad}$ and $\widehat{\mathbf{U}}_{ad}$ we have that $\{(\mathbf{y}^*_k, \mathscr{F}^\varepsilon_{\theta^*_k}(\mathbf{y}^*_k))\}_{k=1}^\infty$ is bounded in $(W^{1,2}(Y_0; W_T) \cap L^\infty(Y_0; W_T)) \times L^\infty(Y_0; L^2(I; \mathbb{R}^m))$. By Gronwall's inequality we can argue that $\{\mathbf{p}^*_k\}_{k=1}^\infty$ is also bounded in $L^\infty(Y_0; W_T)$. Thus, due to the Banach–Alaoglu theorem, there is a subsequence, denoted by the same index, and $(\bar{\mathbf{y}}, \bar{\mathbf{p}}, \bar{\mathbf{u}}) \in L^\infty(Y_0; W_T)^2 \times L^\infty(Y_0; L^2(I; \mathbb{R}^m))$ such that

$$\left(\mathbf{y}^*_k, \mathbf{p}^*_k, \mathscr{F}^{\varepsilon_k}_{\theta^*_k}(\mathbf{y}^*_k) \right) \rightharpoonup^* \left(\bar{\mathbf{y}}, \bar{\mathbf{p}}, \bar{\mathbf{u}} \right) \text{ in } L^\infty(Y_0; W_T)^2 \times L^\infty\left(Y_0; L^2(I; \mathbb{R}^m) \right),$$

and $\dot{\mathbf{y}}_k^* \rightharpoonup \dot{\bar{\mathbf{y}}}$ in $L^2(Y_0; L^2(I;\mathbb{R}^n))$. By the compact embedding of $W^{1,2}(Y_0; W_T)$ into $L^2(Y_0; \mathscr{C}(I;\mathbb{R}^n))$, see [22, Theorem 5.3] the subsequence can be chosen such that $\mathbf{y}_k^* \to \bar{\mathbf{y}}$ strongly in $L^2(Y_0; \mathscr{C}(I;\mathbb{R}^n))$. These properties imply that $(\bar{y}, \bar{u}) \coloneqq (\bar{\mathbf{y}}(y_0), \bar{\mathbf{u}}(y_0))$ satisfies

$$\dot{\bar{y}} = \mathbf{f}(\bar{y}) + \mathbf{g}(\bar{y})\bar{u}, \;\; \bar{y}(0) = y_0, \tag{43}$$

for $\mathscr{L}$-a.e. $y_0 \in Y_0$. This also implies $V^*(0, y_0) \leq J(\bar{y}, \bar{u})$ and thus, together with

$$J\left(\mathbf{y}_k^*, \mathscr{F}_{\theta_k^*}^{\varepsilon_k}(\mathbf{y}_k^*)\right) \to V^*(0,\cdot) \text{ in } L^1(Y_0; W_T),$$

see Theorem 17, we have $(\bar{y}, \bar{u}) \in \operatorname{arg\,min}\, (P_{y_0})$ for $\mathscr{L}$-a.e. $y_0 \in Y_0$. Moreover, again using the strong convergence of $\mathbf{y}_k^*$ in $L^2(Y_0; \mathscr{C}(I;\mathbb{R}^n))$ and recalling the definition of $J(\cdot,\cdot)$ as

$$J(y,u) = (1/2)\left\|\mathbf{Q}_1(y - y_d)\right\|^2_{L^2(I;\mathbb{R}^n)} + (\beta/2)\|u\|^2_{L^2(I;\mathbb{R}^n)} + (1/2)\left|Q_2\left(y(T) - y_d^T\right)\right|^2,$$

for all $y \in W_T$, $u \in L^2(I;\mathbb{R}^m)$, we also conclude the convergence of the $L^2(Y_0; L^2(I;\mathbb{R}^m))$ norm of $\mathscr{F}_{\theta_k^*}^{\varepsilon_k}(\mathbf{y}_k^*)$ towards the norm of $\bar{\mathbf{u}}$. Thus $\mathscr{F}_{\theta_k^*}^{\varepsilon_k}(\mathbf{y}_k^*) \to \bar{\mathbf{u}}$ strongly in $L^2(Y_0; L^2(I;\mathbb{R}^m))$, and $\mathbf{y}_k^* \to \bar{\mathbf{y}}$ strongly in $L^2(Y_0; W_T)$, by Lebegue's bounded convergence theorem.

It remains to address the strong convergence of $\mathbf{p}_k$. For this purpose we show that the functions $[D\mathbf{g}(\mathbf{y}_k^*)^\top \mathscr{F}_{\theta_k^*}^{\varepsilon_k}(\mathbf{y}_k^*(\cdot))]\mathbf{p}_k^*(\cdot)$ converge weakly to $[D\mathbf{g}(\bar{\mathbf{y}}(\cdot))^\top \bar{\mathbf{u}}(\cdot)]\bar{\mathbf{p}}(\cdot)$ in $L^2(Y_0; L^2(I;\mathbb{R}^n))$. Fixing a test function $\varphi \in L^2(Y_0; L^2(I;\mathbb{R}^n))$ we first note that

$$\lim_{k\to\infty}\left(\varphi, \left[D\mathbf{g}(\bar{\mathbf{y}}(\cdot))^\top \bar{\mathbf{u}}(\cdot)\right]\left(\mathbf{p}_k^*(\cdot) - \bar{\mathbf{p}}\right)\right)_{L^2(Y_0;L^2(I;\mathbb{R}^n))} = 0.$$

Second, for $\mathscr{L}$-a.e. $y_0 \in Y_0$ we estimate

$$\begin{aligned}
&\left(\varphi(y_0), \left[D\mathbf{g}\left(\mathbf{y}_k^*(y_0)\right)^\top \mathscr{F}_{\theta_k^*}^{\varepsilon_k}\left(\mathbf{y}_k^*(y_0)\right) - D\mathbf{g}\left(\bar{\mathbf{y}}(y_0)\right)^\top \bar{\mathbf{u}}(y_0)\right]\mathbf{p}_k^*(y_0)\right)_{L^2(I;\mathbb{R}^n)} \\
&\leq C\|\varphi(y_0)\|_{L^2}\left\|\mathbf{p}_k^*(y_0)\right\|_{W_T}\left(\left\|\mathscr{F}_{\theta_k^*}^{\varepsilon_k}\left(\mathbf{y}_k^*(y_0)\right)\right\|_{L^2}\left\|\mathbf{y}_k^*(y_0) - \bar{\mathbf{y}}(y_0)\right\|_{W_T} + \left\|\mathscr{F}_{\theta_k^*}^{\varepsilon_k}\left(\mathbf{y}_k^*(y_0)\right) - \bar{\mathbf{u}}(y_0)\right\|_{L^2}\right) \\
&\leq C\|\varphi(y_0)\|_{L^2}\left(\left\|\mathbf{y}_k^*(y_0) - \bar{\mathbf{y}}(y_0)\right\|_{W_T} + \left\|\mathscr{F}_{\theta_k^*}^{\varepsilon_k}\left(\mathbf{y}_k^*(y_0)\right) - \bar{\mathbf{u}}(y_0)\right\|_{L^2}\right)
\end{aligned}$$

for some $C > 0$ independent of $k \in \mathbb{N}$ and y_0. Here we made use of the boundedness of $\{\mathbf{y}_k^*\}_{k=1}^\infty$ and $\{\mathbf{p}_k^*\}_{k=1}^\infty$ in $L^\infty(Y_0; W_T)$, and of $\{\mathscr{F}_{\theta_k^*}^{\varepsilon_k}(\mathbf{y}_k^*)\}_{k=1}^\infty$ in $L^\infty(Y_0; L^2(I;\mathbb{R}^m))$. Integrating both sides of the inequality w.r.t to $\mathscr{L}$ and utilizing the strong convergence of $\mathbf{y}_k^*$ and $\mathscr{F}_{\theta_k^*}^{\varepsilon_k}(\mathbf{y}_k^*)$ we finally arrive at

$$\lim_{k\to\infty}\left(\varphi, \left[D\mathbf{g}\left(\mathbf{y}_k^*(\cdot)\right)^\top \mathscr{F}_{\theta_k^*}^{\varepsilon_k}\left(\mathbf{y}_k^*(\cdot)\right) - D\mathbf{g}(\bar{\mathbf{y}}(\cdot))^\top \bar{\mathbf{u}}(\cdot)\right]\mathbf{p}_k^*(\cdot)\right)_{L^2(Y_0, L^2(I;\mathbb{R}^n))} = 0.$$

By repeating this argument for the different terms appearing in the adjoint equation we get that $(\bar{y}, \bar{p}, \bar{u}) \coloneqq (\bar{\mathbf{y}}(y_0), \bar{\mathbf{p}}(y_0), \bar{\mathbf{u}}(y_0))$ satisfies

$$-\dot{\bar{p}} = D\mathbf{f}(\bar{y})^\top \bar{p} + \left[D\mathbf{g}(\bar{y})^\top \bar{u}\right]\bar{p} + \mathbf{Q}_1^\top \mathbf{Q}_1(\bar{y} - y_d), \;\; \bar{p}(T) = Q_2^\top Q_2\left(\bar{y}(T) - y_d^T\right)$$

for $\mathscr{L}$-a.e. $y_0 \in Y_0$. Applying Gronwall's inequality we deduce

$$\left\|\mathbf{p}_k(y_0) - \bar{\mathbf{p}}(y_0)\right\|_{W_T} \leq C\left(\left\|\mathbf{y}_k^*(y_0) - \bar{\mathbf{y}}(y_0)\right\|_{W_T} + \|\mathscr{F}_{\theta_k^*}^{\varepsilon_k}\left(\mathbf{y}_k^*(y_0)\right) - \bar{\mathbf{u}}(y_0)\|_{L^2}\right)$$

for $\mathscr{L}$-a.e. $y_0 \in Y_0$ and $C > 0$ independent of y_0 and k. This yields $\mathbf{p}_k \to \bar{\mathbf{p}}$ strongly in $L^2(Y_0; W_T)$. Since the weakly convergent subsequence was chosen arbitrarily in the beginning, this finishes the proof. □

Remark 20. If $\mathbf{g}(y(t)) = B \in \mathbb{R}^{m\times n}$ then the statement of the previous theorem also holds without constraints on the control (i.e. for $\mathbf{U}_{ad} = L^2(Y_0; L^2(I;\mathbb{R}^m))$). In this particular case, the uniform boundedness of $\mathscr{F}^{\varepsilon_k}_{\theta^*_k}(\mathbf{y}^*_k)$ in $L^2(Y_0;L^2(I;\mathbb{R}^m))$ follows from

$$\frac{\beta}{2}\left\|\mathscr{F}^{\varepsilon_k}_{\theta^*_k}\left(\mathbf{y}^*_k\right)\right\|^2_{L^2} \le c\int_{Y_0}\omega(y_0)J\left(\mathbf{y}^*_k,\mathscr{F}^{\varepsilon_k}_{\theta^*_k}\left(\mathbf{y}^*_k(y_0)\right)\right)\mathrm{d}\mathscr{L}(y_0) \le C,$$

see Theorem 17. Moreover the adjoint equation does no longer depend on the control. Repeating the arguments of the last proof yields the subsequential convergence of $(\mathbf{y}^*_k, \mathbf{p}^*_k, \mathscr{F}^{\varepsilon_k}_{\theta^*_k}(\mathbf{y}^*_k))$ towards an element

$$(\bar{\mathbf{y}},\bar{\mathbf{p}},\bar{\mathbf{u}}) \in L^\infty(Y_0;W_T)^2 \times L^2\left(Y_0;L^2\left(I;\mathbb{R}^m\right)\right)$$

such that $(\bar{y},\bar{p},\bar{u}) := (\bar{\mathbf{y}}(y_0),\bar{\mathbf{p}}(y_0),\bar{\mathbf{u}}(y_0))$ satisfy the system of state and adjoint equations as well as $(\bar{y},\bar{u}) \in \operatorname{argmin}\,(P_{y_0})$ for $\mathscr{L}$-a.e. $y_0 \in Y_0$. Then it only remains to argue the additional regularity $\mathbf{u} \in L^\infty(Y_0;L^2(Y_0;W_T))$. This is, however, a direct consequence of the first order necessary optimality condition $\bar{\mathbf{u}} = (-1/\beta)B^\top\bar{p}$ for (P_{y_0}), see Proposition 3.

We point out that the statement of Theorem 19 holds independently of the values of the penalty parameters γ_1,γ_2. If $\gamma_1,\gamma_2 > 0$ then we additionally obtain the following convergence results for the approximate value function $\mathscr{V}^{\varepsilon}_{\theta^*_k}$ and its derivative $\partial_y\mathscr{V}^{\varepsilon}_{\theta^*_k}$ along optimal state trajectories.

Proposition 21. *Let the prerequisites of Theorem 17 hold and let $(\mathbf{y}^*_k,\mathbf{p}^*_k,\theta^*_k)$ denote a sequence of minimizing triplets as described in Theorem 19. Assume that $(\mathbf{y}^*_k,\mathbf{p}^*_k,\mathscr{F}^{\varepsilon_k}_{\theta^*_k}(\mathbf{y}^*_k))$ converges to $(\bar{\mathbf{y}},\bar{\mathbf{p}},\bar{\mathbf{u}})$ in*

$$L^2(Y_0;W_T)^2 \times L^2\left(Y_0;L^2\left(I;\mathbb{R}^m\right)\right) \quad \textit{and} \quad \gamma_1,\gamma_2 > 0.$$

Then we also have

$$\mathscr{V}^{\varepsilon_k}_{\theta^*_k}\left(\mathbf{y}^*_k\right) \to \mathscr{V}^*(\bar{\mathbf{y}}) \textit{ in } L^2\left(Y_0;L^2(I)\right),\ \partial_y\mathscr{V}^{\varepsilon_k}_{\theta^*_k}\left(\mathbf{y}^*_k\right) \to \bar{\mathbf{p}} \textit{ in } L^2\left(Y_0;L^2\left(I;\mathbb{R}^n\right)\right).$$

Proof. Due to the convergence of $\mathbf{y}^*_k \to \bar{\mathbf{y}}$ in $L^2(Y_0;L^2(I;\mathbb{R}^n))$ and $\mathscr{F}^{\varepsilon_k}_{\theta^*_k}(\mathbf{y}^*_k) \to \bar{\mathbf{u}}$ in $L^2(Y_0;L^2(I;\mathbb{R}^m))$, we conclude that

$$J_\bullet\left(\mathbf{y}^*_k,\mathscr{F}^{\varepsilon_k}_{\theta^*_k}\left(\mathbf{y}^*_k\right)\right) \to J_\bullet(\bar{\mathbf{y}},\bar{\mathbf{u}}) = \mathscr{V}^*(\bar{\mathbf{y}}) \text{ in } L^2\left(Y_0;L^2(I)\right).$$

Together with

$$\lim_{k\to\infty}\int_{Y_0}\omega(y_0)\left\|V^{\varepsilon_k}_{\theta^*_k}\left(\mathbf{y}^*_k(y_0)\right) - J_\bullet\left(\mathbf{y}^*_k(y_0),\mathscr{F}^{\varepsilon_k}\left(\mathbf{y}^*_k(y_0)\right)\right)\right\|^2_{L^2}\,\mathrm{d}\mathscr{L}(y_0) = 0,$$

see Theorem 17, we arrive at $\mathscr{V}^{\varepsilon_k}_{\theta^*_k}(\mathbf{y}^*_k) \to \mathscr{V}^*(\bar{\mathbf{y}})$ in $L^2(Y_0;L^2(I))$. The statement on the convergence of $\partial_y\mathscr{V}^{\varepsilon_k}_{\theta^*_k}(\mathbf{y}^*_k)$ follows similarly from the strong convergence of $\mathbf{p}_k$. □

8. Learning from a finite training set

We turn to analysing a discrete version of $(\mathscr{P}_\varepsilon)$. In this case we can proceed without the state-space constraint $\mathbf{y} \in \mathbf{Y}_{ad}$ provided certain growth bounds on $\mathbf{f}$ and $\mathbf{g}$ are satisfied. The numerical realization of $(\mathscr{P}_\varepsilon)$ will always rely on such a discrete approximation. Henceforth we fix a finite ensemble of initial conditions $\{y^i_0 : i = 1,\dots,N\} \subset Y_0$. For positive weights ω_i, $i = 1,\dots,N$, and $\varepsilon > 0$ we consider

$$\inf_{y_i,p_i\in W_T,\theta\in\mathscr{R}_\varepsilon}\left[\sum_{i=1}^N \omega_i J_\varepsilon(y_i,p_i,\theta) + \frac{\gamma_\varepsilon}{2}\|\theta\|^2_{\mathscr{R}_\varepsilon}\right] \tag{$\mathscr{P}^N_\varepsilon$}$$

subject to

$$\begin{aligned}
&\dot{y}_i = \mathbf{f}(y_i) + \mathbf{g}(y_i)\mathscr{F}^{\varepsilon}_{\theta}(y_i),\ y_i(0) = y_0^i \\
&-\dot{p}_i = D\mathbf{f}(y_i)^\top p_i + \left[D\mathbf{g}(y_i)^\top \mathscr{F}^{\varepsilon}_{\theta}(y_i)\right] p_i + \mathbf{Q}_1^\top \mathbf{Q}_1 (y_i - y_d),\ p_i(T) = Q_2^\top Q_2 \left(y_i(T) - y_d^T\right).
\end{aligned}$$

Throughout this section, Assumptions 1 and 4 are supposed to hold. Further ε is supposed to be sufficiently small so that the set of admissible solutions for $(\mathscr{P}^N_\varepsilon)$ is nonempty, compare Theorem 7. It will be convenient to introduce $\mathbf{y} = \mathrm{col}(y_1, \dots, y_N)$, and $\mathbf{p} = \mathrm{col}(p_1, \dots, p_N)$, which replace the ensemble states and costates from the previous sections.

Proposition 22. *Let $\varepsilon > 0$ be sufficiently small and let $(\mathbf{y}^k, \mathbf{p}^k, \theta_k) \in W_T^{2N} \times \mathscr{R}_\varepsilon$ denote an infimizing sequence for $(\mathscr{P}^N_\varepsilon)$. If $\max_i \|y_i^k\|_{L^\infty(I;\mathbb{R}^n)} \le M_\infty$ for some $M_\infty > 0$ independent of $k \in \mathbb{N}$, then Problem $(\mathscr{P}^N_\varepsilon)$ admits at least one minimizer $(\mathbf{y}^*, \mathbf{p}^*, \theta^*)$.*

Proof. Since by assumption $(\mathbf{y}^k, \mathbf{p}^k, \theta_k)$ is an infimizing sequence for $(\mathscr{P}^N_\varepsilon)$ and since $\beta > 0$ we have

$$\max_i \left\|\mathbf{Q_1} y_i^k\right\|^2_{L^2} + \max_i \left\|\mathscr{F}^{\varepsilon}_{\theta_k}\left(y_i^k\right)\right\|^2_{L^2} \le C_N \tag{44}$$

for some $C_N > 0$ depending on N. Moreover there holds

$$\left\|\dot{y}_i^k\right\|_{L^2} \le \left\|\mathbf{f}\left(y_i^k\right)\right\|_{L^2} + \left\|\mathbf{g}(y_i)\mathscr{F}^{\varepsilon}_{\theta_k}\left(y_i^k\right)\right\|_{L^2} \le C(\mathbf{f},\mathbf{g}) M_\infty (1 + C_N)$$

using the uniform L^∞ and L^2 boundedness of y_i^k and $\mathscr{F}^{\varepsilon}_{\theta}(y_i^k)$, respectively. Thus we also have $\|y_k^i\|_{W_T} \le \widehat{C}_N$ for all $k \in \mathbb{N}$, for some $\widehat{C}_N > 0$ which depends on N but not on k and i. The proof can now be completed by the same steps as Theorem 16. □

Remark 23. The L^∞-boundedness of the minimizing sequence y_i^k in Proposition 22 can be be ensured by additional assumptions on the dynamics of the problem. These include:

- Add an additional state constraint $\|y_i\|_{L^\infty} \le \widehat{M}$ to $(\mathscr{P}^N_\varepsilon)$.
- Assume that there are $a_1, a_2, a_3 > 0$ such that

  $$|f(x)| \le a_1 + a_2|x| + a_3|x|^2,\ \|g(x)\| \le a_1 + a_2|x| \quad \forall\, x \in \mathbb{R}^n,$$

  and that Q_1 is positive definite. Then by (44) the family $\{y_i^k\}$ is uniformly w.r.t. $i \in \{1, \dots, n\}$ and $k = 1, \dots$ bounded in $L^2(I;\mathbb{R}^n)$. Further we can readily verify that

  $$\begin{aligned}
  \left\|\dot{y}_i^k\right\|_{L^1} &\le \left\|\mathbf{f}\left(y_i^k\right)\right\|_{L^1} + \left\|\mathbf{g}\left(y_i^k\right)\right\|_{L^1} \\
  &\le 2a_1 T + a_2 \left\|y_i^k\right\|_{L^1} + a_3 \left\|y_i^k\right\|^2_{L^2} + a_2 \left\|y_i^k\right\|_{L^2} \left\|\mathscr{F}^{\varepsilon}_{\theta_k}\left(y_i^k\right)\right\|^2_{L^2} \le M_N
  \end{aligned}$$

  for an N-dependent bound $M_N > 0$. Here we made use of the L^2-boundedness of y_i^k and $\mathscr{F}^{\varepsilon}_{\theta_k}(y_i^k)$ which follows from (44) in the proof of Proposition 22, and the assumption that $Q_1 > 0$. Consequently y_i^k is uniformly bounded in $W^{1,1}(I;\mathbb{R}^n)$ and thus also in $L^\infty(I;\mathbb{R}^n)$.
- Assume that $f(x) = Ax - h(x)$ where $A \in \mathbb{R}^{n\times n}$ and h is monotone i.e. $(x, h(x))_{\mathbb{R}^n} \ge 0$ for all $x \in \mathbb{R}^n$. Moreover assume that Q_1 is positive definite and that

  $$\|g(x)\| \le a_1 + a_2|x| \quad \forall\, x \in \mathbb{R}^n.$$

  In this case, testing the equation satisfied by y_i with y_i, and a Gronwall argument yields

  $$\left|y_i^k(t)\right|^2 \le C_N \left(\left|y_0^i\right|^2 + \left\|y_i^k\right\|^2_{L^2} + \left\|\mathscr{F}^{\varepsilon}_{\theta_k}\left(y_i^k\right)\right\|^2_{L^2}\right)$$

  for some N-dependent $C_N > 0$ and all $t \in I$. Thus, the uniform boundedness of y_i^k in $L^\infty(I;\mathbb{R}^n)$ follows again from the L^2-estimates on y_i^k and $\mathscr{F}^{\varepsilon}_{\theta_k}(y_i^k)$ in (44).

The convergence result as $\varepsilon \to 0^+$ of Theorem 19 can be transferred to the finite training set setting as well.

Proposition 24. *Let the regularisation parameters satisfy $\gamma_\varepsilon \|\theta_\varepsilon\|^2_{\mathscr{R}_\varepsilon} = O(\varepsilon)$. Further let $\varepsilon_k > 0$ be a positive null sequence such that for each $k \in \mathbb{N}$ there exists a solution $(\mathbf{y}^k, \mathbf{p}^k, \theta_k) \in W_T^{2N} \times \mathscr{R}_\varepsilon$ to $(\mathscr{P}^N_{\varepsilon_k})$. If there is $M_\infty > 0$ with $\max_i \|y_i^k\|_{L^\infty} \le M_\infty$ for all $k \in \mathbb{N}$, then $(\mathbf{y}^k, \mathbf{p}^k, \mathscr{F}^{\varepsilon_k}_{\theta_k}(\mathbf{y}^k))$ admits at least one strong accumulation point $(\bar{\mathbf{y}}, \bar{\mathbf{p}}, \bar{\mathbf{u}})$ in $W_T^{2N} \times L^2(I;\mathbb{R}^m)^N$. Each such point satisfies*

$$(\bar{y}_i, \bar{u}_i) \in \operatorname{argmin}\left(P_\beta^{y_0^i}\right), \quad i = 1, \dots, N,$$

as well as

$$\begin{aligned} \dot{\bar{y}}_i &= \mathbf{f}(\bar{y}_i) + \mathbf{g}(\bar{y}_i)\bar{u}_i, \ \bar{y}_i(0) = y_0^i \\ -\dot{\bar{p}}_i &= D\mathbf{f}(\bar{y}_i)^\top \bar{p}_i + \left[D\mathbf{g}(\bar{y}_i)^\top \bar{u}_i\right] \bar{p}_i + \mathbf{Q}_1^\top \mathbf{Q}_1(\bar{y}_i - y_d), \ \bar{p}_i(T) = Q_2^\top Q_2\left(\bar{y}_i(T) - y_d^T\right). \end{aligned}$$

Proof. For every ε_k, with k sufficiently large, denote by $\theta_{\varepsilon_k} \in \mathscr{R}_{\varepsilon_k}$ the corresponding parameters from Assumption 4, by $\mathbf{y}_{\varepsilon_k}$ the associated ensemble solution, see Theorem 7, and by $\mathbf{p}_{\varepsilon_k}$ the adjoint states. For abbreviation we set $y_i^{\varepsilon_k} := \mathbf{y}_{\varepsilon_k}(y_0^i)$ and $p_i^{\varepsilon_k} := \mathbf{p}_{\varepsilon_k}(y_0^i)$. Then, by optimality, we have

$$\sum_{i=1}^N \omega_i J\left(y_i^k, \mathscr{F}^{\varepsilon_k}_{\theta_k}\left(y_i^k\right)\right) \le \sum_{i=1}^N \omega_i J_\varepsilon\left(y_i^{\varepsilon_k}, p_i^{\varepsilon_k}, \theta_{\varepsilon_k}\right) + \frac{\gamma_{\varepsilon_k}}{2}\|\theta_{\varepsilon_k}\|^2_{\mathscr{R}_\varepsilon}. \tag{45}$$

As in the proof of Theorem 17 we see that the righthandside of this inequality converges to $\sum_{i=1}^N \omega_i V^*(0, y_0^i)$ as $k \to +\infty$. Thus it is bounded independently of $k \in \mathbb{N}$. Similarly to Proposition 22 we then conclude the existence of $C_N > 0$ depending on N, but not on k, such that

$$\max_i \left\|\mathbf{Q}_1 y_i^k\right\|^2_{L^2} + \max_i \left\|\mathscr{F}^\varepsilon_{\theta_k}\left(y_i^k\right)\right\|^2_{L^2} \le C_N.$$

Utilizing the state equation this can be improved to a k-independent bound on the W_T-norm of y_i^k. By a Gronwall-type argument the same can be shown for the adjoint states p_i^k. Now fix an arbitrary index $i \in \{1, \dots, N\}$. Summarizing the previous observations we get the uniform boundedness of $(y_i^k, p_i^k, \mathscr{F}^\varepsilon_{\theta_k}(y_i^k))$ in $W_T^2 \times L^2(I;\mathbb{R}^m)$ w.r.t. k, for each $i = 1, \dots, N$. Each of its weak accumulation points $(\bar{y}_i, \bar{p}_i, \bar{u}_i) \in W_T^2 \times L^2(I;\mathbb{R}^m)$ satisfies

$$\dot{\bar{y}} = \mathbf{f}(\bar{y}) + \mathbf{g}(\bar{y})\bar{u}, \ \bar{y}(0) = y_0.$$

From this we conclude that

$$0 \le \sum_{i=1}^N \omega_i V^*\left(0, y_0^i\right) \le \sum_{i=1}^N \omega_i J(\bar{y}_i, \bar{u}_i) \le \varliminf_{k\to\infty} \sum_{i=1}^N \omega_i J\left(y_i^k, \mathscr{F}^{\varepsilon_k}_{\theta_k}\left(y_i^k\right)\right) \le \sum_{i=1}^N \omega_i V^*\left(0, y_0^i\right),$$

Since the second and third of the above inequalities also hold for each summand we conclude that $\lim_{k\to\infty} J(y_i^k, \mathscr{F}^{\varepsilon_k}_{\theta_k}(y_i^k)) \to J(\bar{y}_i, \bar{u}_i)$ as well as $J(\bar{y}_i, \bar{u}_i) = V^*(0, y_0^i)$. Hence

$$(\bar{y}_i, \bar{u}_i) \in \operatorname{argmin}\left(P_\beta^{y_0^i}\right).$$

The proof can now be concluded with minor adaptations to the proof of Theorem 19. □

A result analogous to that of Proposition 21 can also be obtained for Problem $(\mathscr{P}^N_\varepsilon)$. For the sake of brevity we do not present the details.

8.1. *The reduced objective functional*

In order to compute a solution to ($\mathscr{P}_\varepsilon^N$) we will rely on gradient-based optimization methods. For this purpose we introduce a *reduced objective functional* by eliminating the state and adjoint equations in ($\mathscr{P}_\varepsilon^N$). Subsequently, we characterize the derivative of the reduced functional by means of adjoint techniques. To simplify the presentation we fix an arbitrary index $i \in \{1,\dots,N\}$ in the following. Moreover, for abbreviation, we define the mapping

$$\mathbf{A}\colon W_T \times \mathscr{R}_\varepsilon \to \mathscr{B}\left(W_T; L^2\left(I;\mathbb{R}^n\right)\right),\ A(y,\theta) = D\mathbf{f}(y)^\top + \left[D\mathbf{g}(y)^\top \mathscr{F}_\theta^\varepsilon(y)\right].$$

Using this notation, the adjoint equation in ($\mathscr{P}_\varepsilon^N$) can be expressed compactly as

$$-\dot{p}_i = \mathbf{A}(y_i,\theta)p_i + \mathbf{Q}_1^\top \mathbf{Q}_1(y_i - y_d),\ p_i(T) = Q_2^\top Q_2\left(y_i(T) - y_d^T\right).$$

First, we argue the existence of *parameter-to-state operators* for the adjoint and the state equation.

Lemma 25. *Define* $G_i\colon W_T \times W_T \times \mathscr{R}_\varepsilon \to L^2(I;\mathbb{R}^n) \times L^2(I;\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^n$ *by*

$$G_i(y,p,\theta) = \begin{pmatrix} \dot{y} - \mathbf{f}(y) - \mathbf{g}(y)\mathscr{F}_\theta^\varepsilon(y) \\ -\dot{p} - A(y,\theta)p - \mathbf{Q}_1^\top \mathbf{Q}_1(y - y_d) \\ y(0) - y_0^i \\ p(T) - Q_2^\top Q_2\left(y(T) - y_d^T\right) \end{pmatrix}.$$

Let $(\widetilde{y},\widetilde{p},\widetilde{\theta}) \in W_T \times W_T \times \mathscr{R}_\varepsilon$ *satisfy* $G(\widetilde{y},\widetilde{p},\widetilde{\theta}) = 0$. *Then there exists a neighbourhood* $\mathscr{N}_i(\widetilde{y}) \times \mathscr{N}_i(\widetilde{p}) \times \mathscr{N}_i(\widetilde{\theta})$ *as well as* $\mathscr{C}^1$*-mappings* $Y_i\colon \mathscr{N}_i(\widetilde{\theta}) \to \mathscr{N}_i(\widetilde{y}) \subset W_T$, $P_i\colon \mathscr{N}_i(\widetilde{\theta}) \to \mathscr{N}_i(\widetilde{p}) \subset W_T$ *such that*

$$G_i\left(Y_i(\theta), P_i(\theta), \theta\right) = 0 \quad \forall\, \theta \in \mathscr{N}(\widetilde{\theta}).$$

Given $y_i := Y_i(\theta)$ *and* $p_i := P_i(\theta)$, *the Fréchet derivatives of* Y_i *and* P_i *at* $\theta \in \mathscr{N}_i(\widetilde{\theta})$, *in direction* $\delta\theta \in \mathscr{R}_\varepsilon$, *denoted by* $\delta Y_i := Y_i'(\theta)(\delta\theta)$, $\delta P_i := P_i'(\theta)(\delta\theta)$ *satisfy*

$$\begin{aligned} &\delta\dot{Y}_i - \mathbf{A}(y_i,\theta)^\top \delta Y_i - \mathbf{g}(y_i) D_y \mathscr{F}_\theta^\varepsilon(y_i)\delta Y \\ &\qquad = \mathbf{g}(y_i) D_\theta \mathscr{F}_\theta^\varepsilon(y_i)\delta\theta, -\delta\dot{P}_i - \mathbf{A}(y_i,\theta)\delta P_i \\ &\qquad = \left[D_y \mathbf{A}(y_i,\theta)\delta Y_i\right] p_i + \mathbf{Q}_1\mathbf{Q}_1 \delta Y_i + \left[\partial_\theta \mathbf{A}(y_i,\theta)\delta\theta\right] p_i, \delta Y_i(0) \\ &\qquad = 0,\ \delta P_i(T) = Q_2^\top Q_2 \delta Y_i(T). \end{aligned}$$

Proof. This is a direct consequence of the implicit function theorem applied to G noting that the directional derivatives satisfy

$$\left(\partial_y G_i(y,p,\theta)\ \partial_p G_i(y,p,\theta)\right)\begin{pmatrix}\delta Y \\ \delta P\end{pmatrix} = -\partial_\theta G_i(y,p,\theta)\delta\theta.$$

□

Now consider an admissible point $(\widetilde{\mathbf{y}},\widetilde{\mathbf{p}},\theta) \in W_T^{2N} \times \mathscr{R}_\varepsilon$ for ($\mathscr{P}_\varepsilon^N$). For every $i = 1,\dots,N$, let $\mathscr{N}_i(\widetilde{\theta})$ and Y_i, P_i denote the corresponding neighbourhoods and operators from Lemma 25. Setting $\mathscr{N}(\widetilde{\theta}) = \bigcap_{i=1}^N \mathscr{N}_i(\widetilde{\theta})$ define the reduced objective functional

$$\mathscr{J}_N\colon \mathscr{N}(\widetilde{\theta}) \to [0,+\infty),\ \mathscr{J}_N(\theta) = \sum_{i=i}^N \omega_i J_\varepsilon\left(Y_i(\theta), P_i(\theta), \theta\right) + \frac{\gamma_\varepsilon}{2}\|\theta\|_{\mathscr{R}_\varepsilon}^2, \tag{46}$$

and set

$$\Phi_i(t) = \int_0^t \left(V_\theta^\varepsilon(s, y_i(s)) - J_s(y_i,u)\right)\,\mathrm{d}s.$$

Proposition 26. *The functional $\mathscr{J}_N$ from* (46) *is at least of class $\mathscr{C}^1$ on $\mathscr{N}(\widetilde{\theta})$. Given $\theta \in \mathscr{N}(\widetilde{\theta})$, set $y_i := Y_i(\theta)$, $p_i := P_i(\theta)$ as well as $\delta Y_i := Y_i'(\theta)(\delta\theta)$, $\delta P_i := P_i'(\theta)(\delta\theta)$. The directional derivative of $\mathscr{J}_N$ at θ in the direction of $\delta\theta \in \mathscr{R}_\varepsilon$ is given by*

$$\mathscr{J}_N'(\theta)(\delta\theta) = \sum_{i=1}^N \omega_i \left((\widehat{y}_i, \delta Y_i)_{L^2} + \left(\widehat{y}_i^T, \delta Y_i(T)\right)_{\mathbb{R}^n} + \left(\widehat{p}_i, \delta P_i\right)_{L^2} + \left(\widehat{\theta}_i, \delta\theta\right)_{\mathscr{R}_\varepsilon} \right) + \gamma_\varepsilon(\theta, \delta\theta)_{\mathscr{R}_\varepsilon}$$

with

$$\begin{aligned}\widehat{y}_i = (1-\gamma_1\Phi_i)\mathbf{Q}_1\mathbf{Q}_1(y_i - y_d) + \beta(1-\gamma_1\Phi_i)D_y\mathscr{F}_\theta^\varepsilon(y_i)^\top\mathscr{F}_\theta^\varepsilon(y_i)\\ + \gamma_1\left(\mathscr{V}_\theta^\varepsilon(t,y_i) - J_\bullet\left(y_i, \mathscr{F}_\theta^\varepsilon(y_i)\right)\right)\partial_y\mathscr{V}_\theta^\varepsilon(y_i) + \gamma_2 D_{yy}\mathscr{V}_\theta^\varepsilon(y_i)\left(\partial_y\mathscr{V}_\theta^\varepsilon(y_i) - p_i\right),\end{aligned}$$

and

$$\widehat{y}_i^T = \alpha\left(1-\gamma_1\Phi_i(0)\right)Q_2Q_2\left(y_i(T) - y_d^T\right),$$

as well as

$$\widehat{p}_i = \gamma_2\left(p_i - \partial_y\mathscr{V}_\theta^\varepsilon(y_i)\right),$$

and

$$\begin{aligned}\widehat{\theta}_i = \gamma_1 \int_0^T D_\theta V_\theta^\varepsilon(t,y_i(t))^\top\left(V_\theta^\varepsilon(t,y_i(t)) - J_t\left(y_i, \mathscr{F}_\theta^\varepsilon(y_i)\right)\right)\,\mathrm{d}t\\ + \int_0^T \left[\beta\left(1-\gamma_1\Phi_i(t)\right)D_\theta F_\theta^\varepsilon(t,y_i(t))^\top F_\theta^\varepsilon(t,y_i(t)) + \gamma_2 D_{y\theta}V_\theta^\varepsilon(t,y_i(t))^\top\left(\partial_y V_\theta^\varepsilon(t,y_i(t)) - p_i(t)\right)\right]\,\mathrm{d}t.\end{aligned}$$

Proof. The regularity of $\mathscr{J}_N$ follows immediately from Lemma 25 and the chain rule. In order to compute the directional derivative we abbreviate

$$\begin{aligned}F_1(y,u,\theta) &= \frac{\gamma_1}{2}\int_0^T \left|V_\theta^\varepsilon(t,y(t)) - J_t(y,u)\right|^2\,\mathrm{d}t,\\ F_2(y,p,\theta) &= \frac{\gamma_2}{2}\int_0^T \left|\partial_y V_\theta^\varepsilon(t,y(t)) - p(t)\right|^2\,\mathrm{d}t\end{aligned}$$

in the following. Thus we have

$$\begin{aligned}J_\varepsilon(Y_i(\theta), P_i(\theta), \theta) &= J\left(Y_i(\theta), \mathscr{F}_\theta^\varepsilon(Y_i(\theta))\right) + F_1\left(Y_i(\theta), \mathscr{F}_\theta^\varepsilon(Y_i(\theta)), \theta\right) + F_2\left(Y_i(\theta), P_i(\theta), \theta\right)\\ &= G_1(\theta) + G_2(\theta) + G_3(\theta).\end{aligned}$$

We readily verify

$$\begin{aligned}G_1'(\theta)(\delta\theta) = \left(\mathbf{Q}_1\mathbf{Q}_1(y_i - y_d), \delta Y_i\right)_{L^2} + \beta\left(D_y\mathscr{F}_\theta^\varepsilon(y_i)^\top\mathscr{F}_\theta^\varepsilon(y_i), \delta Y_i\right)_{L^2}\\ + \beta\left(D_\theta\mathscr{F}_\theta^\varepsilon(y_i)^\top\mathscr{F}_\theta^\varepsilon(y_i), \delta\theta\right)_{\mathscr{R}_\varepsilon} + \left(Q_2Q_2\left(y_i(T) - y_d^T\right), \delta Y_i(T)\right)_{\mathbb{R}^n}.\end{aligned}$$

Recalling the definition of Φ_i we get

$$G_2'(\theta)(\delta\theta) = \gamma_1\left(E_1 + E_2 + E_3 + E_4\right),$$

where

$$\begin{aligned}E_1 &= \left(\left(\mathscr{V}_\theta^\varepsilon(y_i) - J_\bullet\left(y_i, \mathscr{F}_\theta^\varepsilon(y_i)\right)\right)\partial_y\mathscr{V}_\theta^\varepsilon(y_i), \delta Y_i\right)_{L^2} + \left(D_\theta\mathscr{V}_\theta^\varepsilon(y_i)^\top\left(\mathscr{V}_\theta^\varepsilon(y_i) - J_\bullet\left(y_i, \mathscr{F}_\theta^\varepsilon(y_i)\right)\right), \delta\theta\right)_{\mathscr{R}_\varepsilon}\\ &= \left(\left(\mathscr{V}_\theta^\varepsilon(y_i) - J_\bullet\left(y_i, \mathscr{F}_\theta^\varepsilon(y_i)\right)\right)\partial_y\mathscr{V}_\theta^\varepsilon(t,y_i), \delta Y_i\right)_{L^2}\\ &\quad + \left(\int_0^T D_\theta V_\theta^\varepsilon\left(t,y_i(t)\right)^\top\left(V_\theta^\varepsilon(t,y_i(t)) - J_t\left(y_i, \mathscr{F}_\theta^\varepsilon(y_i)\right)\right)\,\mathrm{d}t, \delta\theta\right)_{\mathscr{R}_\varepsilon},\end{aligned}$$

$$
\begin{aligned}
E_2 = &-\int_0^T \left(V_\theta^\varepsilon(t,y(t)) - J_t(y,u)\right) \\
&\left(\int_t^T \left(Q_1 Q_1 (y(s)-y_d(s)), \delta y(s)\right) \, \mathrm{d}s + \left(Q_2 Q_2 \left(y(T)-y_d^T\right), \delta y(T)\right)_{\mathbb{R}^n}\right) \mathrm{d}t \\
= &-\left(\Phi_i Q_1 Q_1 (y-y_d), \delta y\right)_{L^2} - \Phi_i(0) \left(Q_2 Q_2 (y(T)-y_d^T\right), \delta y(T))_{\mathbb{R}^n},
\end{aligned}
$$

as well as

$$
\begin{aligned}
E_3 &= -\int_0^T \left(V_\theta^\varepsilon(t,y(t)) - J_t(y,u)\right) \left(\beta \int_t^T \left(D_y F_\theta^\varepsilon(s,y_i(s))^\top F_\theta^\varepsilon(s,y_i(s)), \delta Y_i(s)\right)_{\mathbb{R}^n} \, \mathrm{d}s\right) \mathrm{d}t \\
&= -\beta \left(\Phi_i D_y \mathscr{F}_\theta^\varepsilon(y_i)^\top \mathscr{F}_\theta^\varepsilon(y_i), \delta Y_i\right)_{L^2},
\end{aligned}
$$

and

$$
\begin{aligned}
E_4 &= -\int_0^T \left(V_\theta^\varepsilon(t,y(t)) - J_t(y,u)\right) \left(\beta \int_t^T \left(D_\theta F_\theta^\varepsilon\left(s,y_i(s)\right)^\top F_\theta^\varepsilon(s,y_i(s)), \delta\theta\right)_{\mathscr{R}_\varepsilon} \mathrm{d}s\right) \mathrm{d}t \\
&= -\beta \left(\int_0^T \Phi_i(t) D_\theta F_\theta^\varepsilon(t,y_i(t))^\top F_\theta^\varepsilon(t,y_i(t)) \, \mathrm{d}t, \delta\theta\right)_{\mathscr{R}_\varepsilon},
\end{aligned}
$$

by means of partial integration. Finally we calculate

$$
\begin{aligned}
G_3'(\theta)(\delta\theta) = \gamma_2 \left(D_{yy} \mathcal{V}_\theta^\varepsilon(y_i) \left(\partial_y \mathcal{V}_\theta^\varepsilon(y_i) - p_i\right), \delta Y_i\right)_{L^2} - \gamma_2 \left(\partial_y \mathcal{V}_\theta^\varepsilon(y_i) - p_i, \delta P_i\right)_{L^2} \\
+ \gamma_2 \left(\int_0^T D_{y\theta} V_\theta^\varepsilon(t,y_i(t))^\top \left(\partial_y V_\theta^\varepsilon(t,y_i(t)) - p_i(t)\right) \mathrm{d}t, \delta\theta\right)_{\mathscr{R}_\varepsilon}.
\end{aligned}
$$

Summarizing the previous observations, we arrive at the claimed characterization. □

Applying a gradient method to $(\mathscr{P}_\varepsilon^N)$ requires the computation of the gradient $\nabla \mathscr{J}_N(\theta) \in \mathscr{R}_\varepsilon$ which satisfies

$$
\mathscr{J}_N'(\theta)(\delta\theta) = \left(\nabla \mathscr{J}_N(\theta), \delta\theta\right)_{\mathscr{R}_\varepsilon} \quad \forall\, \delta\theta \in \mathscr{R}_\varepsilon.
$$

This can be done by computing $\mathscr{J}_N'(\theta)(e_j)$ for the canonical basis $\{e_j\}_{j=1}^{N_\varepsilon} \subset \mathscr{R}_\varepsilon$. However, such reasoning leads to the necessity to solve $2\dim(\mathscr{R}_\varepsilon)N$ additional ODEs in order to compute the sensitivities $Y_i'(\theta)(e_j)$ and $P_i'(\theta)(e_j)$, respectively. Introducing suitable costate equations, this can be reduced to $2N$ additional equation solves.

Lemma 27. *Let $\widehat{y}_i, \widehat{y}_i^T, \widehat{p}_i$ as well as $\delta Y_i, \delta P_i$ be defined as in Proposition 26. Then there holds*

$$
\left(\widehat{y}_i, \delta Y_i\right)_{L^2} + \left(\widehat{y}_i^T, \delta Y_i(T)\right)_{\mathbb{R}^n} + \left(\widehat{p}_i, \delta P\right)_{L^2} = \left(D_\theta \mathscr{F}_\theta^\varepsilon(y_i)^\top \left(\mathbf{g}(y_i)^\top \zeta_i + \left[D\mathbf{g}(y_i)\kappa_i\right]^\top p_i\right), \delta\theta\right)_{\mathscr{R}_\varepsilon}
$$

where $\zeta_i, \kappa_i \in W_T$ satisfy

$$
\begin{aligned}
-\dot{\zeta}_i &= \mathbf{A}(y_i,\theta)\zeta + D_y \mathscr{F}_\theta^\varepsilon(y_i)^\top \mathbf{g}(y_i)^\top \zeta_i + \left[D_y \mathbf{A}(y_i,\theta)^\top p\right] \kappa_i + \mathbf{Q}_1^\top \mathbf{Q}_1 \kappa_i + \widehat{y}_i \\
\dot{\kappa}_i &= \mathbf{A}(y_i,\theta)^\top \kappa_i + \widehat{p}_i, \\
\zeta_i(T) &= Q_2^\top Q_2 \kappa(T) + \widehat{y}_i^T,\ \kappa_i(0) = 0.
\end{aligned}
$$

Proof. For the sake of readability, we drop the subscript i in the following. By partial integration and Lemma 25 we obtain

$$
\begin{aligned}
\left(\widehat{p}, \delta P\right)_{L^2} &= \left(\dot{\kappa} - \mathbf{A}(y,\theta)^\top \kappa, \delta P\right) = \left(-\dot{\delta P} - \mathbf{A}(y,\theta)\delta P, \kappa\right) + \left(Q_2^\top Q_2 \delta Y(T), \kappa(T)\right)_{\mathbb{R}^n} \\
&= \left(\left[D_y \mathbf{A}(y,\theta)\delta Y\right] p + \mathbf{Q}_1^\top \mathbf{Q}_1 \delta Y + \left[\partial_\theta \mathbf{A}(y,\theta)\delta\theta\right] p, \kappa\right) + \left(\delta Y(T), \zeta(T) - \widehat{y}_i^T\right)_{\mathbb{R}^n}
\end{aligned}
$$

and

$$\begin{aligned}
&(\widehat{y}_i,\delta Y)_{L^2}+\left(\widehat{y}_i^T,\delta Y(T)\right)_{\mathbb{R}^n}\\
&\quad=\left(-\dot{\zeta}-\mathbf{A}(y,\theta)\zeta-D_y\mathscr{F}_\theta^\varepsilon(y)^\top\mathbf{g}(y)^\top\zeta-\left[D_y\mathbf{A}(y,\theta)^\top p\right]\kappa-\mathbf{Q}_1^\top\mathbf{Q}_1\kappa,\delta Y\right)_{L^2}+\left(\widehat{y}_i^T,\delta Y(T)\right)_{\mathbb{R}^n}\\
&\quad=\left(\delta\dot{Y}-\mathbf{A}(y,\theta)^\top\delta Y-\mathbf{g}(y)D_y\mathscr{F}_\theta^\varepsilon(y)\delta Y,\zeta\right)_{L^2}\\
&\qquad-\left(\left[D_y\mathbf{A}(y,\theta)^\top p\right]\kappa+\mathbf{Q}_1^\top\mathbf{Q}_1\kappa,\delta Y\right)_{L^2}-(\delta Y(T),\zeta(T))_{\mathbb{R}^n}\\
&\quad=\left(\mathbf{g}(y)D_\theta\mathscr{F}_\theta^\varepsilon(y)\delta\theta,\zeta\right)_{L^2}-\left(\left[D_y\mathbf{A}(y,\theta)^\top p\right]\kappa+\mathbf{Q}_1^\top\mathbf{Q}_1\kappa,\delta Y\right)_{L^2}-\left(\delta Y(T),\zeta(T)-\widehat{y}_i^T\right)_{\mathbb{R}^n}.
\end{aligned}$$

Adding both equations finally yields

$$\begin{aligned}
\left(\widehat{y}_i,\delta Y\right)_{L^2}+\left(\widehat{y}_i^T,\delta Y(T)\right)_{\mathbb{R}^n}+\left(p_1,\delta P\right)_{L^2}&=\left(\mathbf{g}(y)D_\theta\mathscr{F}_\theta^\varepsilon(y)\delta\theta,\zeta\right)_{L^2}+\left(\left[\partial_\theta\mathbf{A}(y,\theta)\delta\theta\right]p,\kappa\right)_{L^2}\\
&=\left(D_\theta\mathscr{F}_\theta^\varepsilon(y)^\top\left(\mathbf{g}(y)^\top\zeta+[D\mathbf{g}(y)\kappa]^\top p\right),\delta\theta\right)_{\mathscr{R}_\varepsilon}
\end{aligned}$$

which ends the proof of Lemma 27. □

We arrive at the following characterization of the gradient $\nabla\mathscr{J}_N(\theta)$.

Theorem 28. *Let $y_i,p_i,\zeta_i,\kappa_i\in W_T$, $\widehat{\theta}_i\in\mathscr{R}_\varepsilon$ be defined as in Proposition 26 and Lemma 27. The gradient of $\mathscr{J}_N$ at θ is given by*

$$\nabla\mathscr{J}_N(\theta)=\sum_{i=1}^N\omega_i\left(D_\theta\mathscr{F}_\theta^\varepsilon(y_i)^\top\left(\mathbf{g}(y_i)^\top\zeta_i+\left[D\mathbf{g}(y_i)\kappa_i\right]^\top p_i\right)+\widehat{\theta}_i\right)+\gamma_\varepsilon\theta.$$

9. Numerical example

We finish this paper by applying the proposed learning approach to one particular instance of Problem (P_{y_0}). Setting $I=(0,T)$ and $\Omega=(0,2\pi)$, we consider the parabolic bilinear optimal control problem

$$\min_{\mathscr{Y}\in L^2(I\times\Omega),\,u\in L^2(I;\mathbb{R}^3)}\left[\frac{1}{2}\int_I\|\mathscr{Y}(t)-\mathscr{Y}_d(t)\|_{L^2(\Omega)}^2+\frac{\beta}{2}|u(t)|_{\mathbb{R}^3}^2\,\mathrm{d}t\right]+\frac{\alpha}{2}\|\mathscr{Y}(T)-\mathscr{Y}_d(T)\|_{L^2(\Omega)}^2$$

subject to

$$\partial_t\mathscr{Y}-\triangle\mathscr{Y}+\left(u_1\chi_1+u_2\chi_2+u_3\chi_3\right)\mathscr{Y}=0,\tag{47}$$

as well as

$$\mathscr{Y}(t,x)=0\quad\text{on } I\times\partial\Omega,\ \mathscr{Y}(0,x)=\mathscr{Y}_0(x)\quad\text{on }\Omega.$$

Here $\alpha>0,\beta>0$, and $\mathscr{Y}_d$ denotes a given desired state. The dynamics of this infinite-dimensional system can be influenced by choosing a time-dependent three-dimensional control input $u\in L^2(I;\mathbb{R}^3)$ which acts on the subdomains $\Omega_1=(0.5,1)$, $\Omega_2=(2,2.5)$ and $\Omega_3=(4,4.5)$, respectively. The associated characteristic functions are denoted by χ_i, $i=1,\dots,3$.

In order to fit this problem into the setting of the current manuscript, let $\{\lambda_i,\varphi_i\}\in\mathbb{R}_+\times L^2(\Omega)$ denote the first $n\in\mathbb{N}$ normalized eigenpairs of the Dirichlet Laplacian on Ω. Approximating the state dynamics $\mathscr{Y}$ as well as the desired state by

$$\mathscr{Y}(t,x)\approx\sum_{i=1}^n Y_i(t)\varphi_i(t),\ \mathscr{Y}_d(t,x)\approx\sum_{i=1}^n Y_d^i(t)\varphi_i(t),$$

we end up with

$$\min_{Y\in L^2(I;\mathbb{R}^{10}),\,u\in L^2(I;\mathbb{R}^3)}\left[\frac{1}{2}\int_I|Y(t)-Y_d(t)|^2+\frac{\beta}{2}|u(t)|_{\mathbb{R}^3}^2\,\mathrm{d}t+\frac{\alpha}{2}|Y(T)-Y_d(T)|_{\mathbb{R}^{10}}^2\right]\tag{48}$$

subject to

$$\dot{Y}(t) + AY(t) + \sum_{i=1}^{3} u_i M_i Y(t) = 0,\ Y(0) = Y_0.$$

where $(Y_0)_i = (\mathscr{Y}_0, \varphi_i)_{L^2}$, $i = 1, \ldots, n$, and the symmetric matrices $A, M_i \in \mathbb{R}^{n\times n}$ are given by

$$A_{jk} = \begin{cases} 0 & j \neq k \\ \lambda_j & \text{else} \end{cases},\ (M_i)_{jk} = \int_\Omega \phi_j \phi_k \chi_i(x)\,\mathrm{d}x, \quad i = 1,2,3,\ j,k = 1,\ldots,n.$$

Note that by choosing a spectral technique to approximate the infinite dimensional system (47) by a finite dimensional one, we have done justice to the fact that grid based techniques would rapidly lead to systems of dimensions which are challenging or even impossible. Special techniques, as for instance tensor train methods, would then become essential.

9.1. *Learning & validation setup*

In the following, we determine an approximate optimal feedback law for (48) by applying the learning approach detailed in Section 4. The parametrized model V_θ^ε for the value function is given by realizations of residual networks, as described in Section 5.1, with $L_\varepsilon = 2$ layers, $\operatorname{arch}(\theta) = (11, 60, 1)$ and activation function σ given by

$$\sigma(x) = \sin(x) + \cos(x).$$

This yields a total of 1440 trainable parameters. We emphasize that the architecture as well as the activation function were chosen based on numerical testing. In particular, the present tests should not be mistaken as a *quantitative* survey but as a *proof of concept* which highlights the potential of learned feedbacks for optimal control and puts a focus on the role played by the penalty parameters γ_1 and γ_2.

Given a fixed reference vector $\bar{Y}_0$, we randomly generate a set $\mathbf{y}_0$ of 130 initial conditions by sampling uniformly from the closure of $B_1(\bar{Y}_0)$, Subsequently, these are split into a training set $\mathbf{y}_0^t$ of $N = 30$ initial conditions, which is used in the learning problem $(\mathscr{P}_\varepsilon^N)$ together with uniform weights $w_j = 1/N$, and a validation set $\mathbf{y}_0^v = \mathbf{y}_0 \backslash \mathbf{y}_0^T$ which we later utilize to assess the performance of the obtained feedback.

In order to obtain a candidate for the optimal network parameters θ_ε^*, a Barzilai-Borwein method [23], is applied to the learning problems $(\mathscr{P}_\varepsilon^N)$, based on the reduced objective functional introduced in (46) as well as the characterization of its gradient in Theorem 28. For every $Y_0 \in \mathbf{y}_0^t$, this approach entails the computation of the state $Y := Y_\theta(Y_0)$ and the adjoint state $P := P_\theta(Y_0)$ which satisfy

$$\begin{aligned} \dot{Y}(t) + \left(A + \sum_{i=1}^{3} F_\theta^\varepsilon(t, Y(t))_i M_i\right) Y(t) = 0,\ Y(0) = Y_0 \\ -\dot{P}(t) + \left(A + \sum_{i=1}^{3} F_\theta^\varepsilon(t, Y(t))_i M_i\right) P(t) = Y(t) - Y_D(t),\ P(T) = Y(T) - Y_D(T) \end{aligned} \tag{49}$$

as well as the costates $K := K_\theta(Y_0)$ and $Z := Z_\theta(Y_0)$ with

$$\dot{K}(t) + \left(A + \sum_{i=1}^{3} F_\theta^\varepsilon(t, Y(t))_i M_i\right) K(t) = \widehat{P}(t)$$

and

$$-\dot{Z}(t)+\left(A+\sum_{i=1}^{3}F_{\theta}^{\varepsilon}(t,Y(t))_i M_i + D_y F_{\theta}^{\varepsilon}(t,Y(t))^{\top}\begin{pmatrix} Y_j(t)^{\top}M_1\\ Y_j(t)^{\top}M_2\\ Y_j(t)^{\top}M_3\end{pmatrix}\right)Z(t)$$
$$= -D_y F_{\theta}^{\varepsilon}(t,Y(t))^{\top}\begin{pmatrix} Y(t)^{\top}M_1\\ Y(t)^{\top}M_2\\ Y(t)^{\top}M_3\end{pmatrix}Z(t)+K(t)+\widehat{Y}(t)$$

equipped with the boundary conditions

$$K(0)=0,\quad Z(T)=\alpha K(T)+\widehat{Y}_j^T$$

where $\widehat{Y},\widehat{Y}^T$ and $\widehat{P}$ are defined in analogy to Proposition 26. Note that this system is not fully coupled, i.e. in practice, we first solve the nonlinear closed-loop equation using a Radau time-stepping scheme and then, successively treat the adjoint and costate equations by an implicit Euler method. This can be done in parallel for various initial conditions to achieve additional speed-up. Moreover, the adjoint state P and costate K only need to be computed if $\gamma_2>0$. The gradient of the reduced objective functional $\mathscr{J}_N$ in $(\mathscr{P}_{\varepsilon}^N)$ at an admissible θ is then obtained as

$$\frac{1}{30}\sum_{Y\in\mathbf{y}_0^t}\int_I\left(D_{\theta}F_{\theta}^{\varepsilon}(t,Y_{\theta}(Y_0)(t))^{\top}\left(B_Y^{\theta}(t)Z_{\theta}(Y_0)(t)+B_K^{\theta}(t)P_{\theta}(Y_0)(t)\right)\,\mathrm{d}t+\widehat{\theta}(Y_0)\right).$$

where we set

$$B_Y^{\theta}(t):=\begin{pmatrix} Y_{\theta}(Y_0)(t)^{\top}M_1\\ Y_{\theta}(Y_0)(t)^{\top}M_2\\ Y_{\theta}(Y_0)(t)^{\top}M_3\end{pmatrix},\quad B_K^{\theta}(t):=\begin{pmatrix} K_{\theta}(Y_0)(t)^{\top}M_1\\ K_{\theta}(Y_0)(t)^{\top}M_2\\ K_{\theta}(Y_0)(t)^{\top}M_3\end{pmatrix},$$

integration has to be understood componentwise and $\widehat{\theta}(Y_0)$ is as in Proposition 26.

Once the network is determined, we compute the state $Y_{\theta}(Y_0)$ and adjoint $P_{\theta}(Y_0)$ for every $Y_0\in\mathbf{y}_0$ from (49) and set $U_{\theta}(Y_0):=\mathscr{F}_{\hat{\theta}}^{\varepsilon}(Y_{\theta}(Y_0))$. Subsequently we determine a stationary point $(\bar{Y}(Y_0),\bar{U}(Y_0))$ of (48), $Y_0\in\mathbf{y}_0$, by applying a Barzilai–Borwein gradient method to its control-reduced formulation. The associated adjoint state is denoted by $\bar{P}(Y_0)$. At this point, it should be stressed that both, the open loop as well as the feedback learning problem, are nonconvex. As a consequence, we cannot ensure global optimality of the computed stationary points and, in particular, both methods might provide different results. For the present example, open loop and learned feedback controls are comparable. Moreover, for every $Y_0\in\mathbf{y}_0$, we have $J(\bar{Y}(Y_0),\bar{U}(Y_0))\geq J(Y_{\theta}(Y_0),U_{\theta}(Y_0))$. In order to assess the performance of open loop and feedback controls, let $Y_0^{ad}\subset\mathbf{y}_0$ be either $Y_0^{ad}=\mathbf{y}_0^t$ or $Y_0^{ad}=\mathbf{y}_0^v$ and consider the relative difference between the averaged objective functional values:

$$\mathrm{Err}_{\mathscr{J}}:=\frac{\sum_{Y_0\in Y_{ad}}J(Y_{\theta}(Y_0),U_{\theta}(Y_0))-\sum_{Y_0\in Y_{ad}}J\left(\bar{Y}(Y_0),\bar{U}(Y_0)\right)}{\sum_{Y_0\in Y_{ad}}J\left(\bar{Y}(Y_0),\bar{U}(Y_0)\right)}$$

as well as the associated normalized mean squared error of $J(Y_{\theta}(\cdot),U_{\theta}(\cdot))$:

$$\mathrm{Err}_J:=\frac{\sum_{Y_0\in Y_{ad}}\left(J(Y_{\theta}(Y_0),U_{\theta}(Y_0))-J\left(\bar{Y}(Y_0),\bar{U}(Y_0)\right)\right)^2}{\sum_{Y_0\in Y_{ad}}J\left(\bar{Y}(Y_0),\bar{U}(Y_0)\right)^2}.$$

The normalized mean-squared errors of the state, Err_Y, adjoint, Err_P, and of the control, Err_U, are defined analogously. Moreover, to quantify the influence of the penalty parameters γ_1 and γ_2, we define

$$\mathrm{Err}_V:=\frac{\sum_{Y_0\in Y_{ad}}\int_I\left|V_{\theta}^{\varepsilon}(t,Y_{\theta}(Y_0)(t))-J_t(Y_{\theta}(Y_0),U_{\theta}(Y_0)(t))\right|^2\,\mathrm{d}t}{\sum_{Y_0\in Y_{ad}}\int_I|J_t(Y_{\theta}(Y_0),U_{\theta}(Y_0)(t))|^2\,\mathrm{d}t}.$$

as well as

$$\mathrm{Err}_{\partial V} := \frac{\sum_{Y_0 \in Y_{ad}} \int_I \left|\partial_y V_\theta^\varepsilon(t, Y_\theta(Y_0)(t)) - P_\theta(Y_0)(t)\right|^2 \,\mathrm{d}t}{\sum_{Y_0 \in Y_{ad}} \int_I |P_\theta(Y_0)(t)|^2 \,\mathrm{d}t}.$$

For $Y_0^{ad} = Y_0^t$, these terms correspond to the relative sizes of the additional penalties in $(\mathscr{P}_\varepsilon^N)$. Finally, we also want to compare V_θ^ε with the optimal value function V^*. Of course, V^* can neither be given analytically nor can it be computed exactly. As a remedy, we recall that if V^* is sufficiently regular and $(\bar{Y}(Y_0), \bar{U}(Y_0))$ is a minimizing pair of (48) with adjoint state $\bar{P}(Y_0)$, we have

$$V^*\left(t, \bar{Y}(Y_0)(t)\right) = J_t\left(\bar{Y}(Y_0), \bar{U}(Y_0)\right) \quad \text{as well as} \quad \partial_y V^*\left(t, \bar{Y}(Y_0)(t)\right) = \bar{P}(Y_0)(t)$$

for all $t \in I$. As a consequence, setting

$$d\left(V^*, V_\theta^\varepsilon\right) = \frac{\sum_{Y_0 \in Y_{ad}} \int_0^T \left|V_\theta^\varepsilon\left(t, \bar{Y}(Y_0)(t)\right) - J_t\left(\bar{Y}(Y_0), \bar{U}(Y_0)\right)\right|^2 \,\mathrm{d}t}{\sum_{Y_0 \in Y_{ad}} \int_0^T \left|J_t\left(\bar{Y}(Y_0), \bar{U}(Y_0)\right)\right|^2 \,\mathrm{d}t}.$$

as well as

$$d\left(\partial V^*, \partial V_\theta^\varepsilon\right) = \frac{\sum_{Y_0 \in Y_{ad}} \int_0^T \left|\partial_y V_\theta^\varepsilon\left(t, \bar{Y}(Y_0)(t)\right) - \bar{P}(Y_0)(t)\right|^2 \,\mathrm{d}t}{\sum_{Y_0 \in Y_{ad}} \int_0^T \left|\bar{P}(Y_0)(t)\right|^2 \,\mathrm{d}t}.$$

provides a suitable “distance” for the comparison of V^* and V_ε^θ.

9.2. *Validation results*

As a concrete example, we set $T = 2$, $\beta = 0.01$, $\alpha = 0.25$ and $\mathscr{Y}_d(t,x) = x^2/10$, i.e., we try to steer the system towards a parabola. Note that there is no control input $u \in L^2(I;\mathbb{R}^3)$ such that the corresponding solution $\mathscr{Y}$ of the PDE (47) satisfies $\mathscr{Y}(t) = \mathscr{Y}_d$. The parabolic binlinear control problem is approximated using $n = 10$ eigenfunctions. All computations were carried out in Matlab 2019 on a notebook with 32 GB RAM and an Intel®Core™ i7-10870H CPU@2.20 GHz.

In order to compute an approximately optimal feedback law for this problem, we solve $(\mathscr{P}_\varepsilon^N)$ for various penalty parameter configurations $\gamma_1, \gamma_2 \in \{0, 0.1, 1\}$. The resulting normalized errors can be found in Table 1, for $Y_0^{ad} = \mathbf{y}_0^t$, and Table 2, for $Y_0^{ad} = \mathbf{y}_0^v$. Comparing their individual entries, we observe that there is (almost) no difference in performance between the training and the validation sets. This means that, while the utilized networks are rather simple and only comprise a small number of trainable parameters, the corresponding learned feedback controls generalize well to initial conditions which are not contained in the training set.

Indeed, on the one hand *all* computed networks provide feedback controls which perform similarly to their open loop counterparts. This is manifested in very small averaged errors for the objective functional, i.e. $\mathrm{Err}_{\mathscr{J}}$ and Err_J, the states and adjoint states, Err_Y and Err_P, as well as the controls, Err_U. These start to (slowly) deteriorate as γ_1 and/or γ_2 grow. However, cf. the explanation in Section 4, this is expected: for $\gamma_1 > 0$ and/or $\gamma_2 > 0$, the learned feedback has to strike a balance between minimizing $J(Y_\theta(\cdot), U_\theta(\cdot))$ and keeping the penalty terms small, hence the slightly larger error.

On the other hand, the picture looks different once we consider the errors associated to the approximation of the value function, i.e., Err_V, $\mathrm{Err}_{\partial V}$ as well as $d(\partial V^*, \partial V_\theta^\varepsilon)$ and $d(\partial V^*, \partial V_\theta^\varepsilon)$. Here $\gamma_1 > 0$ and/or $\gamma_2 > 0$ have a significant influence on $d(V^*, V_\theta^\varepsilon)$ and $d(\partial V^*, \partial V_\theta^\varepsilon)$ while the other normalized mean squared errors remain relatively small. Moreover, we have $\mathrm{Err}_V \approx d(V^*, V_\theta^\varepsilon)$ and $\mathrm{Err}_{\partial V} \approx d(\partial V^*, \partial V_\theta^\varepsilon)$ on the test as well as on the validation set. Hence, large values for these terms are a reliable indicator for structural differences between V_θ^ε and V^* and/or $\partial_y V_\theta^\varepsilon$ and $\partial_y V^*$, respectively.

Now, while $\gamma_1 = \gamma_2 = 0$ provides a very good approximation to the open loop optimal control, it performs the worst in terms of approximating the optimal value function and its derivative. This is related to two observations. First, in this case, the learning problem $(\mathscr{P}^N_\varepsilon)$ only depends on the derivative $\partial_y V^\theta_\varepsilon$ but *not* on the value function V^ε_θ. Since primitives are not unique, approximating V^* by V^ε_θ is unlikely. Second, due to the absence of V^ε_θ in the problem, some of the parameters in the model are not trainable. In fact, for $\gamma_1 = \gamma_2 = 0$, there holds $\partial_{W_{12}} \mathscr{J}_N(\theta) = 0$ for every admissible θ.

Once we increase γ_1 and γ_2, this is no longer the case. Hence, we observe rapid decrease for $d(V^*, V^\varepsilon_\theta)$ and $d(\partial V^*, \partial V^\varepsilon_\theta)$. Most remarkably, the improvement for both is, to some extend, already visible for $\gamma_1 > 0$ and $\gamma_2 = 0$. In this setting, applying the gradient method neither requires computing the adjoint state P nor the costate K which limits the cost of every gradient step to $2N = 60$ ODE solve. Quite the contrary, increasing $\gamma_2 > 0$ but keeping $\gamma_1 = 0$ fixed, there is *no* improvement for $d(V^*, V^\varepsilon_\theta)$. This further backs up our reasoning given for the case of $\gamma_1 = \gamma_2 = 0$.

Consequently, the computed results indicate that the best balance between finding an optimal control and approximating the value function is achieved by a careful choice of $\gamma_1, \gamma_2 > 0$. Moreover, they highlight two important points: First, the presented learning approach indeed allows to compute semiglobal optimal feedback laws F^ε_θ for higher dimensional problems and, thus, to some extent, alleviates the curse of dimensionality. Second, incorporating additional terms into the learning problem penalizing the violation of the dynamic programming principles (15), allows to compute a good approximation V^ε_θ of the optimal value function on the fly. As stated initially, the present example should be understood as a proof of concept and, following these first promising results, we believe that this approach to feedback learning deserves further investigations, both, from the theoretical and the numerical side. For example, it would be interesting to explore systematic ways of choosing the penalty parameters γ_1, γ_2. However, this goes beyond the scope of the current paper and is left for future work.

Table 1. Results on training set i.e. $Y^{ad}_0 = \mathbf{y}^t_0$.

Penalty	$\mathrm{Err}_{\mathscr{J}}$	Err_Y	Err_P	Err_U
$\gamma_1 = 0, \gamma_2 = 0$	0.15%	0.04%	0.12%	2.4%
$\gamma_1 = 0.1, \gamma_2 = 0.1$	0.36%	0.1%	0.24%	5.5%
$\gamma_1 = 0.1, \gamma_2 = 0$	0.29%	0.1%	0.85%	4.4%
$\gamma_1 = 1, \gamma_2 = 1$	0.64%	0.25%	1%	8.65%
$\gamma_1 = 0, \gamma_2 = 1$	0.1%	0.05%	0.26%	2.1%

Penalty	Err_J	Err_V	$\mathrm{Err}_{\partial V}$	$d(V^\varepsilon_\theta; V^*)$	$d(\partial_y V^\varepsilon_\theta; \partial_y V^*)$
$\gamma_1 = 0, \gamma_2 = 0$	0.0003%	79%	33%	78.8%	33.5%
$\gamma_1 = 0.1, \gamma_2 = 0.1$	0.001%	0.03%	7.4%	0.03%	7%
$\gamma_1 = 0.1, \gamma_2 = 0$	0.001%	0.02%	12.5%	0.02%	12.1%
$\gamma_1 = 1, \gamma_2 = 1$	0.005%	0.007%	4.5%	0.01%	3.5%
$\gamma_1 = 0, \gamma_2 = 1$	0.003%	88.8%	6.4%	88.5%	6.4%

Table 2. Results on validation set i.e. $Y_0^{ad} = \mathbf{y}_0^v$.

Penalty	$\mathrm{Err}_{\mathscr{J}}$	Err_Y	Err_P	Err_U
$\gamma_1 = 0, \gamma_2 = 0$	0.23%	0.06%	0.57%	4.42%
$\gamma_1 = 0.1, \gamma_2 = 0.1$	0.51%	0.15%	1.1%	9.2%
$\gamma_1 = 0.1, \gamma_2 = 0$	0.47%	0.16%	2.8%	8.7%
$\gamma_1 = 1, \gamma_2 = 1$	0.85%	0.35%	6%	13.5%
$\gamma_1 = 0, \gamma_2 = 1$	0.25%	0.1%	1.3%	4.8%

Penalty	Err_J	Err_V	$\mathrm{Err}_{\partial V}$	$d(V_\theta^\varepsilon; V^*)$	$d(\partial_y V_\theta^\varepsilon; \partial_y V^*)$
$\gamma_1 = 0, \gamma_2 = 0$	0.002%	78.7%	33.6%	78.6%	33.9%
$\gamma_1 = 0.1, \gamma_2 = 0.1$	0.007%	0.03%	8.9%	0.03%	7.9%
$\gamma_1 = 0.1, \gamma_2 = 0$	0.008%	0.02%	15.1%	0.02%	13.1%
$\gamma_1 = 1, \gamma_2 = 1$	0.02%	0.009%	9.8%	0.01%	4.2%
$\gamma_1 = 0, \gamma_2 = 1$	0.002%	88%	8.6%	88%	7.1%

Appendix A. Condition (24)

Here we address condition (24). Define $N_\varepsilon = \lceil \frac{2\widehat{M}}{\varepsilon} \rceil$, $\widetilde{M} := \varepsilon N_\varepsilon$ and introduce the equidistant grid $G = \{-\widetilde{M}, (1 - N_\varepsilon)\varepsilon, \dots, -\varepsilon, 0, \varepsilon, \dots, (N_\varepsilon - 1)\varepsilon, \widetilde{M}\}$. Next endow the hypercube $[-\widetilde{M}, \widetilde{M}]^{n+1}$ with the $(n+1)-$ dimensional product of the grid G. These grid points define $\{Q_i\}_{i=1}^{(2N_\varepsilon)^{n+1}}$ closed subhypercubes of dimension ε^n whose union covers $\bar{K} = [0, T] \times \bar{B}_{2\widehat{M}}(0)$.

We extend this $n+1$-dimensional grid by adding $k \geq \lceil \frac{1}{2}\sqrt{n} \rceil + 1$ layers (again all of dimension ε^n), to the surfaces of the preexisting grid, resulting in $\widetilde{N}_\varepsilon = (2N_\varepsilon + 2k)^{n+1}$ hypercubes whose union covers $[-\widetilde{M} - k\varepsilon, \widetilde{M} + k\varepsilon]^{n+1}$. The subhypercubes are ordered in such a manner that the interiors ones $\{Q_i\}_{i=1}^{(2N_\varepsilon + 2(k-1))^{n+1}}$ are assembled first and the ones with a boundary face

$$\{Q_i\}_{i=(2N_\varepsilon+2(k-1)+1)^{n+1}}^{(2N_\varepsilon+2k)^{n+1}}$$

come last. The set of indices corresponding to interior hypercubes are denoted by $\mathscr{I}$, those to boundary hypercubes by $\mathscr{F}$.

Next we introduce a staggered grid and place a node $x_i = (t_i, y_i)$ at the barycenter of each of the Q_i, $i = 1, \dots, (2N_\varepsilon + 2k)^{n+1}$. We shall use the standard mollifier of radius r_ε defined by

$$\psi(x) = \begin{cases} \exp\left(\frac{1}{|\frac{x}{r_\varepsilon}|^2 - 1}\right), & \text{for} |x| \leq r_\varepsilon \\ 0, & \text{for} |x| \leq r_\varepsilon, \end{cases}$$

where $r_\varepsilon = \varepsilon(\frac{1}{2}\sqrt{n} + .1)$. Note that by adding .1 in the previous expression the cube $[-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}]^n$ is contained in the interior of the support of psi. Finally we introduce $\psi_j(x) = \psi(x - x_j)$, for $j \in \mathscr{I} \cup \mathscr{F}$ and

$$\varphi_j = \frac{\psi_j}{\sum_{i \in \mathscr{I} \cup \mathscr{F}} \psi_i}, \text{ for } j \in \mathscr{I}.$$

Let us deduce the following properties:

(i) For each $j \in \mathscr{I} \cup \mathscr{F}$ we have $supp\,\psi_j = \bar{K}_j$ where $K_j = \{x : |x - x_j| < r_\varepsilon\}$.

(ii) By construction there exists $\mathfrak{m}$ such that $\operatorname{card}\{j : \psi_j(x) \neq 0\} \leq \mathfrak{m}, \quad \forall\ x \in \mathbb{R}^{n+1}$, and $\forall\ j \in \mathscr{I} \cup \mathscr{F}$, for each $\varepsilon \in (0, \varepsilon_0]$.

(iii) For each $j \in \mathscr{I}$ the denominator in the definition of φ_j is different from zero. Hence φ_j is well-defined with $supp\, \varphi_j = \bar{K}_j$ for $j \in \mathscr{I}$ and $\varphi_j : \mathbb{R}^{n+1} \to [0,1]$, and it is $\mathscr{C}^\infty$ smooth.

(iv) $\bar{K} \subset \bigcup_{j \in \mathscr{I}} Q_j \subset \bigcup_{j \in \mathscr{I}} K_j$.

(v) $\operatorname{card}\{j : \varphi_j(x) \neq 0\} \leq \mathfrak{m}, \quad \forall\ x \in \mathbb{R}^{n+1}$, and $\forall\ j \in \mathscr{I}$, for each $\varepsilon \in (0, \varepsilon_0]$.

(vi) Due to the choice of r_ε the functions ψ are uniformly bounded from below on Q_j for each j, independent of $\varepsilon \in (0, \varepsilon_0]$. Moreover due to the boundedness of ψ and by the definition of $\mathfrak{m}$, there exists $\nu > 0$ such that

$$\sum_{i \in \mathscr{I} \cup \mathscr{F}} \psi_i(x) \geq \nu, \forall\ x \in Q_j, \text{ with } j \in \mathscr{I} \cup \mathscr{F},$$

and thus in particular $\sum_{i \in \mathscr{I} \cup \mathscr{F}} \psi_i(x) \geq \nu, \forall\ x \in \bar{K}$.

(vii)

$$\operatorname{supp}\ \varphi_j \cap \bar{K} = \emptyset \ \ \forall\ j \in \mathscr{F}.$$

This is a consequence of the fact that for $j \in \mathscr{F}$ we have $dist(x_j, \partial([-\widetilde{M}, \widetilde{M}]^n)) = \varepsilon[(k-1) + \frac{1}{2}]$ and thus $dist(\partial K_j, \partial([-\widetilde{M}, \widetilde{M}]^n)) \leq \varepsilon[(k-1) + \frac{1}{2} - r_\varepsilon] = k - \frac{1}{2}(1+\sqrt{n}) - .1 > \varepsilon(k - 1 - \frac{1}{2}\sqrt{n}) > 0$.

(viii) $\sum_{i \in \mathscr{I}} \varphi_i = 1, \forall\ x \in \bar{K}$. This is a consequence of (vii) and the definition of φ_j.

(ix) $\|D^j \varphi_i\|_{C(\bar{K}_i \cap \bar{K})} \leq \bar{\mu} \varepsilon^{-j}$, for some $\bar{\mu}$ independent of $i \in \mathscr{I}$, and $j \in \{1, 2\}$.

Once we have verified (ix), all the properties demanded in (24) on the partition of unity $\{\varphi_i\}_{i \in \mathscr{I}}$ subordinate to K_i will be satisfied.

In the following calculations we repeatedly use that $\nabla \sum_{i \in \mathscr{I}} \varphi_i(x) = 0$ for $x \in \bar{K}$. This follows from (viii). As short calculation shows that for each $j \in \mathscr{I}$, each $x \in \bar{K}$, and $k, \ell \in \{1, \ldots, n\}$

$$\partial_{x_k} \varphi_j(x) = \frac{\partial_{x_k} \psi_j(x)}{\sum_{i \in \mathscr{I}} \psi_i(x)}, \quad \partial_{x_\ell} \partial_{x_k} \varphi_j(x) = \frac{\partial_{x_\ell} \partial_{x_k} \varphi_j(x) \sum_{i \in \mathscr{I}} \psi_i(x) - \psi_j \sum_{i \in \mathscr{I}} \partial_{x_\ell} \partial_{x_k} \psi_i(x)}{\left(\sum_{i \in \mathscr{I}} \psi_i(x)\right)^2},$$

where we use that $\partial_{x_k} \sum_{i \in \mathscr{I}} \psi_i(x) = 0$ for $x \in \bar{K}$.

To obtain the required estimates we introduce for $\eta > 0$

$$\psi_\eta(x) = \begin{cases} \exp\left(\frac{1}{\left|\frac{x}{\eta}\right|^2 - 1}\right), & \text{for} |x| \leq \eta \\ 0, & \text{for} |x| \leq \eta. \end{cases}$$

Then we have

$$\partial_{x_k} \psi_\eta = -\psi_\eta \frac{2x_k}{\eta^2 \left(\left|\frac{x}{\eta}\right|^2 - 1\right)^2},$$

$$\left(\partial_{x_k}\right)^2 \psi_\eta = \frac{2\psi_\eta}{\eta^2 \left(\left|\frac{x}{\eta}\right|^2 - 1\right)^4} \left[\frac{2x_k^2}{\eta^2} - \left(\left|\frac{x}{\eta}\right|^2 - 1\right)^2 + \frac{4x_k^2}{\eta^2} \left(\left|\frac{x}{\eta}\right|^2 - 1\right)\right],$$

and for $k \neq \ell$

$$\partial_{x_\ell} \partial_{x_k} \psi_\eta = \frac{2\psi_\eta\, x_\ell x_k}{\eta^2 \left(\left|\frac{x}{\eta}\right|^2 - 1\right)^4} \left[\frac{2}{\eta^2} + \frac{4}{\eta^2} \left(\left|\frac{x}{\eta}\right|^2 - 1\right)\right] = \frac{2\psi_\eta\, x_\ell x_k}{\eta^2 \left(\left|\frac{x}{\eta}\right|^2 - 1\right)^4} \left[\frac{-2}{\eta^2} + \frac{4}{\eta^4} |x|^2\right].$$

Considering the behavior of $\partial_{x_k} \psi_\eta$ and $\partial_{x_\ell} \partial_{x_k} \psi_\eta$ separately on the ball $B_{\frac{\eta}{2}}(0)$ and its complement in $B_\eta(0)$, it follows that these functions behave like $O(\frac{1}{\eta})$ and $O(\frac{1}{\eta^2})$. Applying these estimates in the expressions for the first and second derivatives for φ_j and using the lower bound established in (vi) we obtain (ix).

Appendix B. Perturbation results

Here we collect pertinent existence and stability results for dynamical systems. The constant $\widehat{M}(0)$ appearing below relates to Assumption (**A.2**).

Proposition 29. *Let $\mathbf{y} \in \mathscr{C}(Y_0; W_T)$, $\mathbf{y}(y_0) \in \mathscr{Y}_{ad}$ for all $y_0 \in Y_0$, $\delta\mathbf{v} \in \mathscr{C}(Y_0; L^2(I;\mathbb{R}^n))$, as well as $\delta\mathbf{y}_0 \in \mathscr{C}(Y_0;\mathbb{R}^n)$ be given. Moreover let $A\colon I\times\mathbb{R}^n \to \mathbb{R}^{n\times n}$ be continuous, and denote by $\mathbf{A}\colon L^\infty(I;\mathbb{R}^n) \to \mathscr{B}(L^2(I;\mathbb{R}^n))$ the induced Nemitsky operator i.e.*

$$\mathbf{A}(y)\delta y = A(t,y(t))\delta y(t) \quad \forall\, \delta y \in L^2\left(I;\mathbb{R}^n\right),\ y \in L^\infty\left(I;\mathbb{R}^n\right)$$

and a.e. $t\in I$. Then there is $\delta\mathbf{y} \in \mathscr{C}(Y_0; W_T)$ such that

$$\dot{\delta y} = \mathbf{A}(y)\delta y + v,\ \delta y(0) = \delta y_0 \tag{B1}$$

for $y := \mathbf{y}(y_0)$, $\delta v := \delta\mathbf{v}(y_0)$, $\delta y_0 := \delta\mathbf{y}(y_0)$ and all $y_0 \in Y_0$. It satisfies

$$\left\|\delta\mathbf{y}(y_0)\right\|_{W_T} \le C\left(\left\|\delta\mathbf{v}(y_0)\right\|_{L^2} + \left|\delta\mathbf{y}_0(y_0)\right|\right) \tag{B2}$$

for some $C>0$ depending continuously on $\max_{(\tau,y)\in I\times\bar{B}_{2\widehat{M}(0)}} \|A(\tau,y)\|_{\mathbb{R}^{n\times n}}$, and independent of $y_0 \in Y_0$.

Proof. Let $y_0\in Y_0$ be arbitrary but fixed. Then there is a unique solution $\delta y \in W_T$ to (B1) which satisfies

$$\begin{aligned}\frac{1}{2}\left|\delta y(t)\right|^2 &= \frac{1}{2}\left|\delta y(0)\right|^2 + \int_0^t \left(\dot{\delta y}(s), \delta(y)(s)\right)\,\mathrm{d}s\\ &= \frac{1}{2}|\delta y_0|^2 + \int_0^t \left(\delta y(s), A(t,y(t))\delta(y)(s)\right) + \left(\delta v(s), \delta y(s)\right)\,\mathrm{d}s\\ &\le \frac{1}{2}|\delta y_0|^2 + \frac{1}{2}|\delta v|^2_{L^2} + \frac{1}{2}\int_0^t \left(2\max_{(\tau,y)\in I\times\bar{B}_{2\widehat{M}(0)}} \|A(\tau,y)\|_{\mathbb{R}^{n\times n}} + 1\right)|\delta y(s)|^2\,\mathrm{d}s\end{aligned}$$

for all $t\in I$. Setting

$$L := \left(2\max_{(z,y)\in I\times\bar{B}_{2\widehat{M}(0)}} \|A(z,y)\|_{\mathbb{R}^{n\times n}} + 1\right),$$

Gronwall's inequality implies that

$$\|\delta y\|_{L^\infty} \le e^{TL}\left(|\delta y_0| + |\delta v|_{L^2}\right).$$

By (B1) we further get $\|\dot{\delta y}\|_{L^2} \le L(\|\delta y\|_{L^2} + \|v\|_{L^2})$, which implies (B2). Next, let $y_0^k \in Y_0$ denote a convergent sequence with limit y_0. For abbreviation set

$$\delta y_k := \delta\mathbf{y}\left(y_0^k\right),\ y_k := \mathbf{y}\left(y_0^k\right),\ \delta v_k := \delta\mathbf{v}\left(y_0^k\right),\ \delta y_0 := \delta\mathbf{y}_0\left(y_0^k\right)$$

as well as

$$y := \mathbf{y}(y_0),\ \delta v := \delta\mathbf{v}(y_0),\ \delta y_0 := \delta\mathbf{y}_0(y_0).$$

Note that δy_k is uniformly bounded in W_T by (B2). Thus it admits a subsequence, denoted by the same index, with $\delta y_k \rightharpoonup \delta y$ in W_T for some $\delta y \in W_T$. This implies

$$\delta y_k(0) \to \delta y(0) \text{ in } \mathbb{R}^n,\ \delta y_k \to \delta y \text{ in } L^\infty\left(I;\mathbb{R}^n\right),\ \dot{\delta y}_k \rightharpoonup \dot{\delta y} \text{ in } L^2\left(I;\mathbb{R}^n\right).$$

Moreover, due to the continuity of $\mathbf{y}, \delta\mathbf{v}$ and $\delta\mathbf{y}_0$, we get

$$\dot{\delta y}_k = \mathbf{A}(y_k)\delta y_k + \delta v_k \to \mathbf{A}(y)\delta y + \delta v \text{ in } L^2\left(I;\mathbb{R}^n\right),\ \delta y_k(0) \to \delta y_0 \text{ in } \mathbb{R}^n.$$

Summarizing the previous observations we conclude that

$$\dot{\delta y} = \mathbf{A}(y)\delta y + \delta v,\ \delta y(0) = \delta y_0$$

as well as $\delta y_k \to \delta y$ in W_T, and thus $\delta y = \delta \mathbf{y}(y_0)$. By uniqueness of solutions to the above equation $\delta \mathbf{y}(y_0^k) \to \delta \mathbf{y}(y_0)$ for the whole sequence in W_T follows, and therefore $\delta \mathbf{y} \in \mathscr{C}(Y_0; W_T)$. □

Next we address nonlinear systems of the form:

$$\dot{y}_v = \mathbf{f}(y_v) + \mathbf{g}(y_v)\mathscr{F}^*(y_v) + v, \quad y_v(0) = y_0 \tag{B3}$$

where $v \in L^2(I;\mathbb{R}^n)$ is a perturbation.

Proposition 30. *Let Assumption 1 hold. Then there exist an open neighbourhood $V_1 \subset L^2(I;\mathbb{R}^n)$ of 0 and an open neighbourhood $\mathbf{y}_0$ of Y_0 such that* (B3) *admits a unique solution $y_v = \mathbf{y}^v(y_0) \in \mathscr{Y}_{ad}$ for every pair $(v, y_0) \in V_1 \times \mathbf{y}_0$. Moreover the mapping*

$$\mathbf{y}^\bullet(\cdot)\colon V_1 \times \mathbf{y}_0 \to \mathscr{Y}_{ad}, \quad (v, y_0) \mapsto \mathbf{y}^v(y_0) \tag{B4}$$

is continuously Fréchet differentiable.

Proof. Define the mapping

$$G\colon \mathscr{Y}_{ad} \times \mathbb{R}^n \times L^2\left(I;\mathbb{R}^n\right) \to L^2\left(I;\mathbb{R}^n\right) \times \mathbb{R}^n$$

with

$$G(y, y_0, v) = \begin{pmatrix} \dot{y} - \mathbf{f}(y) - \mathbf{g}(y)\mathscr{F}^*(y) - v \\ y(0) - y_0 \end{pmatrix}.$$

Now fix an arbitrary $\bar{y}_0 \in Y_0$ and, utilizing (**A.3**) denote by $\bar{y} = \mathbf{y}^*(y_0) \in \operatorname{int}\mathscr{Y}_{ad}$ the unique solution in $\mathscr{Y}_{ad}$ to the unperturbed closed loop system $G(\bar{y}, \bar{y}_0, 0) = 0$. Since G is of class $\mathscr{C}^1$ in a neighborhood of $(\bar{y}, \bar{y}_0, 0)$ we have

$$D_y G(y, y_0, v)\delta y = \begin{pmatrix} \dot{\delta y} - D\mathbf{f}(y)\delta y - \left[D\mathbf{g}(y)\delta y\right]\mathscr{F}^*(y) - \mathbf{g}(y)\partial_y \mathscr{F}^*(y)\delta y \\ \delta y(0) \end{pmatrix}.$$

It is straightforward that the linearized equation

$$D_y G(\bar{y}, \bar{y}_0, v)\delta y = \begin{pmatrix} \delta v \\ \delta y_0 \end{pmatrix}$$

admits a unique solution $\delta \bar{y} \in W_T$ for every $\delta v \in L^2(I;\mathbb{R}^n)$, $\delta y_0 \in \mathbb{R}^n$. Moreover, applying Gronwall's lemma yields $c > 0$ independent of $\bar{y}$, $\bar{y}_0$ with

$$\|\delta \bar{y}\|_{W_T} \le c(\|\delta v\|_{L^2(I;\mathbb{R}^n)} + |\delta y_0|), \quad \forall\, \delta v \in L^2\left(I;\mathbb{R}^n\right),\ \delta y_0 \in \mathbb{R}^n.$$

Thus from the implicit function theorem we get constants $\kappa_1 = \kappa_1(\bar{y}_0)$ and $\kappa_2 = \kappa_2(\bar{y}_0)$, such that for every $y_0 \in \mathbb{R}^n$ with $|y_0 - \bar{y}_0| < \kappa_1$ and $\|v\|_{L^2(I;\mathbb{R}^n)} < \kappa_2$ there exists $\mathbf{y}^v(y_0) \in \mathscr{Y}_{ad}$ with $G(\mathbf{y}^v(y_0), y_0, v) = 0$. By (**A.1**) it is the unique solution to (B3) in $\mathscr{Y}_{ad}$. Moreover, the mapping

$$\mathbf{y}^\bullet(\cdot)\colon B_{\kappa_2}(0) \times B_{\kappa_1}(\bar{y}_0) \to \mathscr{Y}, \quad (v, y_0) \mapsto \mathbf{y}^v(y_0)$$

is of class $\mathscr{C}^1$. Observe that repeating this argument for every $y_0 \in Y_0$ yields an open covering of Y_0 i.e.

$$Y_0 \subset \bigcup_{\bar{y}_0 \in Y_0} B_{\kappa_1(\bar{y}_0)}(\bar{y}_0).$$

Since Y_0 is compact there exists a finite set of initial conditions $\{\bar{y}_0^i\}_{i=1}^N \subset Y_0$, including 0, such that

$$Y_0 \subset \mathbf{y}_0 := \bigcup_{i=1}^N B_{\kappa_1(\bar{y}_0^i)}\left(\bar{y}_0^i\right).$$

Set $V = \bigcap_{i=1}^N B_{\kappa_2(\bar{y}_0^i)}(0) \subset L^2(I;\mathbb{R}^n)$. Summarizing these arguments yields the existence of a $\mathscr{C}^1$-mapping

$$\mathbf{y}^\cdot(\cdot)\colon V \times \mathbf{y}_0 \to \mathscr{Y}_{ad}, \quad \mathbf{y}^v(y_0) \text{ uniquely solves (B3) in } \mathscr{Y}_{ad}.$$

□

We use the following consequences of the previous proposition.

Corollary 31. *There exists an open neighborhood $V_2 \subset V_1 \subset L^2(I;\mathbb{R}^n)$ of 0 as well as $c > 0$ such that*

$$\left\|\mathbf{y}^{v_1}(y_0) - \mathbf{y}^{v_2}(y_0)\right\|_{W_T} \le c\|v_1 - v_2\|_{L^2(I;\mathbb{R}^n)} \quad \forall\, y_0 \in Y_0,\ v_1 \in V_2, v_2 \in V_2$$

and

$$\left\|\mathbf{y}^{v}(y_0)\right\|_{W_T} \le M_{Y_0} + c\|v\|_{L^2(I;\mathbb{R}^n)} \quad \forall\, y_0 \in Y_0,\ v \in V_2,$$

hold. Here M_{Y_0} denotes the constant from (**A.3**).

Proof. The first assertion follows from the continuous differentiability of $v \to \mathbf{y}^v(y_0)$ and compactness of Y_0. To verify the second we use that $\mathbf{y}^*(y_0) = \mathbf{y}^0(y_0)$ and estimate

$$\left\|\mathbf{y}^{v}(y_0)\right\|_{W_T} \le \left\|\mathbf{y}^{*}(y_0)\right\|_{W_T} + \left\|\mathbf{y}^{v}(y_0) - \mathbf{y}^{0}(y_0)\right\|_{W_T}.$$

The claim now follows from the first inequality and (**A.3**). □

Conflicts of interest

The authors have no conflict of interest to declare.

References

[1] W. H. Fleming, H. M. Soner, *Controlled Markov processes and viscosity solutions*, 2nd ed., Stochastic Modelling and Applied Probability, vol. 25, Springer, 2006.

[2] S. Dolgov, D. Kalise, K. Kunisch, "Tensor decomposition methods for high-dimensional Hamilton-Jacobi-Bellman equations", *SIAM J. Sci. Comput.* **43** (2021), no. 3, p. A1625-A1650.

[3] K. Kunisch, D. Walter, "Semiglobal optimal feedback stabilization of autonomous systems via deep neural network approximation", *ESAIM, Control Optim. Calc. Var.* **27** (2021), article no. 16.

[4] D. Onken, L. Nurbekyan, X. Li, S. W. Fung, S. Osher, L. Ruthotto, "A neural network approach applied to multi-agent optimal control", *preprint*, arXiv:2104.03270, 2021.

[5] O. Bokanowski, J. Garcke, M. Griebel, I. Klompmaker, "An adaptive sparse grid semi-lagrangian scheme for first order Hamilton-Jacobi Bellman equations", *J. Sci. Comput.* **55** (2013), no. 3, p. 575-605.

[6] M. Falcone, R. Ferretti, "Numerical methods for Hamilton-Jacobi type equations", Handbook of Numerical Analysis, vol. 17, North-Holland, 2016, p. 603-626.

[7] D. Kalise, K. Kunisch, Z. Rao, *Hamilton-Jacobi-Bellman equations: Numerical Methods and Applications in Optimal Control*, Radon Series on Computational and Applied Mathematics, vol. 21, Walter de Gruyter, 2018.

[8] D. Kalise, K. Kunisch, "Polynomial approximation of high-dimensional Hamilton-Jacobi-Bellman equations and applications to feedback control of semilinear parabolic PDEs", *SIAM J. Sci. Comput.* **40** (2018), no. 2, p. a629-a652.

[9] K. Fackeldey, M. Oster, L. Sallandt, R. Schneider, "Approximative policy iteration for exit time feedback control problems driven by stochastic differential equations using tensor train format", *preprint*, arXiv:2010.04465, 2020.

[10] P.-L. Lions, J.-C. Rochet, "Hopf formula and multitime Hamilton-Jacobi equations", *Proc. Am. Math. Soc.* **96** (1986), p. 79-84.

[11] Y. T. Chow, W. Li, S. Osher, W. Yin, "Algorithm for Hamilton-Jacobi equations in density space via a generalized Hopf formula", *preprint*, arXiv:1805.01636, 2018.

[12] B. Azmi, D. Kalise, K. Kunisch, "Optimal feedback law recovery by gradient-augmented sparse polynomial regression", *J. Mach. Learn. Res.* **22** (2021), article no. 48.

[13] T. Nakamura-Zimmerer, Q. Gong, W. Kang, "Adaptive deep learning for high-dimensional Hamilton-Jacobi-Bellman equations", *SIAM J. Sci. Comput.* **43** (2021), no. 2, p. a1221-a1247.

[14] D. Bertsekas, *Reinforcement Learning and Optimal Control*, Athena Scientific optimization and computation series, vol. 1, Athena Scientific, 2019.

[15] F. L. Lewis, D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control", *IEEE Circuits Syst. Mag.* **9** (2009), no. 3, p. 32-50.

[16] B. Recht, "A tour of reinforcement learning: The view from continuous control", *preprint*, arXiv:1806.09460, 2018.

[17] K. Vamvoudakis, F. Lewis, S. S. Ge, "Neural networks in feedback control systems", in *Mechanical Engineers' Handbook*, vol. 3, Wiley, 2014, p. 843-894.

[18] M. Falcone, R. Ferretti, *Semi-Lagrangian approximation schemes for linear and Hamilton-Jacobi equations*, Other Titles in Applied Mathematics, vol. 133, Society for Industrial and Applied Mathematics, 2014.
[19] A. Pinkus, "Approximation theory of the MLP model in neural networks", *Acta Numerica* **8** (1999), p. 143-195.
[20] K. Hornik, "Approximation capabilities of multilayer feedforward networks", *Neural Networks* **4** (1991), no. 2, p. 251-257.
[21] J. Diestel, J. J. j. Uhl, *Vector measures*, Mathematical Surveys, vol. 15, American Mathematical Society, 1977.
[22] W. Arendt, M. Kreuter, "Mapping theorems for Sobolev spaces of vector-valued functions", *Stud. Math.* **240** (2018), no. 3, p. 275-299.
[23] J. Barzilai, J. M. Borwein, "Two-point step size gradient methods", *IMA J. Numer. Anal.* **8** (1988), no. 1, p. 141-148.