

LOCAL WELL-POSEDNESS OF THE MORTENSEN OBSERVER

T. BREITEN AND J. SCHRÖDER*

Abstract. The analytical background of nonlinear observers based on minimal energy estimation is discussed. It is shown that locally the derivation of the observer equation based on a trajectory with pointwise minimal energy can be done rigorously. The result is obtained by a local sensitivity analysis of the value function based on Pontryagin's maximum principle and the Hamilton-Jacobi-Bellman equation. The consideration of a differential Riccati equation reveals that locally the second derivative of the value function is a positive definite matrix. The local convexity ensures existence of a trajectory minimizing the energy, which is then shown to satisfy the observer equation.

Mathematics Subject Classification. 49J15, 49L12, 49N60, 93B53.

Received May 15, 2023. Accepted May 27, 2024.

1. INTRODUCTION

We consider a nonlinear perturbed dynamical system of the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + G(x(t) \otimes x(t)) + Fv(t), \quad t \in (0, T], \\ x(0) &= x_0 + \eta \end{aligned} \tag{1.1}$$

where $A \in \mathbb{R}^{n,n}$, $G \in \mathbb{R}^{n,n^2}$, $F \in \mathbb{R}^{n,m}$, $x_0 \in \mathbb{R}^n$ are known and $v \in L^2(0, T; \mathbb{R}^m)$, $\eta \in \mathbb{R}^n$ are assumed to be (deterministic) disturbances. Let us emphasize that the specific quadratic structure of the nonlinearity is mainly due to technical simplification rather than a necessary requirement for the results obtained in this article. In particular, by a process sometimes called *lifting*, a considerably more general class of nonlinear systems can be embedded in a structure of the form (1.1), see, e.g., [1, 2]. For the system (1.1), let us consider a (disturbed) linear observation of the form

$$y(t) = Cx(t) + \mu(t), \tag{1.2}$$

with $C \in \mathbb{R}^{r,n}$ and $\mu \in L^2(0, T; \mathbb{R}^r)$. The goal of this article is the theoretical analysis of the well-known (see [3]) *Mortensen observer*

$$\begin{aligned} \dot{\hat{x}}(t) &= A\hat{x}(t) + G(\hat{x}(t) \otimes \hat{x}(t)) + \alpha \nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}(t), y)^{-1} C^\top (y(t) - C\hat{x}(t)), \quad t \in (0, T], \\ \hat{x}(0) &= x_0 \end{aligned} \tag{1.3}$$

Keywords and phrases: Observer design, minimum energy estimation, Hamilton–Jacobi–Bellman equation, optimal control.
Institute of Mathematics, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany.

* Corresponding author: j.schroeder@tu-berlin.de

which can (formally) be derived by pointwise minimization of the minimal value function of an optimal control problem associated with (1.1). More precisely, for given $t \in (0, T]$ and $\xi \in \mathbb{R}^n$ we have that

$$\begin{aligned} \mathcal{V}(t, \xi, y) = \inf_{\substack{x \in H^1(0, t; \mathbb{R}^n) \\ v \in L^2(0, t; \mathbb{R}^m)}} J(x, v; t, \xi) &:= \frac{1}{2} \|x(0) - x_0\|^2 + \frac{1}{2} \int_0^t \|v\|^2 + \alpha \|y - Cx\|^2 ds, \\ \text{s.t. } e(x, v; t, \xi) &:= (\dot{x} - Ax - G(x \otimes x) - Fv, x(t) - \xi) = 0. \end{aligned} \quad (1.4)$$

An integral part of our analysis is the discussion of the corresponding time-dependent non-homogeneous Hamilton–Jacobi–Bellman (HJB) equation

$$\begin{aligned} \partial_t \mathcal{V}(t, \xi, y) &= -\langle \nabla_\xi \mathcal{V}(t, \xi, y), A\xi + G(\xi \otimes \xi) \rangle - \frac{1}{2} \|F^\top \nabla_\xi \mathcal{V}(t, \xi, y)\|^2 + \frac{\alpha}{2} \|y(t) - C\xi\|^2, \\ \mathcal{V}(0, \xi, y) &= \frac{1}{2} \|\xi - x_0\|^2. \end{aligned} \quad (1.5)$$

The observer (1.3) was initially proposed as a *maximum likelihood filter* in [3] where the author considered a more general setup as we do here, also allowing for nonlinear output operators. As was already noted by Mortensen, it generalizes the Kalman–Bucy filter [4, 5] which is obtained in (1.3) for the special case of linear dynamics, *i.e.*, when $G = 0$ in (1.1). In contrast to similar works on nonlinear (stochastic) filtering [6–9], the maximum likelihood filter relies on a deterministic interpretation of the unknown disturbances η, v and μ thereby connecting it to a deterministic optimal control problem of the form (1.4), see also [10] for a thorough discussion of the linear case. Moreover, the structure (1.3) resembles a state-dependent Luenberger observer [11]. In the literature, (1.3) is also referred to as a *minimum energy estimator* [12, 13] or simply the *Mortensen observer/estimator* [14, 15]. While initially derived on a formal level, in [14] the author provides a rigorous analysis of the optimal control problem as well as the value function under the assumption that the arising nonlinearities are continuously differentiable with globally bounded derivatives, *cf.* also the more detailed general exposition in [16]. In [17], it is shown that if the nonlinearities are globally Lipschitz continuous, then the observer state $\hat{x}(t)$ in (1.3) converges asymptotically to the true state $x(t)$ as $t \rightarrow \infty$. Due to the value function in (1.4) suffering from the curse of dimensionality, different approximations of the Mortensen observer have been proposed and theoretically analyzed. Let us mention [18] which discusses convergence results for nonlinearities with a globally bounded second derivative which is used within an approximate minimum energy estimator. Related to that, in [19] the author discusses convergence results for the extended Kalman filter which is known to be a (first order) polynomial approximation to the Mortensen observer, see also [3]. More recently, in [15] a discrete-time version of the Mortensen observer was analyzed under the assumption of affine dynamics. Let us also point to a numerical realization of the Mortensen observer by a neural network based approximation approach that was studied in [20].

Looking at the observer equation (1.3) it becomes apparent that for the discussion of well-posedness of the observer trajectory it is essential to ensure existence of the Hessian $\nabla_{\xi\xi}^2 \mathcal{V}$ and further to show that it is an invertible matrix. In general however, value functions are notoriously non-smooth, therefore showing the required smoothness is one of the main challenges of this work. For example, the seminal work [21] displays, that HJB equations as in (1.5) generally do not allow for a classical notion of a solution. Early works on regularity of value functions for finite and infinite-dimensional systems can be found in, *e.g.*, [22, 23]. In the latter references, the considered nonlinearities are assumed to be Lipschitz continuous and exhibit a linear growth condition w.r.t. the state variable. In view of the term $G(x \otimes x)$, these results do not directly apply here. Let us also refer to [24, 25] where the authors show local regularity results for HJB solutions based on pointwise regularity. In the context of (unconstrained) infinite-horizon control problems, in [26, 27] it has been shown that the associated (time-independent) value function is infinitely often differentiable. The problem of finite-horizon control problems under state and control constraints was discussed in [28–30]. Here the authors were able to show that the solutions of a parameterized problem depend on said parameter in a differentiable

fashion. However, these results do not yield time-differentiability of the value function considered in this work, because the controls are assumed to be essentially bounded. Further [29, 30] explicitly assume the parameter to be independent of time.

In this article, we will employ some of the ideas from [26, 27] and perform a sensitivity analysis for appropriately chosen time-dependent finite-horizon non-homogeneous control problems as in (1.4). Our main results can be summarized as follows:

- (i) Based on a nominal trajectory obtained for the undisturbed dynamics, we define a local neighborhood on which the HJB equation (1.5) has a classical solution, see Theorem 4.9.
- (ii) For the Hessian $\nabla_{\xi\xi}^2 \mathcal{V}$ of the value function \mathcal{V} , in Theorem 5.1 we analyze a specific differential Riccati equation whose solution we show to be positive definite. As a consequence, in Theorem 5.3, we locally extend this result implying the (local) positive definiteness of $\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, y)$.
- (iii) In Lemma 6.1 we establish an upper bound for minimizers of $\mathcal{V}(t, \cdot, y)$. Together with the local strict convexity of \mathcal{V} a unique minimizer can be identified as a solution of $\nabla_{\xi} \mathcal{V}(t, \cdot, y) = 0$, cf. Proposition 6.2.
- (iv) For continuous y sufficiently close to $C\tilde{x}$ the observer equation can then be derived in a rigorous manner. *Via* a density argument the result is transferred to less regular data. The main result is stated in Theorem 6.6. It shows that if initial perturbation and dynamics as well as output disturbances are sufficiently small, the Mortensen observer (1.3) is well-defined.

The technique of showing regularity of the value function by an application of the inverse mapping theorem that we deploy is a well known tool in the context of sensitivity analysis. A thorough display stating general results can be found for example in [31]. The results presented there could be used to show first order differentiability of the value function with respect to space and time. However, this work further requires time continuity of the spatial derivatives of the value function up to order three. To the best of the authors' knowledge this smoothness can not be obtained directly with any of the available theory, therefore this work contains the technical sensitivity analysis of the specific problem using only elementary results. The above mentioned time continuity of $\nabla_{\xi} \mathcal{V}$, $\nabla_{\xi\xi}^2 \mathcal{V}$ and $\nabla_{\xi\xi\xi}^3 \mathcal{V}$ is shown using an argument of time uniform convergence of the difference quotients utilizing a bound for the spatial derivatives, see Proposition 4.8. Hence it is crucial that all constants in this work remain independent of time. While this poses a challenge throughout the article, it also allows for a stronger result, see Remark 6.7.

Notation. If not mentioned otherwise, $\|\cdot\|$ will denote the Euclidean norm on \mathbb{R}^d , where the dimension d varies. The associated scalar product is denoted by $\langle \cdot, \cdot \rangle$. The i -th canonical basis vector is denoted by e_i . Further we denote by I_d the identity matrix of dimension d and the matrix spectral norm is denoted by $\|\cdot\|_2$. The Kronecker product of two matrices A and B or two vectors x and y is denoted by $A \otimes B$ and $x \otimes y$, respectively. For $1 \leq p \leq \infty$ we denote by $L^p(0, T; \mathbb{R}^d)$ the Lebesgue spaces, while $H^1(0, T; \mathbb{R}^d)$ denotes the Sobolev space of functions with a first weak derivative in $L^2(0, T; \mathbb{R}^d)$. Furthermore $C([0, T]; \mathbb{R}^d)$ and $C^k([0, T]; \mathbb{R}^d)$ denote the spaces of functions $f: [0, T] \rightarrow \mathbb{R}^d$ that are continuous and k times continuously differentiable, respectively. For $d = 1$ the image space is dropped in the notation. The space of all functions from $C^\infty([0, T]; \mathbb{R}^n)$ with compact support is denoted by $C_0^\infty([0, T]; \mathbb{R}^n)$. All mentioned spaces are equipped with their standard norms. The space of solution trajectories will be denoted by V_t and is given by $H^1(0, t; \mathbb{R}^n)$ equipped with the norm $\|\cdot\|_{V_t} = \max(\|\cdot\|_{H^1(0, t; \mathbb{R}^n)}, \|\cdot\|_{L^\infty(0, t; \mathbb{R}^n)})$.

For two Banach spaces X and Y their Cartesian product is denoted by $X \times Y$ and equipped with the norm $\|(x, y)\|_{X \times Y} = \max(\|x\|_X, \|y\|_Y)$. Further $L(X, Y)$ denotes the space of linear and bounded mappings from X to Y . For an element $x \in X$ and a real number $\epsilon > 0$ the open ball of radius ϵ around x is denoted by $\mathcal{U}_\epsilon(x)$. The weak convergence of a sequence x_k to some x is denoted as $x_k \rightharpoonup x$, for $k \rightarrow \infty$. For a function $f: X \rightarrow Y$ its Fréchet derivative is denoted by Df . For a function $f: X_1 \times X_2 \rightarrow Y$, where X_1 and X_2 are Banach spaces, the partial Fréchet derivative with respect to the first variable is denoted by $D_{x_1} f$. Higher order and mixed partial derivatives are denoted with appropriate indices, e.g., $D_{x_1 x_2}^2 f$.

Throughout this work we use c as a generic constant.

2. PRELIMINARIES

For this entire work let $T > 0$ be fixed. In order to formulate the state equation of interest assume that for some given initial value $x_0 \in \mathbb{R}^n$ and matrices $A \in \mathbb{R}^{n,n}$ and $G \in \mathbb{R}^{n,n^2}$ the problem

$$\begin{aligned}\dot{x}(t) &= Ax(t) + G(x(t) \otimes x(t)), \\ x(0) &= x_0,\end{aligned}$$

admits a unique solution $\tilde{x} \in V_T$. This so-called nominal trajectory \tilde{x} will play an important role throughout this work. The state equation is formulated on an interval $[0, t]$, where $t \in (0, T]$ is fixed. It reads

$$\begin{aligned}\dot{x}(s) &= Ax(s) + G(x(s) \otimes x(s)) + Fv(s), \\ x(t) &= \tilde{x}(t) + \xi,\end{aligned}\tag{2.1}$$

where the matrix $F \in \mathbb{R}^{n,m}$ is given. Furthermore the disturbance $v \in L^2(0, t; \mathbb{R}^m)$ is for now fixed. One searches for a solution $x \in V_t$ that satisfies the first equation almost everywhere in $[0, t]$. We point out the subtle difference of the formulations (1.4) and (2.1). In the former ξ denotes the final state of the system while in the latter it denotes the difference from the model trajectory at the final time. This adjustment can be considered as a change of coordinates. Instead of considering neighborhoods around \tilde{x} , we deal with neighborhoods of zero.

We further want to emphasize that the dynamics considered in this work are given by a disturbed initial value problem, see (1.1). This specifically means that the dynamics evolve forward in time, as does the observer trajectory characterized by (1.3). However, the technical discussions of this work are concerned with the state equation of the optimal control problem which is given as a final value problem and therefore is associated with dynamics evolving backwards in time, see (1.4) and (2.1). This discrepancy stems from the idea to define the observer *via* an energy minimization and sets our discussion apart from the dominant literature. There the value function and HJB equation is usually discussed in the context of a forward problem.

Before the state equation can be discussed, a result about a general linear equation is presented. It is of the form

$$\begin{aligned}\dot{x}(s) &= B_t(s)x(s) + f(s), \\ x(0) &= \xi,\end{aligned}\tag{2.2}$$

where $f \in L^2(0, t; \mathbb{R}^n)$ is fixed. The system matrix B_t is not only time-dependent, but also depends on the interval $[0, t]$ on which the equation is considered. However it will be assumed that the spectral norm of the system matrix is bounded uniformly in time and in the parameter t . Since the right hand side is measurable and integrable in s , and fulfills a generalized Lipschitz condition in x , this problem admits a unique solution $x \in V_t$. The following lemma presents an estimate for the solution.

Lemma 2.1. *Let $s \mapsto B_t(s) \in \mathbb{R}^{n,n}$ be a matrix-valued function depending on the parameter $t \in (0, T]$. Let $t \in (0, T]$ and assume that there exists a constant $\kappa > 0$ independent of t and s such that for all $s \in [0, t]$ it holds $\|B_t(s)\|_2 \leq \kappa$. Then there exists a constant $\bar{c}(\kappa) > 0$ independent of t , ξ and f such that the unique solution $x \in V_t$ of (2.2) satisfies*

$$\|x\|_{V_t} \leq \bar{c}(\kappa) (\|\xi\| + \|f\|_{L^2(0, t; \mathbb{R}^n)}).$$

Proof. The proof is done in a standard way by testing the equation with its solution and integrating over time. Utilizing Young's inequality, Gronwall's inequality and the representation of \dot{x} given by the equation yields the assertion. \square

Via a time transformation it can be shown that the same result holds, if one considers a final value problem instead of the initial value problem.

3. THE OPTIMAL CONTROL PROBLEM

The main result of this work is the proof of existence of a solution to the observer equation

$$\begin{aligned} \hat{x}(t) &= A'(t)\hat{x}(t) + G(\hat{x}(t) \otimes \hat{x}(t)) + \alpha \nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}(t), \omega)^{-1} C^\top (\omega(t) - C\hat{x}(t)), \quad t \in (0, T], \\ \hat{x}(0) &= 0, \end{aligned} \quad (3.1)$$

where $A'(t) = A + G(\hat{x}(t) \otimes I_n) + G(I_n \otimes \hat{x}(t))$. The function $\mathcal{V}(t, \xi, \omega)$ is the value function associated with the control problem

$$\begin{aligned} \inf_{(x,v) \in V_t \times L^2(0,t;\mathbb{R}^m)} J(x,v;t,\omega) &:= \frac{1}{2} \|x(0) - x_0\|^2 + \frac{1}{2} \int_0^t \|v\|^2 + \alpha \|\omega - C(x - \tilde{x})\|^2 ds, \\ \text{s.t. } e(x,v;t,\xi) &:= (\dot{x} - Ax - G(x \otimes x) - Fv, x(t) - \tilde{x}(t) - \xi) = 0, \end{aligned} \quad (3.2)$$

where $t \in (0, T]$, $\alpha > 0$, $C \in \mathbb{R}^{r,n}$, $\omega \in L^2(0,t;\mathbb{R}^r)$ and $\xi \in \mathbb{R}^n$ are fixed. The given $x_0 \in \mathbb{R}^n$ is the initial condition of the nominal trajectory \tilde{x} .

Similar to the coordinate change performed for ξ , we utilize a transformation for the measured data. The control problem and the value function are not formulated in terms of the measured data y . Instead they depend on the difference of measured data and modeled output. More precisely, we introduce the variable $\omega = y - C\tilde{x}$. Now the undisturbed model considered on $[0, t]$ corresponds to $\xi = 0$ and $\omega = 0$ instead of $\xi = \tilde{x}(t)$ and $y = C\tilde{x}$. In this sense (3.1) can be understood as an adjusted version of (1.3). Specifically, the solution trajectories are shifted by \tilde{x} . Since both formulations are equivalent, it suffices to show results for the shifted one.

The purpose of Sections 3 to 5 is to prove sufficient regularity of \mathcal{V} such that $\nabla_{\xi\xi}^2 \mathcal{V}$ exists and is an invertible matrix. This however can only be achieved in a local sense, which results in the necessity for an upper bound on $\|\omega\|_{L^2(0,T;\mathbb{R}^r)}$ to ensure solvability of (3.1). More precisely, it will be shown that there exists a constant $\delta > 0$ such that the assertion holds for all ω satisfying $\|\omega\|_{L^2(0,T;\mathbb{R}^r)} < \delta$. In order to achieve this, it is essential that all constants describing locality are independent of time. This goal will pose a challenge throughout this work. We show existence and uniqueness of a solution of the optimal control problem on the time horizon $[0, t]$ for any fixed $t \in (0, T]$. Without extra effort one would obtain the results to hold under the assumption $\|\omega\|_{L^2(0,t;\mathbb{R}^r)} < \delta(t)$, where $\delta \rightarrow 0$, for $t \rightarrow 0$. This issue carries over to $\mathcal{V}(t, \cdot, \cdot)$ being smooth on a domain of radius $\delta(t)$ that vanishes for $t \rightarrow 0$. While it takes some additional effort to avoid such time dependencies it leads to a considerably stronger result. The specific benefits of this effort are summarized in Remark 6.7.

In order to obtain regularity results concerning \mathcal{V} , one first has to perform a standard analysis of the optimal control problem. The remainder of this section is concerned with solvability of the state equation, existence of minimizers and the first order optimality condition expressed *via* the solution of the adjoint state equation.

3.1. State equation

First it will be shown that for $\xi \in \mathbb{R}^n$ and $v \in L^2(0,t;\mathbb{R}^m)$ with sufficiently small norms the state equation (2.1) admits a unique solution. For this subsection let $t \in (0, T]$ be fixed. First consider

$$\dot{x} = -Ax - G(I_n \otimes \tilde{x}_t)x - G(\tilde{x}_t \otimes I_n)x - G(x \otimes x) - f, \quad x(0) = \xi, \quad (3.3)$$

on the interval $[0, t]$, where $\tilde{x}_t(s) := \tilde{x}(t - s)$ is a time transformation of the nominal trajectory on $[0, t]$.

Proposition 3.1. *Let $t \in (0, T]$ and let $\bar{c} = \bar{c}(\|A\|_2 + 2\|G\|_2\|\tilde{x}\|_{L^\infty(0, T; \mathbb{R}^n)})$ be the constant from Lemma 2.1. Then there exists a constant $\delta'_1 > 0$ independent of t such that for all $\xi \in \mathbb{R}^n$ and $f \in L^2(0, t; \mathbb{R}^n)$ satisfying*

$$\|\xi\| + \|f\|_{L^2(0, t; \mathbb{R}^n)} \leq \delta'_1$$

there exists a unique solution $x \in V_t$ of (3.3). It satisfies

$$\|x\|_{V_t} \leq 2\bar{c} (\|\xi\| + \|f\|_{L^2(0, t; \mathbb{R}^n)}).$$

Proof. The proof of existence and the bound is carried out *via* a fixed point argument on the set

$$\mathcal{M} := \{x \in V_t : \|x\|_{V_t} \leq 2\bar{c} (\|\xi\| + \|f\|_{L^2(0, t; \mathbb{R}^n)})\}.$$

Such a strategy to prove existence of a solution can be found for example in [26], Lemma 5. Uniqueness of the solution in V_t is proven in a standard manner. \square

Remark 3.2. For an appropriate estimation of the quadratic term the estimates for both the sup norm and the H^1 -norm are required. The need for the estimate of the sup norm can not be circumvented by the continuous embedding

$$\|x - \tilde{x}\|_{L^\infty(0, t; \mathbb{R}^n)} \leq \check{c}(t)\|x - \tilde{x}\|_{H^1(0, t; \mathbb{R}^n)}.$$

This is due to the unboundedness of the optimal embedding constant $\check{c}(t) = \sinh(t)^{-\frac{1}{2}}$ (see [32]) for t approaching zero. Without the estimate for the sup norm δ_1 would depend on time t with the property $\delta_1(t) \rightarrow 0$ for $t \rightarrow 0$. This issue will come up frequently in the following discussions. Therefore estimates are presented for the V_t -norm instead of the H^1 -norm.

The unique solvability of the state equation follows immediately.

Corollary 3.3. *There exist constants $\delta_1 > 0$ and $\tilde{c} > 0$ independent of t such that for any pair $(\xi, v) \in \mathbb{R}^n \times L^2(0, t; \mathbb{R}^m)$ satisfying $\|\xi\| + \|v\|_{L^2(0, t; \mathbb{R}^m)} \leq \delta_1$ the state equation (2.1) admits a unique solution $x \in V_t$. It satisfies*

$$\|x - \tilde{x}\|_{V_t} \leq \tilde{c} (\|\xi\| + \|v\|_{L^2(0, t; \mathbb{R}^m)}).$$

Proof. Let ξ and v be sufficiently small. According to Proposition 3.1 system (3.3) with $f(s) = Fv(t-s)$ admits a unique solution \underline{x} . It follows that $e(s) := \underline{x}(t-s)$ is the unique solution to

$$\dot{e} = Ae + G(I_n \otimes \tilde{x})e + G(\tilde{x} \otimes I_n)e + G(e \otimes e) + Fv, \quad e(t) = \xi.$$

Then $x = \tilde{x} + e$ is the desired solution and satisfies the estimate of the assertion. \square

3.2. Existence of a minimizer

Next it will be shown that for sufficiently small data the control problem admits a solution. Again the proof can be done using well known techniques and we limit its discussion to a rough sketch. For the general result we refer to [33], Theorem 1.45 and for a very similar statement see [26], Lemma 8.

Proposition 3.4. *Let $t \in (0, T]$. There exists a constant $\delta_2 \in (0, \delta_1]$ independent of t such that for all $(\xi, \omega) \in \mathbb{R}^n \times L^2(0, t; \mathbb{R}^r)$ satisfying $\max(\|\xi\|, \|\omega\|_{L^2(0, t; \mathbb{R}^r)}) \leq \delta_2$ the optimal control problem (3.2) admits a solution*

$(\bar{x}, \bar{v}) \in V_t \times L^2(0, t; \mathbb{R}^m)$. Furthermore, there exists a constant $M_1 > 0$ independent of t , ξ and ω such that any minimizing pair $(z, u) \in V_t \times L^2(0, t; \mathbb{R}^m)$ satisfies

$$\max(\|z - \tilde{x}\|_{V_t}, \|u\|_{L^2(0, t; \mathbb{R}^m)}) \leq M_1 \max(\|\xi\|, \|\omega\|_{L^2(0, t; \mathbb{R}^r)}).$$

Proof. For now set $\delta_2 = \delta_1$. With Corollary 3.3 it follows that the uncontrolled state equation associated with ξ admits a unique solution. Hence one obtains an upper bound for the infimum in (3.2) of the form $j \leq c\delta_2^2$. For a minimizing sequence (x_k, v_k) one can assume $J(x_k, v_k; t, \omega) \leq 2c\delta_2^2$ which yields $\|v_k\|_{L^2(0, t; \mathbb{R}^m)} \leq \sqrt{2}c\delta_2$. A reduction of δ_2 leads to $\|\xi\| + \|v_k\|_{L^2(0, t; \mathbb{R}^m)} \leq \delta_2 + \sqrt{2}c\delta_2 \leq \delta_1$ allowing the application of Corollary 3.3 to the state equation controlled by v_k . Going over to appropriate subsequences of the bounded sequences (x_k) and (v_k) implies the existence of weak limits \bar{x} and \bar{v} . Passing to the weak limit it is shown that $e(\bar{x}, \bar{v}, \xi, t) = 0$. In order to pass to the weak limit in the quadratic term one has to go over to a subsequence that converges strongly in $L^2(0, t; \mathbb{R}^n)$. This is possible because H^1 is compactly embedded in L^2 . Since the cost functional is weakly sequentially lower semi continuous, (\bar{x}, \bar{v}) is a minimizer. The estimate can be derived with the strategies applied to show boundedness of the minimizing sequence. \square

3.3. Necessary optimality condition

In this subsection a first order necessary optimality condition of the optimal control problem will be derived.

Proposition 3.5. *Let $t \in (0, T]$ be fixed and assume $\max(\|\xi\|, \|\omega\|_{L^2(0, t; \mathbb{R}^r)}) \leq \delta_2$ and let $(\bar{x}, \bar{v}) \in V_t \times L^2(0, t; \mathbb{R}^m)$ be a minimizer of the corresponding optimal control problem (3.2). Then there exists a unique adjoint state $p \in V_t$ satisfying*

$$\dot{p} = -A^\top p - (\bar{x}^\top \otimes I_n)G^\top p - (I_n \otimes \bar{x}^\top)G^\top p + \alpha C^\top(\omega - C(\bar{x} - \tilde{x})), \quad (3.4)$$

$$p(0) = x_0 - \bar{x}(0), \quad (3.5)$$

$$\bar{v} + F^\top p = 0, \quad (3.6)$$

where (3.4) is fulfilled almost everywhere in $[0, t]$, (3.5) is fulfilled in \mathbb{R}^n and (3.6) holds in $L^2(0, t; \mathbb{R}^m)$. Furthermore there exists a constant $M_2 > 0$ independent of t , \bar{x} , \bar{v} , and ω such that it holds

$$\|p\|_{V_t} \leq M_2 \max(\|\xi\|, \|\omega\|_{L^2(0, t; \mathbb{R}^r)}).$$

Proof. This standard result can be shown for example by an application of [34], Proposition 1, Section 4.14. \square

4. SENSITIVITY ANALYSIS

For (ξ, ω) satisfying $\max(\|\xi\|, \|\omega\|_{L^2(0, T; \mathbb{R}^r)}) \leq \delta_2$ and $t \in (0, T]$ the value function corresponding to the control problem (3.2) is defined as

$$\mathcal{V}(t, \xi, \omega) = \min_{\substack{x \in V_t \\ v \in L^2(0, t; \mathbb{R}^m)}} J(x, v; t, \omega), \quad \text{subject to: } e(x, v; t, \xi) = 0$$

with initialization defined by $\mathcal{V}(0, \xi, \omega) = \frac{1}{2}\|\xi\|^2$. Note that due to Proposition 3.4 it is well-defined.

The goal of this section is to show sufficient regularity of the value function to ensure that its Hessian $\nabla_{\xi\xi}^2 \mathcal{V}$ is well-defined and differentiable in t .

4.1. Regularity in space and measurement

The regularity of the value function with respect to the last two variables will be shown using the chain rule. To that end it is shown that for sufficiently small data there exists a unique minimizing pair which depends on

the data in a smooth fashion. For a fixed $t \in (0, T]$ this is achieved by an application of the implicit function theorem to the function $P_t: Y_t \times X_t \rightarrow Y_t$, $(w, h) \mapsto \Phi_t(h) - w$, where

$$\begin{aligned} X_t &:= V_t \times L^2(0, t; \mathbb{R}^m) \times V_t \times L^2(0, t; \mathbb{R}^r), \\ Y_t &:= L^2(0, t; \mathbb{R}^r) \times L^2(0, t; \mathbb{R}^n) \times \mathbb{R}^n \times L^2(0, t; \mathbb{R}^n) \times \mathbb{R}^n \times L^2(0, t; \mathbb{R}^m), \end{aligned}$$

and $\Phi_t: X_t \rightarrow Y_t$ is defined by

$$\Phi_t(x, v, p, \omega) = \begin{pmatrix} \omega \\ \dot{x} - Ax - G(x \otimes x) - Fv \\ x(t) - \tilde{x}(t) \\ \dot{p} + A^\top p + (x^\top \otimes I_n)G^\top p + (I_n \otimes x^\top)G^\top p - \alpha C^\top(\omega - C(x - \tilde{x})) \\ p(0) - x_0 + x(0) \\ v + F^\top p \end{pmatrix}. \quad (4.1)$$

Note that it holds $\Phi_t(x, v, p, \omega) = (\omega, 0, \xi, 0, 0, 0)$ if and only if the triple (x, v, p) satisfies the first order necessary optimality condition of the optimal control problem (3.2) associated with ξ and ω .

The strategy of applying the implicit function to the mapping P_t is equivalent to an application of the inverse mapping theorem to Φ_t . The specific version of the implicit function theorem that we apply is presented in [35]. It gives sufficient conditions under which the obtained neighborhoods are given explicitly, which is necessary for the resulting constants to be independent of time. The following technical lemma is needed when verifying the assumptions made in [35]. The proof requires the discussion of a linear quadratic optimal control problem which can be found in Appendix A.

Lemma 4.1. *Denote $h_0 = (\tilde{x}, 0, 0, 0) \in X_t$. Further let $h = (x, v, p, \omega) \in X_t$ be fixed but arbitrary. Then $D\Phi_t(h_0): X_t \rightarrow Y_t$ is bijective. Further there exist constants $\tilde{M}, \check{M} > 0$ independent of t such that*

$$\begin{aligned} \|D\Phi_t(h_0) - D\Phi_t(h)\|_{L(X_t, Y_t)} &\leq \tilde{M}\|h_0 - h\|_{X_t} \text{ and} \\ \|D\Phi_t(h_0)^{-1}\|_{L(Y_t, X_t)} &\leq \check{M}. \end{aligned}$$

Proof. First note that for $h = (x, v, p, \omega) \in X_t$ and $w = (\mu, f, a, -l_1, b, -l_2) \in Y_t$ it holds $D\Phi_t(\tilde{x}, 0, 0, 0)(x, v, p, \omega) = (\mu, f, a, -l_1, b, -l_2)$ if and only if

$$\begin{cases} \omega &= \mu, \\ \dot{x} &= Ax + G(\tilde{x} \otimes I_n)x + G(I_n \otimes \tilde{x})x + Fv + f, \\ x(t) &= a, \\ \dot{p} &= -A^\top p - (\tilde{x}^\top \otimes I_n)G^\top p - (I_n \otimes \tilde{x}^\top)G^\top p + \alpha C^\top(\omega - Cx) - l_1, \\ p(0) &= b - x(0), \\ v + F^\top p + l_2 &= 0. \end{cases} \quad (4.2)$$

We now employ results presented in Appendix A. System (4.2) is equivalent to $\omega = \mu$ and (x, v, p) being the solution of (A.3) associated with $\check{x}_t = \tilde{x}|_{[0, t]}$, $\rho = 0$ and the data given by $w \in Y_t$. Since $\|\tilde{x}|_{[0, t]}\|_{L^\infty(0, t; \mathbb{R}^n)} \leq \|\tilde{x}\|_{L^\infty(0, T; \mathbb{R}^n)}$ and $0 < \beta$, Proposition A.1 yields the unique solvability of the system and hence shows that $D\Phi_t(h_0)$ is bijective. The estimate presented in Proposition A.1 together with the definition of the operator norm further shows the bound for $\|D\Phi_t(h_0)^{-1}\|_{L(Y_t, X_t)}$. To show the first estimate note that for any $(z, u, q, \eta) \in X_t$

satisfying $\|(z, u, q, \eta)\|_{X_t} = 1$ it holds

$$\begin{aligned}
& \|D\Phi_t(h_0)(z, u, q, \eta) - D\Phi_t(h)(z, u, q, \eta)\|_{Y_t} \\
&= \left\| \begin{pmatrix} 0 \\ G((x - \tilde{x}) \otimes z) + G(z \otimes (x - \tilde{x})) \\ 0 \\ ((\tilde{x} - x)^\top \otimes I_n) + (I_n \otimes (\tilde{x} - x)^\top) \\ 0 \\ 0 \end{pmatrix} G^\top q - ((z^\top \otimes I_n) + (I_n \otimes z^\top)) G^\top p \right\|_{Y_t} \\
&\leq 2\|G\|_2 \max(\|z\|_{V_t}\|x - \tilde{x}\|_{V_t}, \|q\|_{V_t}\|x - \tilde{x}\|_{V_t} + \|z\|_{V_t}\|p\|_{V_t}) \\
&\leq 2\|G\|_2 (\|x - \tilde{x}\|_{V_t} + \|p\|_{V_t}) \leq 4\|G\|_2 \|h - h_0\|_{X_t}.
\end{aligned}$$

Taking the supremum over all such $(z, u, q, \eta) \in X_t$ yields the existence of \tilde{M} . \square

It follows the proof that Φ_t has a local inverse, with a domain and image space that are independent of time.

Lemma 4.2. *Let $t \in (0, T]$ be arbitrary. Then there exist constants $\delta_3 > 0$ and $\delta'_3 > 0$ independent of t and three C^∞ functions $\mathcal{X}_t, \mathcal{U}_t$ and \mathcal{P}_t such that for all $(\xi, \omega) \in \mathcal{U}_{\delta_3}(0)$ the quadruple $(\mathcal{X}_t(\xi, \omega), \mathcal{U}_t(\xi, \omega), \mathcal{P}_t(\xi, \omega), \omega)$ is the unique solution to*

$$\Phi_t(x, v, p, \omega) = (\omega, 0, \xi, 0, 0, 0), \quad (x, v, p, \omega) \in \mathcal{U}_{\delta'_3}(\tilde{x}, 0, 0, 0). \quad (4.3)$$

Proof. Let $t \in (0, T]$ be fixed. The proof will be carried out by an application of the implicit function theorem presented in [35] to the C^∞ -mapping $P_t(w, h) = \Phi_t(h) - w$. With $w_0 = 0 \in Y_t$ and $h_0 = (\tilde{x}, 0, 0, 0) \in X_t$ it holds $P_t(w_0, h_0) = 0$. According to Lemma 4.1 $D_h P_t = D\Phi_t$ exists, is continuous in $Y_t \times X_t$ and has a bounded linear inverse. Note that the upper bound of the norm of the inverse \tilde{M} is independent of t . With the non-decreasing functions $g_1(s, i) = \tilde{M}i$ and $g_2(s) = s$ and the constants $\alpha = \frac{1}{2}$, $\delta_3 = \frac{1}{4\tilde{M}\tilde{M}^2}$ and $\delta'_3 = \frac{1}{2\tilde{M}\tilde{M}}$ all assumptions are fulfilled and the assertion holds for the chosen, time-independent constants δ_3 and δ'_3 . Here the functions g_1 and g_2 and the parameter α correspond to the functions and constant with the same name used in [35]. The constants δ_3 and δ'_3 correspond to δ and ϵ in [35] respectively.

At this point let us note that the implicit function theorem presented in [35] only yields continuity of the three functions, but not differentiability. However, comparing the assumptions and the proof with the ones of more classical versions (see for example [34], Thm. 4.E) shows that the neighborhoods are only adjusted within the application of the fixed point argument. The part of the proof that shows that differentiability of order k of P_t carries over to $\mathcal{X}_t, \mathcal{U}_t$ and \mathcal{P}_t does not alter the neighborhoods at all. Hence the assertion of higher regularity holds for the chosen constants. \square

Note that this result does not yet imply that the pair $(\mathcal{X}_t, \mathcal{U}_t)$ minimizes the cost functional. It has only been shown that the triple $(\mathcal{X}_t(\xi, \omega), \mathcal{U}_t(\xi, \omega), \mathcal{P}_t(\xi, \omega))$ satisfies the first order optimality condition. The following proposition ensures that it actually solves the optimal control problem.

Proposition 4.3. *Let $t \in (0, T]$. Then there exists $\delta_4 \in (0, \min(\delta_2, \delta_3, \delta'_3)]$ independent of t such that for all $(\xi, \omega) \in \mathcal{U}_{\delta_4}(0)$ the optimal control problem (3.2) admits exactly one solution. It is given by $(\bar{x}, \bar{v}) = (\mathcal{X}_t(\xi, \omega), \mathcal{U}_t(\xi, \omega))$. The associated adjoint state is given by $\mathcal{P}_t(\xi, \omega)$.*

Proof. For now set $\delta_4 = \min(\delta_2, \delta_3)$. According to Proposition 3.4 there exists a minimizer (\bar{x}, \bar{v}) and it holds

$$\max(\|\bar{x} - \tilde{x}\|_{V_t}, \|\bar{v}\|_{L^2(0, t; \mathbb{R}^m)}) \leq M_1 \max(\|\xi\|, \|\omega\|_{L^2(0, t; \mathbb{R}^r)}) < M_1 \delta_4.$$

Furthermore, from Proposition 3.5 we obtain existence of the associated adjoint state p which satisfies

$$\|p\|_{V_t} \leq M_2 \max(\|\xi\|, \|\omega\|_{L^2(0,t;\mathbb{R}^r)}) < M_2 \delta_4.$$

After a suitable reduction of δ_4 one obtains

$$\max(\|\bar{x} - \tilde{x}\|_{V_t}, \|\bar{v}\|_{L^2(0,t;\mathbb{R}^m)}, \|p\|_{V_t}, \|\omega\|_{L^2(0,t;\mathbb{R}^r)}) < \delta'_3.$$

Since (\bar{x}, \bar{v}, p) is an optimal triple for the control problem (3.2) given by ω and ξ , it holds $\Phi_t(\bar{x}, \bar{v}, p, \omega) = (\omega, 0, \xi, 0, 0, 0)$. Because $(\xi, \omega) \in \mathcal{U}_{\delta_4}(0)$, Lemma 4.2 yields the assertion. \square

With the smoothness of the cost functional an application of the chain rule immediately yields regularity of the value function with respect to ξ and ω .

Corollary 4.4. *For any fixed $t \in (0, T]$ the value function $\mathcal{V}(t, \cdot, \cdot): \mathcal{U}_{\delta_4}(0) \rightarrow \mathbb{R}$ associated with the optimal control problem (3.2) is of class C^∞ .*

We conclude this subsection by a representation of the optimal control evaluated in the final time in terms of the gradient of the value function. Since the proof can be done using well-known techniques, see [22], it is omitted here.

Corollary 4.5. *Let $t \in (0, T]$ and $(\xi, \omega) \in \mathcal{U}_{\frac{1}{2}\delta_4}(0)$. Then the adjoint $p(t) = \mathcal{P}_t(\xi, \omega)(t)$ and the minimizing control $\bar{v}(t) = \mathcal{U}_t(\xi, \omega)(t)$ evaluated in the final time t are characterized by*

$$p(t) = -\nabla_\xi \mathcal{V}(t, \xi, \omega) \text{ and } \bar{v}(t) = F^\top \nabla_\xi \mathcal{V}(t, \xi, \omega).$$

4.2. Regularity in time

This subsection aims at showing time regularity of the value function. One of the essential steps is to prove time uniform bounds for the partial derivatives of the value function \mathcal{V} . This will be done by a characterization of the triples of partial derivatives of \mathcal{X}_t , \mathcal{U}_t and \mathcal{P}_t as solutions of linear quadratic optimal control problems which is presented in Appendix A.

Proposition 4.6. *Let $t \in (0, T]$ and $(\xi, \omega) \in \mathcal{U}_{\delta_4}(0)$. Let $\mu \in L^2(0, t; \mathbb{R}^r)$ and $z_1, z_2, z_3, z_4 \in \mathbb{R}^n$ and for $i = 1, \dots, 4$ denote $z^i = (z_1, \dots, z_i)$. Then there exist $\check{M}_1, \check{M}_2, \check{M}_3, \check{M}_4, \check{M}_5, \check{M}_6 > 0$ independent of t such that for $i = 1, \dots, 4$ it holds*

$$\begin{aligned} \left| D_{\xi^i}^i \mathcal{V}(t, \xi, \omega) z^i \right| &\leq \check{M}_i \prod_{k=1}^i \|z_k\|, \\ \left| D_{\omega \xi}^2 \mathcal{V}(t, \xi, \omega)(\mu, z_1) \right| &\leq \check{M}_5 \|\mu\|_{L^2(0,t;\mathbb{R}^r)} \|z_1\|, \\ \left| D_{\omega \xi^2}^3 \mathcal{V}(t, \xi, \omega)(\mu, z_2, z_1) \right| &\leq \check{M}_6 \|\mu\|_{L^2(0,t;\mathbb{R}^r)} \|z_1\| \|z_2\|. \end{aligned}$$

Proof. For any $t \in (0, T]$ and $(\xi, \omega) \in \mathcal{U}_{\delta_4}(0)$ it holds

$$\mathcal{V}(t, \xi, \omega) = J(\mathcal{X}_t(\xi, \omega), \mathcal{U}_t(\xi, \omega); t, \omega).$$

Forming the appropriate derivatives of this equation, applying the chain rule to the derivative on the right hand side and utilizing the bounds from Proposition 3.4 and Lemma B.1 yields the assertion for $t > 0$. Noting that $\mathcal{V}(0, \xi, \omega) = \frac{1}{2} \|\xi\|^2$ concludes the proof. \square

With these tools at hand it is now possible to show time continuity of the value function. To obtain this result we assume ω to be essentially bounded.

Proposition 4.7. *Let $(\xi, \omega) \in \mathcal{U}_{\frac{1}{2}\delta_4}(0) \subset \mathbb{R}^n \times L^2(0, T; \mathbb{R}^r)$ and assume $\omega \in L^\infty(0, T; \mathbb{R}^r)$. Then the mapping $t \mapsto \mathcal{V}(t, \xi, \omega)$ is continuous in $[0, T]$.*

Proof. Let $t \in [0, T)$ be arbitrary. Then for sufficiently small $\tau > 0$ it holds $t + \tau \in (0, T)$ and with Bellman's principle it holds

$$\begin{aligned} \mathcal{V}(t + \tau, \xi, \omega) - \mathcal{V}(t, \xi, \omega) &= \mathcal{V}(t, \mathcal{X}_{t+\tau}(\xi, \omega)(t) - \tilde{x}(t), \omega) - \mathcal{V}(t, \xi, \omega) \\ &+ \frac{1}{2} \int_t^{t+\tau} \|\mathcal{U}_{t+\tau}(\xi, \omega)\|^2 + \alpha \|\omega - C(\mathcal{X}_{t+\tau}(\xi, \omega) - \tilde{x})\|^2 ds. \end{aligned} \quad (4.4)$$

To see that for $\tau \searrow 0$ the first summand converges to zero, note that $\mathcal{X}_{t+\tau}(\xi, \omega) \in C^1([0, t + \tau]; \mathbb{R}^n)$. This is due to the fact that the optimal control $\mathcal{U}_{t+\tau}(\xi, \omega) = -F^\top \mathcal{P}_{t+\tau}(\xi, \omega)$ is continuous. It follows

$$\begin{aligned} \|\mathcal{X}_{t+\tau}(\xi, \omega)(t) - \tilde{x}(t) - \xi\| &= \left\| \mathcal{X}_{t+\tau}(\xi, \omega)(t + \tau) - \int_t^{t+\tau} \dot{\mathcal{X}}_{t+\tau}(\xi, \omega)(s) ds - \tilde{x}(t) - \xi \right\| \\ &\leq \|\tilde{x}(t + \tau) - \tilde{x}(t)\| + \tau \|\dot{\mathcal{X}}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)}. \end{aligned} \quad (4.5)$$

For any $s \in [0, t + \tau]$ it holds

$$\begin{aligned} \|\dot{\mathcal{X}}_{t+\tau}(\xi, \omega)(s)\| &\leq \|A\mathcal{X}_{t+\tau}(\xi, \omega)(s)\| + \|G(\mathcal{X}_{t+\tau}(\xi, \omega)(s) \otimes \mathcal{X}_{t+\tau}(\xi, \omega)(s))\| + \|F\mathcal{U}_{t+\tau}(\xi, \omega)(s)\| \\ &\leq \|A\|_2 \|\mathcal{X}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)} + \|G\|_2 \|\mathcal{X}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)}^2 \\ &+ \|F\|_2^2 \|\mathcal{P}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)}. \end{aligned}$$

By Proposition 3.4 and Proposition 3.5 it follows

$$\begin{aligned} \|\dot{\mathcal{X}}_{t+\tau}(\xi, \omega)(s)\| &\leq \|A\|_2 \left(\|\tilde{x}\|_{L^\infty(0, t+\tau; \mathbb{R}^n)} + M_1 \max(\|\xi\|, \|\omega\|_{L^2(0, t+\tau, \mathbb{R}^r)}) \right) \\ &+ 2\|G\|_2 \left(\|\tilde{x}\|_{L^\infty(0, t+\tau; \mathbb{R}^n)}^2 + M_1^2 \max(\|\xi\|^2, \|\omega\|_{L^2(0, t+\tau, \mathbb{R}^r)}^2) \right) \\ &+ M_2 \|F\|_2^2 \max(\|\xi\|^2, \|\omega\|_{L^2(0, t+\tau, \mathbb{R}^r)}^2). \end{aligned}$$

Since $t + \tau < T$, there exists some $c > 0$ independent of τ such that

$$\|\dot{\mathcal{X}}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)} \leq c$$

holds for all sufficiently small τ . Therefore the right hand side of (4.5) tends to zero for $\tau \searrow 0$. The fact that $\nu \mapsto \mathcal{V}(t, \nu, \omega)$ is continuous in $\nu = \xi$ then implies that the first summand of (4.4) goes to zero for $\tau \searrow 0$.

For the second summand it holds

$$\begin{aligned} &\int_t^{t+\tau} \|\mathcal{U}_{t+\tau}(\xi, \omega)\|^2 + \alpha \|\omega - C(\mathcal{X}_{t+\tau}(\xi, \omega) - \tilde{x})\|^2 ds \\ &\leq \int_t^{t+\tau} \|F\|_2^2 \|\mathcal{P}_{t+\tau}(\xi, \omega)\|^2 + 2\alpha \|\omega\|_{L^\infty(0, T; \mathbb{R}^r)}^2 + 2\alpha \|C\|_2^2 \|\mathcal{X}_{t+\tau}(\xi, \omega) - \tilde{x}\|^2 ds \\ &\leq \tau \left(\|F\|_2^2 \|\mathcal{P}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)}^2 + 2\alpha \|\omega\|_{L^\infty(0, T; \mathbb{R}^r)}^2 + 2\alpha \|C\|_2^2 \|\mathcal{X}_{t+\tau}(\xi, \omega) - \tilde{x}\|_{L^\infty(0, t+\tau; \mathbb{R}^n)}^2 \right). \end{aligned}$$

With Proposition 3.4 and Proposition 3.5 all terms on the right hand side can be estimated from above by a constant independent of τ . Hence the right hand side goes to zero for $\tau \searrow 0$. This shows that for all $t \in [0, T)$

it holds

$$\mathcal{V}(t + \tau, \xi, \omega) \rightarrow \mathcal{V}(t, \xi, \omega), \quad \text{for } \tau \searrow 0. \quad (4.6)$$

Now let $t \in (0, T]$ be fixed but arbitrary. Then for sufficiently small $\tau > 0$ it holds $t - \tau \in (0, T)$. With Bellman's principle it follows

$$\begin{aligned} \mathcal{V}(t, \xi, \omega) - \mathcal{V}(t - \tau, \xi, \omega) &= \mathcal{V}(t - \tau, \mathcal{X}_t(\xi, \omega)(t - \tau) - \tilde{x}(t - \tau), \omega) - \mathcal{V}(t - \tau, \xi, \omega) \\ &\quad + \int_{t-\tau}^t \|\mathcal{U}_t(\xi, \omega)\|^2 + \alpha \|\omega - C(\mathcal{X}_t(\xi, \omega) - \tilde{x})\|^2 ds. \end{aligned} \quad (4.7)$$

To see convergence of the first summand note that

$$\begin{aligned} \mathcal{X}_t(\xi, \omega)(t - \tau) - \tilde{x}(t - \tau) &= \mathcal{X}_t(\xi, \omega)(t) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds - \tilde{x}(t - \tau) \\ &= \xi + \tilde{x}(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds. \end{aligned}$$

Since $\mathcal{X}_t(\xi, \omega) \in C^1([0, t]; \mathbb{R}^n)$, one has that $\|\dot{\mathcal{X}}_t(\xi, \omega)\|_{L^\infty(0, t; \mathbb{R}^n)}$ is finite. From here on assume that τ is small enough such that $t - \tau \in (0, T)$ and

$$\|\mathcal{X}_t(\xi, \omega)(t - \tau) - \tilde{x}(t - \tau)\| \leq \|\xi\| + \|\tilde{x}(t) - \tilde{x}(t - \tau)\| + \tau \|\dot{\mathcal{X}}_t(\xi, \omega)\|_{L^\infty(0, t; \mathbb{R}^n)} < \delta_4.$$

With Taylor there exists some $\theta \in [0, 1]$ such that

$$\begin{aligned} &\mathcal{V}(t - \tau, \mathcal{X}_t(\xi, \omega)(t - \tau) - \tilde{x}(t - \tau), \omega) - \mathcal{V}(t - \tau, \xi, \omega) \\ &= \left\langle \nabla_\xi \mathcal{V} \left(t - \tau, \xi + \theta \left(\tilde{x}(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds \right), \omega \right), \left(\tilde{x}(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds \right) \right\rangle. \end{aligned}$$

Then Proposition 4.6 yields

$$\begin{aligned} &|\mathcal{V}(t - \tau, \mathcal{X}_t(\xi, \omega)(t - \tau) - \tilde{x}(t - \tau), \omega) - \mathcal{V}(t - \tau, \xi, \omega)| \\ &\leq \check{M}_1 \left\| \tilde{x}(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds \right\| \leq \check{M}_1 \|\tilde{x}(t) - \tilde{x}(t - \tau)\| + \check{M}_1 \tau \|\dot{\mathcal{X}}_t(\xi, \omega)\|_{L^\infty(0, t; \mathbb{R}^n)}. \end{aligned}$$

Therefore the first summand in (4.7) tends to zero for $\tau \searrow 0$. The second summand can be treated with the same arguments as the second summand of (4.4). This finally shows that for all $t \in (0, T]$ it holds

$$\mathcal{V}(t - \tau, \xi, \omega) \rightarrow \mathcal{V}(t, \xi, \omega), \quad \text{for } \tau \searrow 0$$

and the assertion is shown. \square

Because all constants are independent of time the continuity of the value function carries over to its spatial derivatives. In the following proofs we frequently apply Taylor's theorem as presented in [34], Theorem 4.C and characterize the integral remainder term using the mean value theorem for real valued functions, cf., [36], Section 18 Satz 7.

Proposition 4.8. *Let $(\xi, \omega) \in \mathcal{U}_{\frac{1}{2}\delta_4}(0) \subset \mathbb{R}^n \times L^2(0, T; \mathbb{R}^r)$ and assume $\omega \in L^\infty(0, T; \mathbb{R}^r)$. Then the mappings $t \mapsto \nabla_\xi \mathcal{V}(t, \xi, \omega)$, $t \mapsto \nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega)$ and $t \mapsto \nabla_{\xi\xi\xi}^3 \mathcal{V}(t, \xi, \omega)$ are continuous in $[0, T]$.*

Proof. First, let us show that for any $i \in \{1, \dots, n\}$ the partial derivative $t \mapsto \frac{\partial}{\partial \xi_i} \mathcal{V}(t, \xi, \omega)$ has the claimed regularity. Note that for any h with $|h|$ small enough, we have $\|\xi + he_i\| < \frac{1}{2}\delta_4$. Then Proposition 4.7 shows that for such h it holds $f_h \in C([0, T])$, where

$$f_h(t) := \frac{1}{h} (\mathcal{V}(t, \xi + he_i, \omega) - \mathcal{V}(t, \xi, \omega)).$$

By definition f_h converges pointwise to $\frac{\partial}{\partial \xi_i} \mathcal{V}(\cdot, \xi, \omega)$ for $h \rightarrow 0$. Let us show that f_h is a Cauchy sequence in $C([0, T])$. The fact that $C([0, T])$ is a complete space then implies the uniform convergence of f_h and the pointwise limit coincides with the uniform limit, which is an element of $C([0, T])$.

Specifically we need to show that for any $\epsilon > 0$ there exists $h_0 > 0$ such that for all h_1, h_2 with $|h_1|, |h_2| < h_0$ it follows

$$\|f_{h_1} - f_{h_2}\|_{C([0, T])} < \epsilon.$$

Taylor's Theorem yields existence of $\theta_1, \theta_2 \in [0, 1]$ such that

$$\begin{aligned} & \sup_{s \in [0, T]} \left| \frac{1}{h_1} (\mathcal{V}(s, \xi + h_1 e_i, \omega) - \mathcal{V}(s, \xi, \omega)) - \frac{1}{h_2} (\mathcal{V}(s, \xi + h_2 e_i, \omega) - \mathcal{V}(s, \xi, \omega)) \right| \\ &= \sup_{s \in [0, T]} \left| \frac{1}{h_1} \nabla_\xi \mathcal{V}(s, \xi + \theta_1 h_1 e_i, \omega) h_1 e_i - \frac{1}{h_2} \nabla_\xi \mathcal{V}(s, \xi + \theta_2 h_2 e_i, \omega) h_2 e_i \right| \\ &= \sup_{s \in [0, T]} \left| \frac{\partial}{\partial \xi_i} \mathcal{V}(s, \xi + \theta_1 h_1 e_i, \omega) - \frac{\partial}{\partial \xi_i} \mathcal{V}(s, \xi + \theta_2 h_2 e_i, \omega) \right| \end{aligned}$$

After another application of Taylor's Theorem it follows the existence of $\theta_3 \in [0, 1]$ such that

$$\begin{aligned} \|f_{h_1} - f_{h_2}\|_{C([0, T])} &= \sup_{s \in [0, T]} \left| \nabla_\xi \frac{\partial}{\partial \xi_i} \mathcal{V}(s, \xi + \theta_2 h_2 e_i + \theta_3 (\theta_1 h_1 e_i - \theta_2 h_2 e_i), \omega) (\theta_1 h_1 e_i - \theta_2 h_2 e_i) \right| \\ &= (\theta_1 h_1 - \theta_2 h_2) \sup_{s \in [0, T]} \left| \frac{\partial^2}{\partial \xi_i^2} \mathcal{V}(s, \xi + (\theta_2 h_2 + \theta_1 \theta_3 h_1 - \theta_2 \theta_3 h_2) e_i, \omega) \right| \end{aligned}$$

A sufficiently small upper bound h_0 for $|h_1|$ and $|h_2|$ yields

$$\|\xi + (\theta_2 h_2 + \theta_1 \theta_3 h_1 - \theta_2 \theta_3 h_2) e_i\| \leq \|\xi\| + 2|h_2| + |h_1| < \frac{1}{2}\delta_4.$$

With the bound from Proposition 4.6 and a possible further decrease of h_0 it follows

$$\|f_{h_1} - f_{h_2}\|_{C([0, T])} \leq (|h_1| + |h_2|) \check{M}_2 < \epsilon.$$

This concludes the proof that f_h is a Cauchy sequence in a complete space. Hence the pointwise limit $\mathcal{V}(\cdot, \xi, \omega)$ is also the uniform limit and an element of $C([0, T])$.

Since Proposition 4.6 includes bounds for the spatial derivatives of \mathcal{V} up to order four, the assertion for derivatives of order two and three can be shown analogously. \square

We can finally show that for sufficiently small (ξ, ω) the value function is classically differentiable in time. This implies in particular that locally there exists a strong solution to the HJB which is given by the derivative of the value function. For some of the technical proofs, we refer to Appendix B.

Theorem 4.9. *Let $(\xi, \omega) \in \mathcal{U}_{\frac{1}{2}\delta_4}(0) \subset \mathbb{R}^n \times L^2(0, T; \mathbb{R}^r)$ and assume ω to be continuous. Then the mapping $t \mapsto \mathcal{V}(t, \xi, \omega)$ is differentiable in $(0, T]$. Its derivative in $t \in (0, T]$ is given by*

$$\partial_t \mathcal{V}(t, \xi, \omega) = -\langle \nabla_\xi \mathcal{V}(t, \xi, \omega), h(t, \xi) \rangle - \frac{1}{2} \|F^\top \nabla_\xi \mathcal{V}(t, \xi, \omega)\|^2 + \frac{\alpha}{2} \|\omega(t) - C\xi\|^2, \quad (4.8)$$

where $h(t, \xi) = A\xi + G(\tilde{x}(t) \otimes \xi) + G(\xi \otimes \tilde{x}(t)) + G(\xi \otimes \xi)$.

Proof. For sufficiently small $\tau > 0$ it holds

$$\begin{aligned} \mathcal{X}_{t+\tau}(\xi, \omega)(t) - \tilde{x}(t) &= \mathcal{X}_{t+\tau}(\xi, \omega)(t + \tau) - \int_t^{t+\tau} \dot{\mathcal{X}}_{t+\tau}(\xi, \omega) \, ds - \tilde{x}(t) \\ &= \xi + \tilde{x}(t + \tau) - \tilde{x}(t) - \int_t^{t+\tau} \dot{\mathcal{X}}_{t+\tau}(\xi, \omega) \, ds. \end{aligned}$$

Hence with Bellman's principle it follows

$$\begin{aligned} \frac{1}{\tau} (\mathcal{V}(t + \tau, \xi, \omega) - \mathcal{V}(t, \xi, \omega)) &= \frac{1}{\tau} \mathcal{V} \left(t, \xi + \tilde{x}(t + \tau) - \tilde{x}(t) - \int_t^{t+\tau} \dot{\mathcal{X}}_{t+\tau}(\xi, \omega) \, ds, \omega \right) - \frac{1}{\tau} \mathcal{V}(t, \xi, \omega) \\ &\quad + \frac{1}{2\tau} \int_t^{t+\tau} \|\mathcal{U}_{t+\tau}(\xi, \omega)\|^2 + \alpha \|\omega - C(\mathcal{X}_{t+\tau}(\xi, \omega) - \tilde{x})\|^2 \, ds. \end{aligned}$$

With Lemma B.5 for $\tau \searrow 0$ the second term converges to

$$\frac{1}{2} \|\mathcal{U}_t(\xi, \omega)(t)\|^2 + \frac{\alpha}{2} \|\omega(t) - C(\mathcal{X}_t(\xi, \omega)(t) - \tilde{x}(t))\|^2.$$

A Taylor expansion shows that there exists some $\theta(\tau) \in [0, 1]$ such that the first term is equal to

$$\begin{aligned} &\frac{1}{\tau} \left\langle \nabla_\xi \mathcal{V} \left(t, \xi + \theta(\tau) \left(\tilde{x}(t + \tau) - \tilde{x}(t) - \int_t^{t+\tau} \dot{\mathcal{X}}_{t+\tau}(\xi, \omega) \, ds \right), \omega \right), \tilde{x}(t + \tau) - \tilde{x}(t) - \int_t^{t+\tau} \dot{\mathcal{X}}_{t+\tau}(\xi, \omega) \, ds \right\rangle \\ &= \left\langle \nabla_\xi \mathcal{V} \left(t, \xi + \theta(\tau) \left(\tilde{x}(t + \tau) - \tilde{x}(t) - \int_t^{t+\tau} \dot{\mathcal{X}}_{t+\tau}(\xi, \omega) \, ds \right), \omega \right), \frac{\tilde{x}(t + \tau) - \tilde{x}(t)}{\tau} - \frac{1}{\tau} \int_t^{t+\tau} \dot{\mathcal{X}}_{t+\tau}(\xi, \omega) \, ds \right\rangle. \end{aligned}$$

Due to Lemma B.4 and the continuity of $\nabla_\xi \mathcal{V}$ with respect to ξ (shown in Cor. 4.4) the right hand side converges to

$$\left\langle \nabla_\xi \mathcal{V}(t, \xi, \omega), \dot{\tilde{x}}(t) - \dot{\mathcal{X}}_t(\xi, \omega)(t) \right\rangle.$$

It remains to show convergence of $\frac{1}{\tau} (\mathcal{V}(t, \xi, \omega) - \mathcal{V}(t - \tau, \xi, \omega))$. First note that it holds

$$\mathcal{X}_t(\xi, \omega)(t - \tau) - \tilde{x}(t - \tau) = \mathcal{X}_t(\xi, \omega)(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) \, ds = \xi + \tilde{x}(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) \, ds.$$

With Bellman's principle it follows

$$\begin{aligned} \frac{1}{\tau} (\mathcal{V}(t, \xi, \omega) - \mathcal{V}(t - \tau, \xi, \omega)) &= \frac{1}{\tau} (\mathcal{V}(t - \tau, \mathcal{X}_t(\xi, \omega)(t - \tau) - \tilde{x}(t - \tau), \omega) - \mathcal{V}(t - \tau, \xi, \omega)) \\ &\quad + \frac{1}{2\tau} \int_{t-\tau}^t \|\mathcal{U}_t(\xi, \omega)\|^2 + \alpha \|\omega - C(\mathcal{X}_t(\xi, \omega) - \tilde{x})\|^2 ds. \end{aligned} \quad (4.9)$$

A second order Taylor expansion shows that the first term of the right hand side is equal to

$$\begin{aligned} &= \frac{1}{\tau} \mathcal{V} \left(t - \tau, \xi + \tilde{x}(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds, \omega \right) - \frac{1}{\tau} \mathcal{V}(t - \tau, \xi, \omega) \\ &= \frac{1}{\tau} \left\langle \nabla_{\xi} \mathcal{V}(t - \tau, \xi, \omega), \tilde{x}(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds \right\rangle \\ &\quad + \frac{1}{\tau} \int_0^1 (1 - \theta) \left(\tilde{x}(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds \right)^{\top} \\ &\quad \nabla_{\xi\xi}^2 \mathcal{V} \left(t - \tau, \xi + \theta \left(\tilde{x}(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds \right), \omega \right) \\ &\quad \left(\tilde{x}(t) - \tilde{x}(t - \tau) - \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds \right) d\theta. \end{aligned}$$

Since $s \mapsto \nabla_{\xi} \mathcal{V}(s, \xi, \omega)$ and $s \mapsto \dot{\mathcal{X}}_t(\xi, \omega)(s)$ are continuous in $s = t$, for $\tau \searrow 0$ the first order term converges to

$$\left\langle \nabla_{\xi} \mathcal{V}(t, \xi, \omega), \dot{\tilde{x}}(t) - \dot{\mathcal{X}}_t(\xi, \omega)(t) \right\rangle.$$

For sufficiently small τ Proposition 4.6 yields that the absolute value of the second order term is bounded from above by

$$\tau \int_0^1 \check{M}_2 \left\| \frac{\tilde{x}(t) - \tilde{x}(t - \tau)}{\tau} - \frac{1}{\tau} \int_{t-\tau}^t \dot{\mathcal{X}}_t(\xi, \omega)(s) ds \right\|^2 d\theta.$$

Since the term in the integral converges for $\tau \searrow 0$, it is bounded. Therefore the right hand side converges to zero for $\tau \searrow 0$. Since $\mathcal{U}_t(\xi, \omega)(s)$, $\mathcal{X}_t(\xi, \omega)(s)$, $\omega(s)$ and $\tilde{x}(s)$ are continuous in $s = t$, the second term in the right hand side of (4.9) converges to

$$\frac{1}{2} \|\mathcal{U}_t(\xi, \omega)(t)\|^2 + \frac{\alpha}{2} \|\omega(t) - C(\mathcal{X}_t(\xi, \omega)(t) - \tilde{x}(t))\|^2,$$

for $\tau \searrow 0$. This finally proves

$$\begin{aligned} \partial_t \mathcal{V}(t, \xi, \omega) &= \lim_{\tau \rightarrow 0} \frac{\mathcal{V}(t + \tau, \xi, \omega) - \mathcal{V}(t, \xi, \omega)}{\tau} \\ &= \left\langle \nabla_{\xi} \mathcal{V}(t, \xi, \omega), \dot{\tilde{x}}(t) - \dot{\mathcal{X}}_t(\xi, \omega)(t) \right\rangle + \frac{1}{2} \|\mathcal{U}_t(\xi, \omega)(t)\|^2 + \frac{\alpha}{2} \|\omega(t) - C(\mathcal{X}_t(\xi, \omega)(t) - \tilde{x}(t))\|^2 \\ &= - \langle \nabla_{\xi} \mathcal{V}(t, \xi, \omega), h(t, \xi) \rangle - \langle \nabla_{\xi} \mathcal{V}(t, \xi, \omega), FF^{\top} \nabla_{\xi} \mathcal{V}(t, \xi, \omega) \rangle + \frac{1}{2} \|F^{\top} \nabla_{\xi} \mathcal{V}(t, \xi, \omega)\|^2 + \frac{\alpha}{2} \|\omega(t) - C\xi\|^2, \end{aligned}$$

where we used Corollary 4.5.

This shows the assertion for all $t \in (0, T)$. The result can be extended to $t = T$ by a consideration of this entire work on the time interval $[0, \hat{T}]$ for some $\hat{T} > T$. This argument is repeated implicitly whenever we extend regularity results to the right hand boundary of the time interval. \square

5. REGULARITY OF THE HESSIAN

In this section it will be shown that for sufficiently small data the Hessian of the value function is an invertible matrix at any time t . We begin with a characterization of the Hessian $\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)$.

Proposition 5.1. *For all $t \in (0, T]$ it holds*

$$\partial_t \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0) = \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0) B(t) + B^\top(t) \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0) - \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0) F F^\top \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0) + \alpha C^\top C,$$

where $B(t) := -A - G(\tilde{x}(t) \otimes I_n) - G(I_n \otimes \tilde{x}(t))$.

Proof. Since $\omega = 0$ is continuous, (4.8) holds with $(\xi, \omega) = (\xi, 0)$ with $\|\xi\|$ small enough. Then taking partial derivatives with respect to ξ_i and ξ_j and evaluating in $(\xi, \omega) = (0, 0)$ yields

$$\begin{aligned} \frac{\partial^2}{\partial \xi_j \partial \xi_i} \partial_t \mathcal{V}(t, 0, 0) &= - \left\langle \frac{\partial^2}{\partial \xi_j \partial \xi_i} \nabla_\xi \mathcal{V}(t, 0, 0), h(t, 0) \right\rangle - \left\langle \frac{\partial}{\partial \xi_i} \nabla_\xi \mathcal{V}(t, 0, 0), \frac{\partial}{\partial \xi_j} h(t, 0) \right\rangle \\ &\quad - \left\langle \frac{\partial}{\partial \xi_j} \nabla_\xi \mathcal{V}(t, 0, 0), \frac{\partial}{\partial \xi_i} h(t, 0) \right\rangle - \left\langle \nabla_\xi \mathcal{V}(t, 0, 0), \frac{\partial^2}{\partial \xi_j \partial \xi_i} h(t, 0) \right\rangle \\ &\quad - \left\langle F^\top \frac{\partial^2}{\partial \xi_j \partial \xi_i} \nabla_\xi \mathcal{V}(t, 0, 0), F^\top \nabla_\xi \mathcal{V}(t, 0, 0) \right\rangle \\ &\quad - \left\langle F^\top \frac{\partial}{\partial \xi_i} \nabla_\xi \mathcal{V}(t, 0, 0), F^\top \frac{\partial}{\partial \xi_j} \nabla_\xi \mathcal{V}(t, 0, 0) \right\rangle + \alpha \langle C^\top C e_i, e_j \rangle. \end{aligned}$$

Note that for any $\eta \in \mathbb{R}^n$ it holds $\mathcal{V}(t, \eta, 0) \geq 0$ and $\mathcal{V}(t, 0, 0) = 0$. It follows that $\eta = 0$ is a minimizer of $\mathcal{V}(t, \cdot, 0)$ and therefore it holds $\nabla_\xi \mathcal{V}(t, 0, 0) = 0$ as well as $h(t, 0) = 0$. Consequently, we obtain

$$\begin{aligned} \frac{\partial^2}{\partial \xi_j \partial \xi_i} \partial_t \mathcal{V}(t, 0, 0) &= - \left\langle \frac{\partial}{\partial \xi_i} \nabla_\xi \mathcal{V}(t, 0, 0), \frac{\partial}{\partial \xi_j} h(t, 0) \right\rangle - \left\langle \frac{\partial}{\partial \xi_j} \nabla_\xi \mathcal{V}(t, 0, 0), \frac{\partial}{\partial \xi_i} h(t, 0) \right\rangle \\ &\quad - \left\langle F^\top \frac{\partial}{\partial \xi_i} \nabla_\xi \mathcal{V}(t, 0, 0), F^\top \frac{\partial}{\partial \xi_j} \nabla_\xi \mathcal{V}(t, 0, 0) \right\rangle + \alpha \langle C^\top C e_i, e_j \rangle. \end{aligned}$$

With Lemma B.6 (iii) we can switch the order of differentiation on the left hand side. Writing the resulting n^2 equations as a matrix equation yields the assertion. \square

Exploiting the fact that the Hessian is continuous in time and the fact that $\nabla_{\xi\xi}^2(0, \xi, \omega) = I_d$ holds for all (ξ, ω) , it is shown that the Hessian in $(0, 0)$ is invertible at all times.

Proposition 5.2. *The Hessian $\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)$ is positive definite for all $t \in [0, T]$.*

Proof. Since the set of symmetric positive definite matrices is open in the space of symmetric matrices with respect to any matrix norm $\|\cdot\|$, there exists some $\epsilon > 0$ such that any symmetric matrix A satisfying $\|I_n - A\| < \epsilon$ is positive definite. Note that Proposition 4.8 shows that $\nabla_{\xi\xi}^2 \mathcal{V}(\cdot, 0, 0)$ is continuous in 0. Hence there exists some $t_0 \in (0, T)$ such that

$$\|I_n - \nabla_{\xi\xi}^2 \mathcal{V}(t_0, 0, 0)\| = \|\nabla_{\xi\xi}^2 \mathcal{V}(0, 0, 0) - \nabla_{\xi\xi}^2 \mathcal{V}(t_0, 0, 0)\| < \epsilon,$$

implying that $\nabla_{\xi\xi}^2 \mathcal{V}(t_0, 0, 0)$ is positive definite. With Proposition 5.1 it follows that $\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)$ is the unique solution on $(0, T]$ of the differential Riccati equation

$$\begin{aligned}\dot{\Pi}(t) &= \Pi(t)B(t) + B(t)^\top \Pi(t) - \Pi(t)FF^\top \Pi(t) + \alpha C^\top C, \\ \Pi(t_0) &= \nabla_{\xi\xi}^2 \mathcal{V}(t_0, 0, 0).\end{aligned}$$

Since $\nabla_{\xi\xi}^2 \mathcal{V}(t_0, 0, 0)$ is positive definite, from [37], Proposition 1.1 we obtain that the Hessian $\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)$ is positive definite for all $t \in (0, T]$. The positive definiteness of the matrix $\nabla_{\xi\xi}^2 \mathcal{V}(0, 0, 0) = I_n$ is clear and the assertion is shown. \square

Finally it can be shown that for sufficiently small (ξ, ω) the Hessian is an invertible matrix.

Theorem 5.3. *There exists a constant $\delta_5 \in (0, \frac{1}{2}\delta_4]$ such that for any $t \in [0, T]$ and $(\xi, \omega) \in \mathcal{U}_{\delta_5}(0)$ the Hessian $\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega)$ is positive definite.*

Proof. The goal of this proof is to show that one can choose a time-independent δ_5 such that for all $t \in [0, T]$ and $(\xi, \omega) \in \mathcal{U}_{\delta_5}(0)$ it holds

$$\|\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega) - \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)\|_{n,n} \leq \frac{1}{2} \|\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)^{-1}\|_{n,n}^{-1}.$$

Then [38], Proposition 7 in Section 1.23 yields that $\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega)$ is an invertible matrix.

As a first step we will show continuity of the mapping $s \mapsto \nabla_{\xi\xi}^2 \mathcal{V}(s, 0, 0)^{-1}$, which then implies existence of a minimum over $[0, T]$. First note that Proposition 4.8 ensures continuity of $s \mapsto \nabla_{\xi\xi}^2 \mathcal{V}(s, 0, 0)$ in $[0, T]$. Since the mapping $A \mapsto \det(A)$ is a continuous mapping, we conclude that there exists $s^* \in [0, T]$ such that for all $t \in [0, T]$ it holds

$$\det(\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)) \geq \min_{s \in [0, T]} \det(\nabla_{\xi\xi}^2 \mathcal{V}(s, 0, 0)) = \det(\nabla_{\xi\xi}^2 \mathcal{V}(s^*, 0, 0)) > 0,$$

where the last inequality follows with the invertibility of $\nabla_{\xi\xi}^2 \mathcal{V}(s^*, 0, 0)$ shown in Proposition 5.2. Finally with Cramer's rule it follows that $s \mapsto \nabla_{\xi\xi}^2 \mathcal{V}(s, 0, 0)^{-1}$ is continuous in $[0, T]$.

For now set $\delta_5 = \frac{1}{2}\delta_4$. Let $t \in (0, T]$ be fixed and let (ξ, ω) satisfy $\max(\|\xi\|, \|\omega\|_{L^2(0, t; \mathbb{R}^r)}) < \delta_5$. Let $\|\cdot\|_{n,n}$ denote the maximum norm on $\mathbb{R}^{n,n}$. Further let $i, j \in \{1, \dots, n\}$ be such that

$$\max_{\substack{k, l \in \\ \{1, \dots, n\}}} \left| \frac{\partial^2}{\partial \xi_k \partial \xi_l} \mathcal{V}(t, \xi, \omega) - \frac{\partial^2}{\partial \xi_k \partial \xi_l} \mathcal{V}(t, 0, 0) \right| = \left| \frac{\partial^2}{\partial \xi_i \partial \xi_j} \mathcal{V}(t, \xi, \omega) - \frac{\partial^2}{\partial \xi_i \partial \xi_j} \mathcal{V}(t, 0, 0) \right|.$$

Taylor's theorem implies the existence of $\theta_\xi(t) \in [0, 1]$ and $\theta_\omega(t) \in [0, 1]$ such that

$$\begin{aligned}\|\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega) - \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)\|_{n,n} &= \left| \frac{\partial^2}{\partial \xi_i \partial \xi_j} \mathcal{V}(t, \xi, \omega) - \frac{\partial^2}{\partial \xi_i \partial \xi_j} \mathcal{V}(t, 0, 0) \right| \\ &\leq \left| \frac{\partial^2}{\partial \xi_i \partial \xi_j} \mathcal{V}(t, \xi, \omega) - \frac{\partial^2}{\partial \xi_i \partial \xi_j} \mathcal{V}(t, 0, \omega) \right| + \left| \frac{\partial^2}{\partial \xi_i \partial \xi_j} \mathcal{V}(t, 0, \omega) - \frac{\partial^2}{\partial \xi_i \partial \xi_j} \mathcal{V}(t, 0, 0) \right| \\ &= \left| D_\xi \frac{\partial^2}{\partial \xi_i \partial \xi_j} \mathcal{V}(t, \theta_\xi(t)\xi, \omega) \xi \right| + \left| D_\omega \frac{\partial^2}{\partial \xi_i \partial \xi_j} \mathcal{V}(t, 0, \theta_\omega(t)\omega) \omega \right| \\ &= |D_{\xi^3}^3 \mathcal{V}(t, \theta_\xi(t)\xi, \omega)(\xi, e_i, e_j)| + |D_{\omega^3}^3 \mathcal{V}(t, 0, \theta_\omega(t)\omega)(\omega, e_i, e_j)| \\ &\leq \check{M}_3 \|\xi\| + \check{M}_6 \|\omega\|_{L^2(0, t; \mathbb{R}^r)} \leq (\check{M}_3 + \check{M}_6) \delta_5.\end{aligned}$$

Hence an appropriate decrease of δ_5 ensures

$$\|\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega) - \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)\|_{n,n} \leq \frac{1}{2} \min_{s \in [0, T]} \|\nabla_{\xi\xi}^2 \mathcal{V}(s, 0, 0)^{-1}\|_{n,n}^{-1} \leq \frac{1}{2} \|\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)^{-1}\|_{n,n}^{-1}.$$

Note that the minimum is a constant independent of t . Therefore δ_5 can be chosen independent of t as well. It is proven that $\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega)$ is an invertible matrix.

To conclude the proof it remains to show that it is positive definite. Assuming the contrary implies that there exists a non-positive eigenvalue. Since the matrix is invertible it must be non-zero and therefore negative. Consider the function $f: [0, 1] \rightarrow \mathbb{R}$, where $f(1)$ is that negative eigenvalue and $f(s)$ is the corresponding eigenvalue of $\nabla_{\xi\xi}^2 \mathcal{V}(t, s\xi, s\omega)$. Since eigenvalues depend continuously on the matrix entries, f is a continuous function. As shown earlier $\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)$ is a symmetric positive definite matrix implying $f(0) > 0$. By assumption it holds $f(1) < 0$. With the continuity of f it follows the existence of $s \in (0, 1)$ such that $f(s) = 0$. Then $\nabla_{\xi\xi}^2 \mathcal{V}(t, s\xi, s\omega)$ has a zero eigenvalue and is therefore not invertible. This leads to a contradiction because $s(\xi, \omega) \in \mathcal{U}_{\delta_5}(0)$. \square

Remark 5.4.

- (i) Note that it is essential to enforce $\delta_5 < \delta_4$. Therefore a time-dependence of δ_4 would have carried over to δ_5 .
- (ii) To obtain the positive definiteness of $\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)$ in Proposition 5.2 we heavily depend on the fact that $\omega = 0 \in L^2(0, T; \mathbb{R}^r)$ is continuous and essentially bounded. Only due to this regularity Theorem 4.9 ensures that the HJB equation holds and eventually leads to Proposition 5.1. However, in Theorem 5.3 we only assume, that $\omega \in L^2(0, T; \mathbb{R}^r)$ has a sufficiently small norm. The assumption of continuity is no longer required.

6. THE MORTENSEN OBSERVER

This section contains the main result of this work. We will show that for sufficiently small ω and fixed arbitrary t the value function admits a unique minimizer in ξ . The resulting trajectory will be characterized as a solution of the observer equation.

Lemma 6.1. *There exist constants $\delta_6 \in (0, \delta_5]$ and $\hat{c}_1 > 0$ such that for every $t \in [0, T]$ and $\omega \in L^2(0, t; \mathbb{R}^r)$ with $\|\omega\|_{L^2(0, t; \mathbb{R}^r)} < \delta_6$ the following holds: The mapping $\xi \mapsto \mathcal{V}(t, \cdot, \omega)$ admits at most one global minimizer and it holds*

$$\xi \in \mathbb{R}^n \text{ is a minimizer of } \mathcal{V}(t, \cdot, \omega) \iff \nabla_{\xi} \mathcal{V}(t, \xi, \omega) = 0 \text{ and } \|\xi\| < \hat{c}_1 \delta_6.$$

Proof. For now set $\delta_6 = \delta_5$ and let $t \in (0, T]$ be arbitrary. First note that

$$\inf_{\xi} \mathcal{V}(t, \xi, \omega) \leq \mathcal{V}(t, 0, \omega) = \min_{\substack{x, v \text{ st.} \\ e(x, v; t, 0) = 0}} J(x, v; t, \omega) \leq J(\tilde{x}, 0; t, \omega) \leq \frac{\alpha}{2} \|\omega\|_{L^2(0, t; \mathbb{R}^r)}^2 < \frac{\alpha}{2} \delta_6^2.$$

Assume $\xi^* \in \mathbb{R}^n$ to be a minimizer. It follows

$$\inf_{\substack{x, v \text{ st.} \\ e(x, v; t, \xi^*) = 0}} J(x, v; t, \omega) = \mathcal{V}(t, \xi^*, \omega) < \frac{\alpha}{2} \delta_6^2.$$

Hence there exist $x \in V_t$ and $v \in L^2(0, t; \mathbb{R}^m)$ such that $e(x, v; t, \xi^*) = 0$ and $J(x, v; t, \omega) < \frac{\alpha}{2} \delta_6^2$. Then $\frac{1}{2} \|x(0) - \tilde{x}(0)\|^2 < \frac{\alpha}{2} \delta_6^2$ and $\frac{1}{2} \|v\|_{L^2(0, t; \mathbb{R}^m)}^2 < \frac{\alpha}{2} \delta_6^2$. With the argument used in the proof of Proposition 3.1 and a possible

decrease of δ_6 it can be shown that there exists $c_1 > 0$ independent of t, x, v, ω and ξ^* such that

$$\|x - \tilde{x}\|_{L^\infty(0,t;\mathbb{R}^n)} \leq c_1 (\|x(0) - \tilde{x}(0)\| + \|v\|_{L^2(0,t;\mathbb{R}^m)}).$$

It follows

$$\begin{aligned} \frac{1}{2}\|\xi^*\|^2 &\leq \frac{1}{2}\|x(t) - \tilde{x}(t)\|^2 \leq \frac{1}{2}\|x - \tilde{x}\|_{L^\infty(0,t;\mathbb{R}^n)}^2 \leq c_1^2 (\|x(0) - \tilde{x}(0)\|^2 + \|v\|_{L^2(0,t;\mathbb{R}^m)}^2) \\ &\leq 2c_1^2 J(x, v; t, \omega) < c_1^2 \alpha \delta_6^2. \end{aligned}$$

Therefore the search for a minimizer can be restricted to the open convex set

$$\mathcal{K} := \left\{ \xi \in \mathbb{R}^n : \|\xi\| < \sqrt{2\alpha} c_1 \delta_6 \right\}.$$

A possible decrease of δ_6 ensures $\mathcal{K} \subset \mathcal{U}_{\delta_5}(0)$. According to Theorem 5.3 the Hessian $\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega)$ is positive definite for all $\xi \in \mathcal{K}$, implying that $\mathcal{V}(t, \cdot, \omega)$ is strictly convex on \mathcal{K} . The assertion is shown for $t \in (0, T]$. Noting that $\mathcal{V}(0, \xi, \omega) = \frac{1}{2}\|\xi\|^2$ concludes the proof. \square

For any fixed $t \in (0, T]$ an application of the implicit function theorem to $\nabla_\xi \mathcal{V}(t, \cdot, \cdot)$ shows existence of a unique minimizer.

Proposition 6.2. *There exists a constant $\delta_7 \in (0, \delta_6]$ such that for every $t \in [0, T]$ and $\omega \in L^2(0, t; \mathbb{R}^m)$ with $\|\omega\|_{L^2(0,t;\mathbb{R}^m)} < \delta_7$ the mapping $\xi \mapsto \mathcal{V}(t, \xi, \omega)$ admits exactly one minimizer $\xi^* \in \mathbb{R}^n$. Further ξ^* depends continuously on ω and there exists a constant $\hat{c}_2 > 0$ such that it holds $\|\xi^*\| \leq \hat{c}_2 \delta_7$.*

Proof. Due to Lemma 6.1 we only have to show existence of $\delta_7 \in (0, \delta_6]$ such that for any $t \in (0, T]$ and $\omega \in L^2(0, t; \mathbb{R}^r)$ satisfying $\|\omega\|_{L^2(0,t;\mathbb{R}^r)} < \delta_7$ there exists $\xi^* \in \mathcal{U}_{\hat{c}_2 \delta_7}(0)$ with $\nabla_\xi \mathcal{V}(t, \xi^*, \omega) = 0$. For a fixed $t \in (0, T]$ we will apply [35], Theorem to the mapping $(\xi, \omega) \mapsto P(\xi, \omega) := \nabla_\xi \mathcal{V}(t, \xi, \omega)$, which is a continuous map on $\Omega = \mathcal{U}_{\delta_5}(0) \subset \mathbb{R}^n \times L^2(0, t; \mathbb{R}^r)$.

First note that $D_\xi P = \nabla_{\xi\xi}^2 \mathcal{V}(t, \cdot, \cdot)$ exists and is continuous in Ω . According to Theorem 5.3 the derivative $D_\xi P(0, 0) = \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)$ is invertible and with Lemma B.7 it holds $\|\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)^{-1}\|_2 < M_V$. Since all matrix norms are equivalent, there exists a constant $c_1 > 0$ such that $\|A\|_2 \leq c_1 \|A\|_{n,n}$ for all $A \in \mathbb{R}^{n,n}$. Hence the arguments made in the proof of Theorem 5.3 yield that for all $(\xi, \omega) \in \Omega$ it holds

$$\|\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega) - \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)\|_2 \leq c_1 (\tilde{M}_3 \|\xi\| + \tilde{M}_6 \|\omega\|_{L^2(0,t;\mathbb{R}^r)}) =: g_1(\|\xi\|, \|\omega\|_{L^2(0,t;\mathbb{R}^r)}).$$

After noting that it holds $\nabla_\xi \mathcal{V}(t, 0, 0) = 0$ a similar argument shows that for all $\omega \in \mathcal{U}_{\delta_5}(0)$ one has

$$\|\nabla_\xi \mathcal{V}(t, 0, \omega)\| \leq c_2 \tilde{M}_5 \|\omega\|_{L^2(0,t;\mathbb{R}^r)} =: g_2(\|\omega\|_{L^2(0,t;\mathbb{R}^r)}),$$

where $c_2 > 0$ is such that $\|\xi\| \leq c_2 \max\{|\xi_i| : i = 1, \dots, n\}$ for all $\xi \in \mathbb{R}^n$.

For the verification of the last assumption made in [35] we set $\alpha = \frac{1}{2}$, $\delta_7 = \delta_6$ and $\epsilon(\delta_7) = 2c_2 M_V \tilde{M}_5 \delta_7$. A possible decrease of δ_7 ensures $\epsilon(\delta_7) < \delta_5$ and

$$\delta_7 M_V c_1 (\tilde{M}_6 + 2c_2 \tilde{M}_3 \tilde{M}_5 M_V) \leq \frac{1}{2}.$$

Now [35], Theorem, yields the existence of a continuous operator F that maps $\mathcal{U}_{\delta_7}(0) \subset L^2(0, t; \mathbb{R}^r)$ into $\overline{\mathcal{U}_{\epsilon(\delta_7)}(0)}$ with the property $\nabla_\xi \mathcal{V}(t, F(\omega), \omega) = 0$. Another decrease of δ_7 ensures $\epsilon(\delta_7) < \hat{c}_1 \delta_6$. Now for any $\omega \in L^2(0, t; \mathbb{R}^r)$ with $\|\omega\|_{L^2(0,t;\mathbb{R}^r)} < \delta_7$ there exists $\xi^* = F(\omega)$ satisfying $\|\xi^*\| < \hat{c}_1 \delta_6$ and $\nabla_\xi \mathcal{V}(t, \xi^*, \omega) = 0$. Lemma 6.1 yields the assertion for $t \in (0, T]$. The result obviously holds for $t = 0$. \square

With this result at hand we can define $\hat{x}(t)$ as the trajectory given pointwise in time as the minimizer of the value function with respect to ξ .

Corollary 6.3. *Assume $\omega \in L^2(0, T; \mathbb{R}^r)$ satisfies $\|\omega\|_{L^2(0, T; \mathbb{R}^r)} < \delta_7$. Then for all $t \in [0, T]$ the expression $\arg \min_{\xi} \mathcal{V}(t, \xi, \omega) \in \mathbb{R}^n$ is well defined and we define the function*

$$\hat{x}_\omega: [0, T] \rightarrow \mathbb{R}^n, \quad \hat{x}_\omega(t) = \arg \min_{\xi} \mathcal{V}(t, \xi, \omega).$$

For all $t \in [0, T]$ it satisfies $\|\hat{x}_\omega(t)\| \leq \hat{c}_2 \delta_7 < \hat{c}_1 \delta_6 < \delta_5$.

In the following we will characterize this trajectory as a solution of the observer equation. We first do this for continuous data ω before using a density argument to extend the result to $\omega \in L^2(0, T; \mathbb{R}^r)$. As a first step it is shown that for continuous ω the trajectory \hat{x}_ω is differentiable in $(0, T)$.

Lemma 6.4. *Assume $\omega \in L^2(0, T; \mathbb{R}^r)$ satisfies $\|\omega\|_{L^2(0, T; \mathbb{R}^r)} < \delta_7$ and is continuous in $(0, T)$. Then \hat{x}_ω is differentiable in all $t \in (0, T)$. For any $t \in (0, T)$ it holds*

$$\dot{\hat{x}}_\omega(t) = h(t, \hat{x}_\omega(t)) + \alpha \nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_\omega(t), \omega)^{-1} C^\top (\omega(t) - C\hat{x}_\omega(t)),$$

where again $h(t, \xi) = A\xi + G(\hat{x}(t) \otimes I_n)\xi + G(I_n \otimes \hat{x}(t))\xi + G(\xi \otimes \xi)$.

Proof. With Lemma B.6 (i) it holds that the mappings $(t, \xi) \mapsto \nabla_\xi \mathcal{V}(t, \xi, \omega)$ and $(t, \xi) \mapsto \nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega)$ are continuous in $(0, T) \times \mathcal{U}_{\delta_5}(0)$ and Corollary 6.3 ensures $\|\hat{x}_\omega(t)\| < \delta_5$ for all t . Lemma B.6 (ii) further yields that $\nabla_\xi \mathcal{V}(\cdot, \cdot, \omega)$ is differentiable in t and that $\partial_t \nabla_\xi \mathcal{V}(t, \xi, \omega) = \nabla_\xi \partial_t \mathcal{V}(t, \xi, \omega)$. Taking a spatial derivative on both sides of the HJB equation one gets

$$\nabla_\xi \partial_t \mathcal{V}(t, \xi, \omega) = -\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega) h(t, \xi) - Dh(t, \xi) \nabla_\xi \mathcal{V}(t, \xi, \omega) - \nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega) FF^\top \nabla_\xi \mathcal{V}(t, \xi, \omega) - \alpha C^\top (\omega(t) - C\xi), \quad (6.1)$$

where $h(t, \xi)$ is defined as in Theorem 4.9. Since all terms on the right hand side are continuous in $(0, T) \times \mathcal{U}_{\delta_5}(0)$, the same holds for $(t, \xi) \mapsto \partial_t \nabla_\xi \mathcal{V}(t, \xi, \omega)$ and it follows $\nabla_\xi \mathcal{V}(\cdot, \cdot, \omega) \in C^1((0, T) \times \mathcal{U}_{\delta_5}(0); \mathbb{R}^n)$.

Let $t_0 \in (0, T)$ be arbitrary. In the following we apply the implicit function theorem [34], Theorem 4.E to the mapping $P(t, \xi) = \nabla_\xi \mathcal{V}(t, \xi, \omega)$ in the point $(t_0, \hat{x}_\omega(t_0))$. To that end we note that $P(t_0, \hat{x}_\omega(t_0)) = \nabla_\xi \mathcal{V}(t_0, \hat{x}_\omega(t_0)) = 0$ because $\hat{x}_\omega(t_0)$ minimizes $\xi \mapsto \mathcal{V}(t_0, \xi, \omega)$. Furthermore the norm bound on $\hat{x}_\omega(t_0)$ together with Theorem 5.3 ensures that $D_\xi P(t_0, \hat{x}_\omega(t_0)) = \nabla_{\xi\xi}^2 \mathcal{V}(t_0, \hat{x}_\omega(t_0), \omega)$ is an invertible matrix. Hence the implicit function theorem yields classic differentiability of \hat{x}_ω in a neighborhood of t_0 . Since $t_0 \in (0, T]$ was chosen arbitrarily, the announced regularity is shown. To obtain the formula for the derivative note that for all $t \in (0, T)$ it holds

$$\nabla_\xi \mathcal{V}(t, \hat{x}_\omega(t), \omega) = 0. \quad (6.2)$$

Taking the derivative with respect to t and applying the chain rule yields

$$\nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_\omega(t), \omega) \dot{\hat{x}}_\omega(t) = -\partial_t \nabla_\xi \mathcal{V}(t, \hat{x}_\omega(t), \omega).$$

As mentioned above we can change the order of differentiation on the right hand side. With (6.1) and (6.2) it follows

$$\nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_\omega(t), \omega) \dot{\hat{x}}_\omega(t) = \nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_\omega(t), \omega) h(t, \hat{x}_\omega(t)) + \alpha C^\top (\omega(t) - C\hat{x}_\omega(t)).$$

The bounds on $\|\omega\|_{L^2(0, T; \mathbb{R}^r)}$ and $\|\hat{x}_\omega(t)\|$ ensure the existence of the inverse of the Hessian and the assertion is shown. \square

With this lemma we can show that for continuous ω the trajectory \widehat{x}_ω is a weak solution of (3.1).

Proposition 6.5. *Assume $\omega \in L^2(0, T; \mathbb{R}^r)$ satisfies $\|\omega\|_{L^2(0, T; \mathbb{R}^r)} < \delta_7$ and is continuous in $[0, T]$. Then \widehat{x}_ω lies in $H^1(0, T; \mathbb{R}^n)$ and is a weak solution to the observer equation (3.1).*

Proof. Due to Lemma 6.4 we already know that \widehat{x}_ω is differentiable almost everywhere in $[0, T]$ with a derivative that lies in $L^2(0, T; \mathbb{R}^n)$. Furthermore it holds $\widehat{x}_\omega(0) = \arg \min_{\xi} \frac{1}{2} \|\xi\|^2 = 0$. It remains to show that \widehat{x}_ω is absolutely continuous in $[0, T]$. To that end we first show that it is continuous in $t = 0$.

Let $(t_k) \subset (0, T)$ be an arbitrary sequence with the property $t_k \rightarrow 0$ for $k \rightarrow \infty$. We will show that $\widehat{x}_\omega(t_k) \rightarrow 0 = \widehat{x}_\omega(0)$ for $k \rightarrow \infty$. First note that for all k it holds $\|\widehat{x}_\omega(t_k)\| < \widehat{c}_2 \delta_7$ which implies that $(\widehat{x}_\omega(t_k))$ is a bounded sequence in \mathbb{R}^n . Hence any arbitrary subsequence admits a converging subsequence. It will be denoted by $(\widehat{x}_\omega(t_k))$ with the limit $\widehat{x} \in \mathbb{R}^n$. With Lemma B.6 (i) it holds that $(t, \xi) \mapsto \nabla_\xi \mathcal{V}(t, \xi, \omega)$ is continuous in $[0, T] \times \mathcal{U}_{\delta_5}(0)$. Then it follows

$$\widehat{x} = \nabla_\xi \mathcal{V}(0, \widehat{x}, \omega) = \lim_{k \rightarrow \infty} \nabla_\xi \mathcal{V}(t_k, \widehat{x}_\omega(t_k), \omega) = 0,$$

where the first equality is due the fact that in $t = 0$ it holds $\mathcal{V}(0, \xi, \omega) = \frac{1}{2} \|\xi\|^2$ for all $\xi \in \mathbb{R}^n$. This implies that every subsequence of $\widehat{x}_\omega(t_k)$ admits a subsequence that converges to 0. Hence the sequence itself converges to zero. Therefore \widehat{x}_ω is continuous in $t = 0$. It remains to show that it is absolutely continuous in $[0, T]$.

To that end let $\epsilon > 0$ be arbitrary. Then there exists $\delta > 0$ such that it holds

$$|a| < \delta \implies \|\widehat{x}_\omega(a)\| < \frac{1}{2} \epsilon.$$

Further note that \widehat{x}_ω is bounded on $(0, T)$. This is due to Lemma 6.4, the bound $\|\widehat{x}_\omega(t)\| \leq \widehat{c}_2 \delta_7$ and the assumed continuity of ω . Let (a_k, b_k) be a finite sequence of pairwise disjoint subintervals from $[0, T]$ with $a_k < b_k$ for $k = 1, \dots, N$. Without loss of generality assume that $a_1 = 0$. With Taylor's Theorem it holds

$$\sum_{k=1}^N \|\widehat{x}_\omega(b_k) - \widehat{x}_\omega(a_k)\| = \|\widehat{x}_\omega(b_1)\| + \sum_{k=2}^N \|\widehat{x}_\omega(b_k) - \widehat{x}_\omega(a_k)\| \leq \|\widehat{x}_\omega(b_k)\| + \max_{s \in (0, T)} \|\widehat{x}'_\omega(s)\| \sum_{k=2}^N b_k - a_k.$$

After an appropriate decrease of δ it holds $\delta \max_{s \in (0, T)} \|\widehat{x}'_\omega(s)\| < \frac{1}{2} \epsilon$. Then for any such finite sequence of subintervals it holds

$$\sum_{k=1}^N b_k - a_k < \delta \implies \sum_{k=1}^N \|\widehat{x}_\omega(b_k) - \widehat{x}_\omega(a_k)\| < \epsilon,$$

which shows that \widehat{x}_ω is absolutely continuous in $[0, T]$. The almost everywhere existing classical derivative given in Lemma 6.4 is an element of $L^2(0, t; \mathbb{R}^n)$. For one-dimensional domains this implies that \widehat{x}_ω admits a weak derivative which almost everywhere agrees with the classical one and the assertion follows. \square

Finally a density argument is used to show that the assertion of Proposition 6.5 also holds for $\omega \in L^2(0, T; \mathbb{R}^r)$.

Theorem 6.6. *Assume $\omega \in L^2(0, T; \mathbb{R}^r)$ satisfies $\|\omega\|_{L^2(0, T; \mathbb{R}^r)} < \delta_7$. Then \widehat{x}_ω lies in $H^1(0, T; \mathbb{R}^n)$ and is a weak solution to the observer equation (3.1).*

Remark 6.7. Before we turn to the main result of this work, some remarks are in order. The assumptions of Theorem 6.6 also underline the importance of avoiding time-dependence of the constants throughout this work. If the constant δ_7 was time-dependent with the property $\delta_7(t) \rightarrow 0, t \rightarrow 0$ the result would be considerably

weaker. One would have to assume that for all $t \in (0, T]$ it holds

$$\int_0^t \|\omega(s)\|^2 ds < \delta_7(t)^2.$$

This means in particular that any continuous ω would need to satisfy $\omega(0) = 0$. Translating this to the original coordinates and setting used in (1.1) and (1.2) this would imply $C\eta = -\mu(0)$, meaning that the modeled observation of the disturbance in the initial state and the observation error in time zero line up. It would also significantly increase the technical difficulties in the subsequent proof.

Proof. Let ω_k be a sequence from $C([0, T]; \mathbb{R}^r)$ such that $\|\omega - \omega_k\|_{L^2(0, T; \mathbb{R}^r)} \rightarrow 0$ for $k \rightarrow \infty$ and $\|\omega_k\|_{L^2(0, T; \mathbb{R}^r)} < \delta_7$ for all k . Consider the sequence $(\hat{x}_{\omega_k}) \subset H^1(0, T; \mathbb{R}^n)$. According to Proposition 6.2 for any fixed $t \in [0, T]$ the mapping $\omega \mapsto \hat{x}_\omega(t)$ is continuous implying that \hat{x}_{ω_k} converges to \hat{x}_ω pointwise. For all $k \in \mathbb{N}$ and $t \in [0, T]$ it holds $\|\hat{x}_{\omega_k}(t)\| \leq \hat{c}_2 \delta_7$, hence the dominated convergence theorem implies that for all $\varphi \in C_0^\infty(0, T; \mathbb{R}^n)$ it holds

$$\int_0^T \langle \hat{x}_\omega(t), \dot{\varphi}(t) \rangle dt = \lim_{k \rightarrow \infty} \int_0^T \langle \hat{x}_{\omega_k}(t), \dot{\varphi}(t) \rangle dt = - \lim_{k \rightarrow \infty} \int_0^T \langle \dot{\hat{x}}_{\omega_k}(t), \varphi(t) \rangle dt,$$

where $\dot{\hat{x}}_{\omega_k}$ is given by the right hand side of the observer equation. It remains to show that

$$\begin{aligned} & \int_0^T \langle h(t, \hat{x}_\omega(t)) + \alpha \nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_\omega(t), \omega)^{-1} C^\top (\omega(t) - C\hat{x}_\omega(t)), \varphi(t) \rangle dt \\ & - \int_0^T \langle h(t, \hat{x}_{\omega_k}(t)) + \alpha \nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_{\omega_k}(t), \omega_k)^{-1} C^\top (\omega_k(t) - C\hat{x}_{\omega_k}(t)), \varphi(t) \rangle dt \end{aligned}$$

converges to zero for $k \rightarrow \infty$. First consider

$$\left| \int_0^T \langle h(t, \hat{x}_\omega(t)) - h(t, \hat{x}_{\omega_k}(t)), \varphi(t) \rangle dt \right| \leq \|\varphi\|_{L^\infty(0, T; \mathbb{R}^n)} \int_0^T \|h(t, \hat{x}_\omega(t)) - h(t, \hat{x}_{\omega_k}(t))\| dt.$$

Note that for any $\xi \in \mathbb{R}^n$ and $t \in [0, T]$ it holds $h(t, \xi) \leq (\|A\|_2 + 2\|G\|_2 \|\hat{x}\|_{L^\infty(0, T; \mathbb{R}^n)}) \|\xi\| + \|G\|_2 \|\xi\|^2$. Further h is continuous in its second argument, hence the pointwise convergence of \hat{x}_{ω_k} and dominated convergence imply convergence to zero. Next consider

$$\begin{aligned} & \left| \int_0^T \langle \nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_\omega(t), \omega)^{-1} C^\top \omega(t) - \nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_{\omega_k}(t), \omega_k)^{-1} C^\top \omega_k(t), \varphi(t) \rangle dt \right| \\ & \leq \|\varphi\|_{L^\infty(0, T; \mathbb{R}^n)} \int_0^T \|\nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_\omega(t), \omega)^{-1} C^\top \omega(t) - \nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_{\omega_k}(t), \omega_k)^{-1} C^\top \omega_k(t)\| dt. \end{aligned}$$

The integral can be estimated from above by

$$\begin{aligned} & \|C\|_2 \int_0^T \|\nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_\omega(t), \omega)^{-1} - \nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_{\omega_k}(t), \omega_k)^{-1}\|_2 \|\omega(t)\| dt \\ & + \|C\|_2 \int_0^T \|\nabla_{\xi\xi}^2 \mathcal{V}(t, \hat{x}_{\omega_k}(t), \omega_k)^{-1}\|_2 \|\omega(t) - \omega_k(t)\| dt. \end{aligned}$$

With the uniform bound of $\|\nabla_{\xi\xi}^2 \mathcal{V}(t, \widehat{x}_{\omega_k}(t), \omega_k)^{-1}\|_2$ from Lemma B.7 and the L^2 -convergence of ω_k the second summand converges to zero. Lemma B.7 further implies that for all $k \in \mathbb{N}$ and $t \in [0, T]$ it holds

$$\|\nabla_{\xi\xi}^2 \mathcal{V}(t, \widehat{x}_\omega(t), \omega)^{-1} - \nabla_{\xi\xi}^2 \mathcal{V}(t, \widehat{x}_{\omega_k}(t), \omega_k)^{-1}\|_2 \leq 2M\gamma.$$

Since \mathcal{V} is C^∞ in ξ and ω , the dominated convergence theorem yields that the first summand also converges to zero. The convergence of the term involving $-C\widehat{x}_\omega$ instead of ω can be shown analogously. Hence it is shown that the weak derivative of \widehat{x}_ω is given by

$$h(t, \widehat{x}_\omega(t)) + \alpha \nabla_{\xi\xi}^2 \mathcal{V}(t, \widehat{x}_\omega(t), \omega) C^\top (\omega(t) - C\widehat{x}_\omega(t)) \in L^2(0, T; \mathbb{R}^n).$$

By definition it holds $\widehat{x}_\omega(0) = 0$, therefore $\widehat{x}_\omega \in H^1(0, T; \mathbb{R}^n)$ is a weak solution of (3.1). \square

7. CONCLUSION

In this work we were able to show that for quadratic systems the formal derivation of the observer equation associated with the Mortensen observer can be done in a rigorous manner, assuming that the initial model is close enough to the actual system from which the data is measured. The assumption of a quadratic right hand side simplifies the presentation of this technical work and can be relaxed using the techniques presented here. The strategies applied mainly require the right hand side to have locally Lipschitz and finitely many non vanishing derivatives. The results can therefore be generalized to dynamics with polynomial right hand sides. By taking advantage of a series representation it might be possible to further extend the results to analytic right hand sides. This, however, would require additional technical efforts to ensure that all bounds and regularities hold. Our findings motivate and justify two approaches to approximate the non-linear observer numerically. On the one hand the observer trajectory can be obtained by a minimization of the value function in every time point, on the other hand it is given as a solution of the observer equation.

APPENDIX A. LINEAR QUADRATIC CONTROL PROBLEM

This section of the Appendix is concerned with a general linear quadratic control problem. Its solution is characterized as the solution of a system of equations. Further the optimizing triple of state, control and adjoint is estimated by the data. The control problem is given as

$$\min_{(x,v) \in V_t \times L^2(0,t;\mathbb{R}^m)} J(x,v), \text{ subject to } e(x,v) = 0, \quad (\text{A.1})$$

where the cost functional $J: V_t \times L^2(0,t;\mathbb{R}^m) \rightarrow \mathbb{R}$ and the constraint $e: V_t \times L^2(0,t;\mathbb{R}^m) \rightarrow L^2(0,t;\mathbb{R}^n) \times \mathbb{R}^n$ are defined as

$$\begin{aligned} J(x,v) &= \frac{1}{2} \|x(0) - b\|^2 + \frac{1}{2} \|v\|_{L^2(0,t;\mathbb{R}^m)}^2 + \frac{\alpha}{2} \|Cx - \mu\|_{L^2(0,t;\mathbb{R}^r)}^2 + \int_0^t \langle G(x \otimes x), \rho \rangle + \langle l_1, x \rangle + \langle l_2, v \rangle \, ds, \\ e(x,v) &= (\dot{x} - Ax - G(\check{x}_t \otimes I_n)x - G(I_n \otimes \check{x}_t)x - Fv - f, x(t) - a). \end{aligned} \quad (\text{A.2})$$

Here $t \in (0, T]$ and the data consisting of $a, b \in \mathbb{R}^n$, $\mu \in L^2(0,t;\mathbb{R}^r)$, $f \in L^2(0,t;\mathbb{R}^n)$ and $l_1 \in L^2(0,t;\mathbb{R}^n)$, $l_2 \in L^2(0,t;\mathbb{R}^m)$ is given.

Proposition A.1. *Assume that $\check{x}_s \in V_s$ is well-defined for any $s \in (0, T]$ and that $\|\check{x}_s\|_{L^\infty(0,s;\mathbb{R}^n)} \leq \kappa$ for some time-independent constant $\kappa > 0$ and all $s \in (0, T]$. Further assume that $\rho \in V_t$ satisfies*

$\|\rho\|_{L^\infty(0,t;\mathbb{R}^n)} \leq \beta := (4\bar{c}(\kappa)^2\|G\|_2)^{-1}$. Then the system

$$\begin{aligned}
\dot{x} &= Ax + G(\check{x}_t \otimes I_n)x + G(I_n \otimes \check{x}_t)x + Fv + f, \\
x(t) &= a, \\
\dot{q} &= -A^\top q - (\check{x}_t^\top \otimes I_n)G^\top q - (I_n \otimes \check{x}_t^\top)G^\top q - \alpha C^\top(Cx - \mu) \\
&\quad - (x^\top \otimes I_n)G^\top \rho - (I_n \otimes x^\top)G^\top \rho - l_1, \\
q(0) &= b - x(0), \\
v + F^\top q + l_2 &= 0,
\end{aligned} \tag{A.3}$$

admits exactly one solution $(\bar{x}, \bar{v}, \bar{q})$. It is given by the minimizing control of (A.1) and its corresponding state and adjoint state. Further there exists a constant $\bar{M} > 0$ independent of $t, a, b, \mu, f, l_1, l_2, \rho$ und \check{x}_t such that

$$\begin{aligned}
&\max \left(\|\bar{v}\|_{L^2(0,t;\mathbb{R}^m)}, \|\bar{x}\|_{V_t}, \|\bar{q}\|_{V_t} \right) \\
&\leq \bar{M} \left(\|a\| + \|b\| + \|\mu\|_{L^2(0,t;\mathbb{R}^m)} + \|f\|_{L^2(0,t;\mathbb{R}^n)} + \|l_1\|_{L^2(0,t;\mathbb{R}^n)} + \|l_2\|_{L^2(0,t;\mathbb{R}^m)} \right).
\end{aligned}$$

Proof. Since the control problem is closely related to the well-studied linear-quadratic control problem, the proof will only be sketched. To see unique solvability of (A.1) it needs to be shown that the cost functional is coercive and convex along the kernel of the constraint. This condition is clearly fulfilled for all linear and non-negative quadratic terms. The upper bound on $\|\rho\|_{L^\infty(0,t;\mathbb{R}^n)}$ ensures that both properties hold for the entire cost functional. The necessary and sufficient optimality condition for (x, v, q) to be the minimizer is then given by (A.3).

The norm of the control \bar{v} can then be estimated from above in terms of the cost functional, which in turn is estimated from above by the cost functional evaluated in the zero control and its corresponding trajectory. These estimates can be carried over to the trajectory and adjoint state associated with \bar{v} . \square

APPENDIX B. TECHNICAL AUXILIARY RESULTS

This section collects auxiliary results and their technical proofs required in this work.

Lemma B.1. *Let $t \in (0, T]$ and $(\xi, \omega) \in \mathcal{U}_{\delta_4}(0) \subset \mathbb{R}^n \times L^2(0, t; \mathbb{R}^r)$. Let $\mu \in L^2(0, t; \mathbb{R}^r)$ and $z_1, z_2, z_3, z_4 \in \mathbb{R}^n$ and for $i = 1, \dots, 4$ denote $z^i = (z_1, \dots, z_i)$. Then there exist constants $\bar{M}_1, \bar{M}_2, \bar{M}_3, \bar{M}_4 > 0$ and $\check{M}_0, \check{M}_1, \check{M}_2 > 0$ independent of $t, \xi, \omega, \mu, z_1, z_2, z_3$ and z_4 such that*

$$\max \left(\|D_{\xi^i}^i \mathcal{X}_t(\xi, \omega) z^i\|_{V_t}, \|D_{\xi^i}^i \mathcal{P}_t(\xi, \omega) z^i\|_{V_t}, \|D_{\xi^i}^i \mathcal{U}_t(\xi, \omega) z^i\|_{L^2(0,t;\mathbb{R}^m)} \right) \leq \bar{M}_i \prod_{k=1}^i \|z_k\|$$

holds for $i = 1, 2, 3, 4$. Further it holds

$$\max \left(\|D_\omega \mathcal{X}_t(\xi, \omega) \mu\|_{V_t}, \|D_\omega \mathcal{P}_t(\xi, \omega) \mu\|_{V_t}, \|D_\omega \mathcal{U}_t(\xi, \omega) \mu\|_{L^2(0,t;\mathbb{R}^m)} \right) \leq \check{M}_0 \|\mu\|_{L^2(0,t;\mathbb{R}^r)},$$

and for $i = 1, 2$ it holds

$$\begin{aligned}
&\max \left(\|D_{\omega \xi^i}^{i+1} \mathcal{X}_t(\xi, \omega)(\mu, z^i)\|_{V_t}, \|D_{\omega \xi^i}^{i+1} \mathcal{P}_t(\xi, \omega)(\mu, z^i)\|_{V_t}, \|D_{\omega \xi^i}^{i+1} \mathcal{U}_t(\xi, \omega)(\mu, z^i)\|_{L^2(0,t;\mathbb{R}^m)} \right) \\
&\leq \check{M}_i \|\mu\|_{L^2(0,t;\mathbb{R}^r)} \prod_{k=1}^i \|z_k\|.
\end{aligned}$$

Proof. First note that it holds $\|\mathcal{X}_t(\xi, \omega)\|_{L^\infty(0,t;\mathbb{R}^n)} \leq M_1\delta_4 + \|\tilde{x}\|_{L^\infty(0,T;\mathbb{R}^n)}$ and $\|\mathcal{P}_t(\xi, \omega)\|_{L^\infty(0,t;\mathbb{R}^n)} \leq M_2\delta_4 \leq \beta$. To enforce the second estimate a reduction of δ_4 might be necessary. For the sake of a better overview we abstain from introducing a new variable to reflect this. By construction it holds

$$\Phi_t(\mathcal{X}_t(\xi, \omega), \mathcal{U}_t(\xi, \omega), \mathcal{P}_t(\xi, \omega), \omega) = (\omega, 0, \xi, 0, 0, 0). \quad (\text{B.1})$$

Since Φ_t is C^∞ on X_t and $\mathcal{X}_t, \mathcal{U}_t$ and \mathcal{P}_t are C^∞ on $\mathcal{U}_{\delta_4}(0)$, it follows

$$D_\xi(\Phi_t(\mathcal{X}_t(\xi, \omega), \mathcal{U}_t(\xi, \omega), \mathcal{P}_t(\xi, \omega), \omega)) z_1 = (0, 0, z_1, 0, 0, 0).$$

With the chain rule one gets that the triple $(D_\xi \mathcal{X}_t(\xi, \omega) z_1, D_\xi \mathcal{U}_t(\xi, \omega) z_1, D_\xi \mathcal{P}_t(\xi, \omega) z_1)$ solves (A.3) with $\check{x}_t = \mathcal{X}_t(\xi, \omega)$, $\rho = \mathcal{P}_t(\xi, \omega)$, $f = 0$, $a = z_1$, $\mu = 0$, $l_1 = 0$, $b = 0$, $l_2 = 0$. Proposition A.1 yields the estimate for the first order ξ -derivatives. Taking two ξ -derivatives of (B.1) shows that $(D_{\xi^2} \mathcal{X}_t(\xi, \omega) z^2, D_{\xi^2} \mathcal{U}_t(\xi, \omega) z^2, D_{\xi^2} \mathcal{P}_t(\xi, \omega) z^2)$ solves (A.3) with

$$\begin{aligned} \check{x}_t &= \mathcal{X}_t(\xi, \omega), \quad \rho = \mathcal{P}_t(\xi, \omega), \quad a = 0, \quad \mu = 0, \quad b = 0, \quad l_2 = 0, \\ f &= G(D_\xi \mathcal{X}_t(\xi, \omega) z_1 \otimes D_\xi \mathcal{X}_t(\xi, \omega) z_2) + G(D_\xi \mathcal{X}_t(\xi, \omega) z_2 \otimes D_\xi \mathcal{X}_t(\xi, \omega) z_1), \\ l_1 &= (D_\xi \mathcal{X}_t(\xi, \omega) z_1^\top \otimes I_n) G^\top \mathcal{P}_t(\xi, \omega) z_2 + (I_n \otimes D_\xi \mathcal{X}_t(\xi, \omega) z_1^\top) G^\top \mathcal{P}_t(\xi, \omega) z_2 \\ &\quad + (D_\xi \mathcal{X}_t(\xi, \omega) z_2^\top \otimes I_n) G^\top \mathcal{P}_t(\xi, \omega) z_1 + (I_n \otimes D_\xi \mathcal{X}_t(\xi, \omega) z_2^\top) G^\top \mathcal{P}_t(\xi, \omega) z_1. \end{aligned}$$

Again the desired estimates are obtained with Proposition A.1. The remaining estimates are shown analogously. \square

Lemma B.2. *Let $(\xi, \omega) \in \mathcal{U}_{\frac{1}{2}\delta_4}(0) \subset \mathbb{R}^n \times L^2(0, T; \mathbb{R}^r)$. Then for all $t \in (0, T)$ it holds*

$$\frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t)\| \, ds \rightarrow 0, \quad \tau \searrow 0.$$

Proof. Let $t \in (0, T)$ be fixed and let $\tau > 0$ be small enough such that $t + \tau < T$. Define $\eta(z) = \mathcal{X}_{t+\tau}(\xi, \omega)(t + \tau - z)$. Then

$$\dot{\eta}(z) = -A\eta(z) - G(\eta(z) \otimes \eta(z)) - F\mathcal{U}_{t+\tau}(\xi, \omega)(t + \tau - z). \quad (\text{B.2})$$

Note that for all $z \in (0, t + \tau - s)$ it holds

$$\langle \dot{\eta}(z), \eta(z) - \mathcal{X}_t(\xi, \omega)(t) \rangle = \frac{d}{dz} \frac{1}{2} \|\eta(z) - \mathcal{X}_t(\xi, \omega)(t)\|^2.$$

Hence testing (B.2) with $\eta(z) - \mathcal{X}_t(\xi, \omega)(t)$ and integrating over $(0, t + \tau - s)$ for some fixed $s \in [t, t + \tau)$ yields

$$\begin{aligned} \frac{1}{2} \|\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t)\|^2 &= \frac{1}{2} \|\eta(t + \tau - s) - \mathcal{X}_t(\xi, \omega)(t)\|^2 \\ &\leq \frac{1}{2} \|\eta(0) - \mathcal{X}_t(\xi, \omega)(t)\|^2 + \int_0^{t+\tau-s} (\|A\|_2 \|\eta\| + \|G\|_2 \|\eta\|^2 + \|F\|_2^2 \|\mathcal{P}_{t+\tau}(\xi, \omega)\|) (\|\eta\| + \|\xi\| + \|\tilde{x}(t)\|) \, ds \\ &\leq \frac{1}{2} \|\tilde{x}(t + \tau) - \tilde{x}(t)\|^2 + \tau \left(\|A\|_2 \|\eta\|_{L^\infty(0, t+\tau; \mathbb{R}^n)} + \|G\|_2 \|\eta\|_{L^\infty(0, t+\tau; \mathbb{R}^n)}^2 + \|F\|_2^2 \|\mathcal{P}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)} \right) \\ &\quad (\|\eta\|_{L^\infty(0, t+\tau; \mathbb{R}^n)} + \|\xi\| + \|\tilde{x}(t)\|). \end{aligned}$$

Since all the norms on the right hand side can be bounded uniformly in τ , there exists some constant $c > 0$ such that for all $s \in [t, t + \tau)$ it holds

$$\frac{1}{2} \|\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t)\| \leq \frac{1}{\sqrt{2}} \|\tilde{x}(t + \tau) - \tilde{x}(t)\| + \sqrt{\tau}c.$$

Hence

$$\begin{aligned} & \frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t)\| \, ds \\ & \leq \frac{1}{\tau} \int_t^{t+\tau} \sqrt{2} \|\tilde{x}(t + \tau) - \tilde{x}(t)\| + 2\sqrt{\tau}c \, ds = \sqrt{2} \|\tilde{x}(t + \tau) - \tilde{x}(t)\| + 2\sqrt{\tau}c. \end{aligned}$$

The continuity of \tilde{x} yields the assertion. \square

Lemma B.3. *Let $(\xi, \omega) \in \mathcal{U}_{\frac{1}{2}\delta_4}(0) \subset \mathbb{R}^n \times L^2(0, T; \mathbb{R}^r)$ and assume ω to be continuous. Then for all $t \in (0, T)$ it holds*

$$\frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{P}_{t+\tau}(\xi, \omega)(s) - \mathcal{P}_t(\xi, \omega)(t)\| \, ds \rightarrow 0, \quad \tau \searrow 0.$$

Proof. The proof is done analogously to the proof of Lemma B.2 using the transformation $\eta(z) = \mathcal{P}_{t+\tau}(\xi, \omega)(t + \tau - z)$. The initial term of this transformation does not lead to a term involving \tilde{x} but instead $\mathcal{P}_{t+\tau}(\xi, \omega)(t + \tau) - \mathcal{P}_t(\xi, \omega)(t)$. With Lemma 4.5 this is equal to $-\nabla_\xi \mathcal{V}(t + \tau, \xi, \omega) + \nabla_\xi \mathcal{V}(t, \xi, \omega)$. Proposition 4.8 ensures the convergence to zero. \square

Lemma B.4. *Let $(\xi, \omega) \in \mathcal{U}_{\frac{1}{2}\delta_4}(0) \subset \mathbb{R}^n \times L^2(0, T; \mathbb{R}^r)$ and assume ω to be continuous. Then for all $t \in (0, T)$ it holds*

$$\frac{1}{\tau} \int_t^{t+\tau} \dot{\mathcal{X}}_{t+\tau}(\xi, \omega)(s) \, ds \rightarrow \dot{\mathcal{X}}_t(\xi, \omega)(t), \quad \tau \searrow 0.$$

Proof. Let $t \in (0, T)$ be fixed and let $\tau > 0$ be small enough such that $t + \tau < T$. Then

$$\begin{aligned} & \left\| \frac{1}{\tau} \int_t^{t+\tau} \dot{\mathcal{X}}_{t+\tau}(\xi, \omega)(s) \, ds - \dot{\mathcal{X}}_t(\xi, \omega)(t) \right\| \\ & \leq \frac{1}{\tau} \int_t^{t+\tau} \|A(\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t))\| + \|F(\mathcal{U}_{t+\tau}(\xi, \omega)(s) - \mathcal{U}_t(\xi, \omega)(t))\| \\ & \quad + \|G(\mathcal{X}_{t+\tau}(\xi, \omega)(s) \otimes (\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t)))\| + \|G((\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t)) \otimes \mathcal{X}_t(\xi, \omega)(t))\| \, ds \\ & \leq (\|A\|_2 + \|G\|_2 \|\mathcal{X}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)} + \|G\|_2 \|\mathcal{X}_t(\xi, \omega)(t)\|) \\ & \quad \frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t)\| \, ds + \|F\|_2^2 \frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{P}_{t+\tau}(\xi, \omega)(s) - \mathcal{P}_t(\xi, \omega)(t)\| \, ds. \end{aligned}$$

The assertion follows with Lemma B.2, Lemma B.3 and the uniform bound on $\|\mathcal{X}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)}$ derived from Proposition 3.4. \square

Lemma B.5. *Let $(\xi, \omega) \in \mathcal{U}_{\frac{1}{2}\delta_4}(0) \subset \mathbb{R}^n \times L^2(0, T; \mathbb{R}^r)$ and assume ω to be continuous. Then for all $t \in (0, T)$ and $\tau \searrow 0$ it holds*

$$\frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{U}_{t+\tau}(\xi, \omega)\|^2 + \alpha \|\omega - C(\mathcal{X}_{t+\tau}(\xi, \omega) - \tilde{x})\|^2 ds \rightarrow \|\mathcal{U}_t(\xi, \omega)(t)\|^2 + \alpha \|\omega(t) - C(\mathcal{X}_t(\xi, \omega)(t) - \tilde{x}(t))\|^2.$$

Proof. Let $t \in (0, T)$ be fixed and let $\tau > 0$ be small enough such that $t + \tau < T$. For the first summand consider

$$\begin{aligned} & \left| \frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{U}_{t+\tau}(\xi, \omega)(s)\|^2 ds - \|\mathcal{U}_t(\xi, \omega)(t)\|^2 \right| \\ & \leq \frac{1}{\tau} \int_t^{t+\tau} |\langle \mathcal{U}_{t+\tau}(\xi, \omega)(s) + \mathcal{U}_t(\xi, \omega)(t), \mathcal{U}_{t+\tau}(\xi, \omega)(s) - \mathcal{U}_t(\xi, \omega)(t) \rangle| ds \\ & \leq (\|F\|_2 \|\mathcal{P}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)} + \|\mathcal{U}_t(\xi, \omega)(t)\|) \|F\|_2 \frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{P}_{t+\tau}(\xi, \omega)(s) - \mathcal{P}_t(\xi, \omega)(t)\| ds. \end{aligned}$$

Lemma B.3 and the uniform bound on $\|\mathcal{P}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)}$ derived from Proposition 3.5 yield the convergence. For the second summand consider

$$\begin{aligned} & \left| \frac{1}{\tau} \int_t^{t+\tau} \|\omega(s) - C\mathcal{X}_{t+\tau}(\xi, \omega)(s) + C\tilde{x}(s)\|^2 ds - \|\omega(t) - C\mathcal{X}_t(\xi, \omega)(t) + C\tilde{x}(t)\|^2 \right| \\ & = \left| \frac{1}{\tau} \int_t^{t+\tau} \|\omega(s)\|^2 + \|C\tilde{x}(s)\|^2 + 2\langle \omega(s), C\tilde{x}(s) \rangle ds - \|\omega(t)\|^2 - \|C\tilde{x}(t)\|^2 - 2\langle \omega(t), C\tilde{x}(t) \rangle \right| \\ & + \left| \frac{1}{\tau} \int_t^{t+\tau} \|C\mathcal{X}_{t+\tau}(\xi, \omega)(s)\|^2 - \|C\mathcal{X}_t(\xi, \omega)(t)\|^2 - 2\langle \omega(s), C\mathcal{X}_{t+\tau}(\xi, \omega)(s) \rangle \right. \\ & \left. + 2\langle \omega(t), C\mathcal{X}_t(\xi, \omega)(t) \rangle - 2\langle C\tilde{x}(s), C\mathcal{X}_{t+\tau}(\xi, \omega)(s) \rangle + 2\langle C\tilde{x}(t), C\mathcal{X}_t(\xi, \omega)(t) \rangle ds \right|. \end{aligned}$$

Since ω and \tilde{x} are continuous, the first two summands of the right hand side converge to zero for $\tau \searrow 0$. The remaining part is smaller or equal to

$$\begin{aligned} & \frac{1}{\tau} \int_t^{t+\tau} \left(|\langle C\mathcal{X}_{t+\tau}(\xi, \omega)(s) + C\mathcal{X}_t(\xi, \omega)(t), C\mathcal{X}_{t+\tau}(\xi, \omega)(s) - C\mathcal{X}_t(\xi, \omega)(t) \rangle| \right. \\ & + 2|\langle \omega(s), C\mathcal{X}_{t+\tau}(\xi, \omega)(s) - C\mathcal{X}_t(\xi, \omega)(t) \rangle| + 2|\langle \omega(t) - \omega(s), C\mathcal{X}_t(\xi, \omega)(t) \rangle| \\ & \left. + 2|\langle C\mathcal{X}_{t+\tau}(\xi, \omega)(s) - C\mathcal{X}_t(\xi, \omega)(t), C\tilde{x}(s) \rangle| + 2|\langle C\mathcal{X}_t(\xi, \omega)(t), C\tilde{x}(t) - C\tilde{x}(s) \rangle| \right) ds \\ & \leq \|C\|_2^2 (\|\mathcal{X}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)} + \|\mathcal{X}_t(\xi, \omega)(t)\|) \frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t)\| ds \\ & + 2\|\omega\|_{L^\infty(0, T; \mathbb{R}^r)} \|C\|_2 \frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t)\| ds \\ & + 2\|C\|_2 \|\mathcal{X}_t(\xi, \omega)(t)\| \left(\frac{1}{\tau} \int_t^{t+\tau} \|\omega(s) - \omega(t)\| ds \right) 2\|C\|_2^2 \|\mathcal{X}_t(\xi, \omega)(t)\| \left(\frac{1}{\tau} \int_t^{t+\tau} \|\tilde{x}(s) - \tilde{x}(t)\| ds \right) \\ & + 2\|C\|_2^2 \|\tilde{x}\|_{L^\infty(0, T; \mathbb{R}^n)} \frac{1}{\tau} \int_t^{t+\tau} \|\mathcal{X}_{t+\tau}(\xi, \omega)(s) - \mathcal{X}_t(\xi, \omega)(t)\| ds. \end{aligned}$$

The assertion follows with the uniform bound on $\|\mathcal{X}_{t+\tau}(\xi, \omega)\|_{L^\infty(0, t+\tau; \mathbb{R}^n)}$, Lemma B.2 and the continuity of ω and \tilde{x} . \square

Lemma B.6. *Let $\omega \in L^2(0, T; \mathbb{R}^r)$ be continuous and satisfy $\|\omega\|_{L^2(0, T; \mathbb{R}^r)} < \frac{1}{2}\delta_4$. Then it holds:*

- (i) *The mappings $(t, \xi) \mapsto \nabla_\xi \mathcal{V}(t, \xi, \omega)$, $(t, \xi) \mapsto \nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega)$ and $(t, \xi) \mapsto \nabla_{\xi\xi\xi}^3 \mathcal{V}(t, \xi, \omega)$ are continuous from $[0, T] \times \mathcal{U}_{\frac{1}{2}\delta_4}(0)$ to their respective image spaces \mathbb{R}^n , $\mathbb{R}^{n,n}$ and $L(\mathbb{R}^n, \mathbb{R}^{n,n})$.*
- (ii) *For any fixed $\xi \in \mathcal{U}_{\frac{1}{2}\delta_4}(0)$ and any $i \in \{1, \dots, n\}$ the function $s \mapsto \frac{\partial}{\partial \xi_i} \mathcal{V}(s, \xi, \omega)$ is differentiable in $(0, T)$ and it holds*

$$\partial_t \frac{\partial}{\partial \xi_i} \mathcal{V}(t, \xi, \omega) = \frac{\partial}{\partial \xi_i} \partial_t \mathcal{V}(t, \xi, \omega).$$

- (iii) *For any fixed $\xi \in \mathcal{U}_{\frac{1}{2}\delta_4}(0)$ and any $i, j \in \{1, \dots, n\}$ the function $s \mapsto \frac{\partial^2}{\partial \xi_j \partial \xi_i} \mathcal{V}(s, \xi, \omega)$ is differentiable in $(0, T)$ and it holds*

$$\partial_t \frac{\partial^2}{\partial \xi_j \partial \xi_i} \mathcal{V}(t, \xi, \omega) = \frac{\partial^2}{\partial \xi_j \partial \xi_i} \partial_t \mathcal{V}(t, \xi, \omega).$$

Proof. (i) We will show continuity for the partial derivatives which then implies (i). To that end let $(t, \xi) \in [0, T] \times \mathcal{U}_{\frac{1}{2}\delta_4}(0)$ be arbitrary. Then for all $(\Delta t, \Delta \xi)$ such that $(t + \Delta t, \xi + \Delta \xi) \in [0, T] \times \mathcal{U}_{\frac{1}{2}\delta_4}(0)$ and any $i \in \{1, \dots, n\}$ it holds

$$\begin{aligned} & \left| \frac{\partial}{\partial \xi_i} \mathcal{V}(t + \Delta t, \xi + \Delta \xi, \omega) - \frac{\partial}{\partial \xi_i} \mathcal{V}(t, \xi, \omega) \right| \\ & \leq \left| \frac{\partial}{\partial \xi_i} \mathcal{V}(t + \Delta t, \xi + \Delta \xi, \omega) - \frac{\partial}{\partial \xi_i} \mathcal{V}(t + \Delta t, \xi, \omega) \right| + \left| \frac{\partial}{\partial \xi_i} \mathcal{V}(t + \Delta t, \xi, \omega) - \frac{\partial}{\partial \xi_i} \mathcal{V}(t, \xi, \omega) \right|. \end{aligned}$$

Due to the time continuity of the gradient shown in Proposition 4.8 the second summand converges to zero for $\Delta t \rightarrow 0$. Taylor's Theorem implies existence of $\theta \in [0, 1]$ such that the first summand is equal to

$$\left| \nabla_\xi \left[\frac{\partial}{\partial \xi_i} \mathcal{V} \right] (t + \Delta t, \xi + \theta \Delta \xi, \omega) \Delta \xi \right|.$$

For sufficiently small $\Delta \xi$ Proposition 4.6 can be used to estimate this expression by $\tilde{M}_2 \|\Delta \xi\|$, which converges to zero for $\Delta \xi \rightarrow 0$. Utilizing the results from Proposition 4.6 and Proposition 4.8 the continuity of the partial derivatives of order two and three can be shown analogously.

(iii) Let $\epsilon > 0$ be small enough such that for any $x, y \in (-\epsilon, \epsilon)$ it holds $\|xe_i + ye_j + \xi\| < \frac{1}{2}\delta_4$. Define the function

$$f: (0, T) \times (-\epsilon, \epsilon)^2 \rightarrow \mathbb{R}, \quad (t, x, y) \mapsto \mathcal{V}(t, xe_i + ye_j + \xi, \omega).$$

Step 1:

Show that for any fixed $\bar{t} \in (0, T)$ and $y \in (-\epsilon, \epsilon)$ it holds

$$\frac{\partial^2}{\partial x \partial t} f(\bar{t}, 0, y) = \frac{\partial^2}{\partial t \partial x} f(\bar{t}, 0, y).$$

Let $t \in (0, T)$ and $x \in (-\epsilon, \epsilon)$ be arbitrary and note that due to Corollary 4.4 and Theorem 4.9

$$\frac{\partial}{\partial x} f(t, x, y) = \frac{\partial}{\partial \xi_i} \mathcal{V}(t, xe_i + ye_j + \xi, \omega) \quad \text{and} \quad \frac{\partial}{\partial t} f(t, x, y) = \partial_t \mathcal{V}(t, xe_i + ye_j + \xi, \omega)$$

exist. Taking a partial derivative of (4.8) and noting that \mathcal{V} is C^∞ with respect to ξ in $\mathcal{U}_{\frac{1}{2}\delta_4}(0)$ ensures the existence of

$$\begin{aligned} \frac{\partial^2}{\partial x \partial t} f(t, x, y) &= \frac{\partial}{\partial \xi_i} \partial_t \mathcal{V}(t, xe_i + ye_j + \xi, \omega) = - \left\langle \frac{\partial}{\partial \xi_i} \nabla_\xi \mathcal{V}(t, xe_i + ye_j + \xi, \omega), h(t, xe_i + ye_j + \xi) \right\rangle \\ &\quad - \left\langle \nabla_\xi \mathcal{V}(t, xe_i + ye_j + \xi, \omega), \frac{\partial}{\partial \xi_i} h(t, xe_i + ye_j + \xi) \right\rangle \\ &\quad - \left\langle F^\top \frac{\partial}{\partial \xi_i} \nabla_\xi \mathcal{V}(t, xe_i + ye_j + \xi, \omega), F^\top \nabla_\xi \mathcal{V}(t, xe_i + ye_j + \xi, \omega) \right\rangle \\ &\quad + \alpha \langle C^\top C e_i, xe_i + ye_j + \xi \rangle - \alpha \langle C^\top \omega(t), e_i \rangle. \end{aligned} \tag{B.3}$$

It remains to show that this function is continuous in $(t, x) = (\bar{t}, 0)$, where $\bar{t} \in (0, T)$ is arbitrary. This will be achieved by an examination of the terms on the right hand side. With the assumption ϵ (i) yields the continuity of the first and second order spatial derivatives. Continuity of the remaining terms in (B.3) can be seen easily. Hence [39], Theorem 9.41, yields the existence of $\frac{\partial^2}{\partial t \partial x} f(\bar{t}, 0, y)$ and it holds $\frac{\partial^2}{\partial x \partial t} f(\bar{t}, 0, y) = \frac{\partial^2}{\partial t \partial x} f(\bar{t}, 0, y)$. The derivative of the HJB equation ensures that $\frac{\partial^2}{\partial x \partial t} f(\bar{t}, 0, y)$ is differentiable in $(-\epsilon, \epsilon)$ with respect to y . Since the derived equation holds for all $y \in (-\epsilon, \epsilon)$ it follows

$$\frac{\partial^3}{\partial y \partial x \partial t} f(\bar{t}, 0, y) = \frac{\partial^3}{\partial y \partial t \partial x} f(\bar{t}, 0, y). \tag{B.4}$$

Step 2:

Show that for any $\bar{t} \in (0, T)$ it holds

$$\frac{\partial^3}{\partial y \partial t \partial x} f(\bar{t}, 0, 0) = \frac{\partial^3}{\partial t \partial y \partial x} f(\bar{t}, 0, 0). \tag{B.5}$$

This will be achieved by an application of [39], Theorem 9.41 to the function

$$(0, T) \times (-\epsilon, \epsilon) \rightarrow \mathbb{R}, \quad (t, y) \mapsto \frac{\partial}{\partial x} f(t, 0, y).$$

Let $t \in (0, T)$ and $y \in (-\epsilon, \epsilon)$ be arbitrary. Since \mathcal{V} is C^∞ with respect to ξ on an appropriate neighborhood of zero, the partial derivative

$$\frac{\partial}{\partial y} \frac{\partial}{\partial x} f(t, 0, y) = \frac{\partial^2}{\partial \xi_j \partial \xi_i} \mathcal{V}(t, ye_j + \xi, \omega)$$

exists. Furthermore Step 1 ensures the existence of the two partial derivatives

$$\frac{\partial}{\partial t} \frac{\partial}{\partial x} f(t, 0, y) \quad \text{and} \quad \frac{\partial^2}{\partial y \partial t} \frac{\partial}{\partial x} f(t, 0, y).$$

It remains to show that the function

$$(t, y) \mapsto \frac{\partial^3}{\partial y \partial t \partial x} f(t, 0, y)$$

is continuous in $(t, y) = (\bar{t}, 0)$, where $\bar{t} \in (0, T)$ is arbitrary. To that end form derivatives with respect to ξ_i and ξ_j of (4.8) to obtain

$$\begin{aligned} \frac{\partial^3}{\partial y \partial t \partial x} f(t, 0, y) &= \frac{\partial^3}{\partial y \partial x \partial t} f(t, 0, y) = \frac{\partial^2}{\partial \xi_j \partial \xi_i} \partial_t \mathcal{V}(t, ye_j + \xi, \omega) \\ &= - \left\langle \frac{\partial^2}{\partial \xi_j \partial \xi_i} \nabla_\xi \mathcal{V}(t, ye_j + \xi, \omega), h(t, ye_j + \xi) \right\rangle - \left\langle \frac{\partial}{\partial \xi_i} \nabla_\xi \mathcal{V}(t, ye_j + \xi, \omega), \frac{\partial}{\partial \xi_j} h(t, ye_j + \xi) \right\rangle \\ &\quad - \left\langle \frac{\partial}{\partial \xi_j} \nabla_\xi \mathcal{V}(t, ye_j + \xi, \omega), \frac{\partial}{\partial \xi_i} h(t, ye_j + \xi) \right\rangle - \left\langle \nabla_\xi \mathcal{V}(t, ye_j + \xi, \omega), \frac{\partial^2}{\partial \xi_j \partial \xi_i} h(t, ye_j + \xi) \right\rangle \quad (\text{B.6}) \\ &\quad - \left\langle F^\top \frac{\partial^2}{\partial \xi_j \partial \xi_i} \nabla_\xi \mathcal{V}(t, ye_j + \xi, \omega), F^\top \nabla_\xi \mathcal{V}(t, ye_j + \xi, \omega) \right\rangle \\ &\quad - \left\langle F^\top \frac{\partial}{\partial \xi_i} \nabla_\xi \mathcal{V}(t, ye_j + \xi, \omega), F^\top \frac{\partial}{\partial \xi_j} \nabla_\xi \mathcal{V}(t, ye_j + \xi, \omega) \right\rangle + \alpha \langle C^\top C e_i, e_j \rangle \end{aligned}$$

and consider the term on the right hand side. Their continuity can be shown analogously to Step 1. Therefore $(t, y) \mapsto \frac{\partial^3}{\partial y \partial t \partial x} f(t, 0, y)$ is continuous in $(\bar{t}, 0)$ and [39], Theorem 9.41 implies (B.5).

Step 3:

Now (B.4) with $y = 0$ together with (B.5) yields

$$\frac{\partial^2}{\partial \xi_j \partial \xi_i} \partial_t \mathcal{V}(\bar{t}, \xi, \omega) = \frac{\partial^3}{\partial y \partial x \partial t} f(\bar{t}, 0, 0) = \frac{\partial^3}{\partial t \partial y \partial x} f(\bar{t}, 0, 0) = \partial_t \frac{\partial^2}{\partial \xi_j \partial \xi_i} \mathcal{V}(\bar{t}, \xi, \omega).$$

The proof for (ii) can be done using the same strategies used for showing (iii) and is omitted here. \square

Lemma B.7. *Let $t \in [0, T]$ and let $(\xi, \omega) \in \mathcal{U}_{\delta_5}(0) \subset \mathbb{R}^n \times L^2(0, t; \mathbb{R}^r)$. Then there exists a constant $M_{\mathcal{V}} > 0$ independent of t , ξ and ω such that*

$$\|\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega)^{-1}\|_2 < M_{\mathcal{V}}.$$

Proof. First note that δ_5 was chosen such that $\|\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega) - \nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)\|_2 \leq \frac{1}{2} \|\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)^{-1}\|_2^{-1}$ holds for all (ξ, ω) satisfying $\max(\|\xi\|, \|\omega\|_{L^2(0, t; \mathbb{R}^r)}) < \delta_5$. It follows

$$\|\nabla_{\xi\xi}^2 \mathcal{V}(t, \xi, \omega)^{-1}\|_2 < 2 \|\nabla_{\xi\xi}^2 \mathcal{V}(t, 0, 0)^{-1}\|_2.$$

This can be seen by utilizing the Neumann series, see for example [34], Proposition 1 in Section 3.5 and [38], Proposition 7 in Section 1.23. As mentioned in the proof of Theorem 5.3 the mapping $s \mapsto \nabla_{\xi\xi}^2 \mathcal{V}(s, 0, 0)^{-1}$ is continuous in $[0, T]$. Hence the right hand side admits a maximum and setting $M_{\mathcal{V}} = 2 \max_{s \in [0, T]} \|\nabla_{\xi\xi}^2 \mathcal{V}(s, 0, 0)^{-1}\|_2$ yields the assertion. \square

Lemma B.8. *Let $t \in [0, T]$. Then there exists a constant $L_{\mathcal{V}} > 0$ independent of t such that for all $x, y \in \mathbb{R}^n$ and $\omega \in L^2(0, t; \mathbb{R}^r)$ satisfying $\|x\|, \|y\| < \delta_5$ and $\|\omega\|_{L^2(0, t; \mathbb{R}^r)} < \delta_5$ it holds*

$$\|\nabla_{\xi\xi}^2 \mathcal{V}(t, y, \omega)^{-1} - \nabla_{\xi\xi}^2 \mathcal{V}(t, x, \omega)^{-1}\|_2 \leq L_{\mathcal{V}} \|y - x\|.$$

Proof. Note that $\mathcal{V}(t, \cdot, \omega)$, $\det(\cdot)$ and $\text{adj}(\cdot)$ are C^∞ -functions on their respective domains, where $\text{adj}(A)$ is the adjugate matrix of A . Hence Cramer's rule implies that $z \mapsto \nabla_{\xi\xi}^2 \mathcal{V}(t, z, \omega)^{-1}$ is C^∞ in $\mathcal{U}_{\delta_5}(0)$. Then for $x, y \in \mathbb{R}^n$ satisfying $\|x\|, \|y\| < \delta_5$ it holds

$$\begin{aligned} & \|\nabla_{\xi\xi}^2 \mathcal{V}(t, y, \omega)^{-1} - \nabla_{\xi\xi}^2 \mathcal{V}(t, x, \omega)^{-1}\|_2 = \left\| \int_0^1 D_\xi \nabla_{\xi\xi}^2 \mathcal{V}(t, x + \tau(y - x), \omega)^{-1} (y - x) \, d\tau \right\|_2 \\ & \leq \int_0^1 \|D_\xi \nabla_{\xi\xi}^2 \mathcal{V}(t, x + \tau(y - x), \omega)^{-1}\|_{L(\mathbb{R}^n, \mathbb{R}^{n \times n})} \, d\tau \|y - x\| \\ & = \int_0^1 \sup_{\|z\|=1} \|D_\xi \nabla_{\xi\xi}^2 \mathcal{V}(t, x + \tau(y - x), \omega)^{-1} z\|_2 \, d\tau \|y - x\| \\ & \leq \int_0^1 \sup_{\|z\|=1} \sum_{i=1}^n |z_i| \|D_\xi \nabla_{\xi\xi}^2 \mathcal{V}(t, x + \tau(y - x), \omega)^{-1} e_i\|_2 \, d\tau \|y - x\| \\ & \leq \int_0^1 \sum_{i=1}^n \left\| \frac{\partial}{\partial \xi_i} \nabla_{\xi\xi}^2 \mathcal{V}(t, x + \tau(y - x), \omega)^{-1} \right\|_2 \, d\tau \|y - x\| \\ & \leq \int_0^1 \sum_{i=1}^n \|\nabla_{\xi\xi}^2 \mathcal{V}(t, x + \tau(y - x), \omega)^{-1}\|_2^2 \left\| \frac{\partial}{\partial \xi_i} \nabla_{\xi\xi}^2 \mathcal{V}(t, x + \tau(y - x), \omega) \right\|_2 \, d\tau \|y - x\| \\ & \leq \int_0^1 \sum_{i=1}^n M_{\mathcal{V}}^2 \left\| \frac{\partial}{\partial \xi_i} \nabla_{\xi\xi}^2 \mathcal{V}(t, x + \tau(y - x), \omega) \right\|_2 \, d\tau \|y - x\|, \end{aligned}$$

where Lemma B.7 was used in the last estimate. It remains to show an appropriate estimate for $\left\| \frac{\partial}{\partial \xi_i} \nabla_{\xi\xi}^2 \mathcal{V}(t, y + \tau(x - y), \omega) \right\|_2$. To that end let $\nu \in \mathbb{R}^n$ fulfill $\|\nu\| < \delta_5$ and let $j, k \in \{1, \dots, n\}$ be such that

$$\max_{h, l \in \{1, \dots, n\}} \left| \frac{\partial^3}{\partial \xi_i \partial \xi_h \partial \xi_l} \mathcal{V}(t, \nu, \omega) \right| = \left| \frac{\partial^3}{\partial \xi_i \partial \xi_j \partial \xi_k} \mathcal{V}(t, \nu, \omega) \right|.$$

With Proposition 4.6 it follows

$$\begin{aligned} & \left\| \frac{\partial}{\partial \xi_i} \nabla_{\xi\xi}^2 \mathcal{V}(t, \nu, \omega) \right\|_2 \leq n \left\| \frac{\partial}{\partial \xi_i} \nabla_{\xi\xi}^2 \mathcal{V}(t, \nu, \omega) \right\|_{n, n} = n \left| \frac{\partial^3}{\partial \xi_i \partial \xi_j \partial \xi_k} \mathcal{V}(t, \nu, \omega) \right| \\ & = n |D_\xi^3 \mathcal{V}(t, \nu, \omega)(e_i, e_j, e_k)| \leq n \check{M}_3. \end{aligned}$$

Finally it follows

$$\|\nabla_{\xi\xi}^2 \mathcal{V}(t, y, \omega)^{-1} - \nabla_{\xi\xi}^2 \mathcal{V}(t, x, \omega)^{-1}\|_2 \leq M_{\mathcal{V}}^2 n^2 \check{M}_3 \|x - y\| =: L_{\mathcal{V}} \|x - y\|.$$

□

ACKNOWLEDGEMENTS

We thank K. Kunisch (KFU Graz) for many helpful comments and several discussions on earlier versions of this manuscript. We thank L. Pfeiffer (INRIA Paris-Saclay) for many helpful suggestions and pointers on the subject of sensitivity analysis. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 504768428.

REFERENCES

- [1] P. Benner and T. Breiten, Two-sided projection methods for nonlinear model order reduction. *SIAM J. Sci. Comput.* **37** (2015) B239–B260.
- [2] C. Gu, QLMOR: a projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **30** (2011) 1307–1320.
- [3] R.E. Mortensen, Maximum-likelihood recursive nonlinear filtering. *J. Optim. Theory Appl.* **2** (1968) 386–394.
- [4] R.E. Kalman, A new approach to linear filtering and prediction problems. *Trans. ASME Ser. D. J. Basic Eng.* **82** (1960) 35–45.
- [5] R.E. Kalman and R.S. Bucy, New results in linear filtering and prediction theory. *Trans. ASME Ser. D. J. Basic Eng.* **83** (1961) 95–108.
- [6] K. Ito and K. Xiong, Gaussian filter for nonlinear filtering problems. *IEEE Trans. Autom. Control* **45** (2000) 910–927.
- [7] H.J. Kushner, On the differential equations satisfied by conditional probability densities of Markov processes, with applications. *SIAM J. Control Optim.* **2** (1962) 106–119.
- [8] H.J. Kushner, Approximations to optimal nonlinear filters. *IEEE Trans. Autom. Control* **12** (1967) 546–556.
- [9] M. Zakai, On the optimal filtering of diffusion processes. *Z. Wahrschein. Verw. Gebiete* **11** (1969) 230–243.
- [10] J.C. Willems, Deterministic least squares filtering. *J. Econometrics* **118** (2004) 341–373.
- [11] D.G. Luenberger, An introduction to observers. *IEEE Trans. Autom. Control* **16** (1971) 596–602.
- [12] O. Hijab, *Minimum energy estimation*, Ph.D. dissertation, University of California, Berkeley (1980).
- [13] A.J. Krener, Minimum covariance, minimax and minimum energy estimators, in *Stochastic Control Theory and Stochastic Differential Systems*, edited by M. Kohlmann and W. Vogel. Vol. 16 of *Lecture Notes in Control and Information Science*. Springer, Berlin–New York (1979) 490–495.
- [14] W.H. Fleming, Deterministic nonlinear filtering. *Ann. Sc. Norm. Super. Pisa - Cl. Sci.* **25** (1997) 435–454.
- [15] P. Moireau, A discrete-time optimal filtering approach for non-linear systems as a stable discretization of the Mortensen observer. *ESAIM Control Optim. Calc. Var.* **24** (2018) 1815–1847.
- [16] W.H. Fleming and H.M. Soner, *Controlled Markov Processes and Viscosity Solutions*. Vol. 25 of *Stochastic Modelling and Applied Probability*. Springer, New York (2006).
- [17] A.J. Krener The convergence of the extended Kalman filter, in *Directions in Mathematical Systems Theory and Optimization*. Vol. 286 of *Lect. Notes Control Inf. Sci.* Springer, Berlin (2003) 173–182.
- [18] J.S. Baras, A. Bensoussan and M.R. James, Dynamic observers as asymptotic limits of recursive filters: special cases. *SIAM J. Appl. Math.* **48** (1988) 1147–1158.
- [19] A.J. Krener The convergence of the minimum energy estimator, in *New Trends in Nonlinear Dynamics and Control and their Applications*. Springer, Berlin (2003) 187–208.
- [20] T. Breiten and K. Kunisch, Neural network based nonlinear observers. *Syst. Control Lett.* **148** (2021).
- [21] M.G. Crandall and P.L. Lions, Viscosity solutions of Hamilton–Jacobi equations. *Trans. Am. Math. Soc.* **277** (1983) 1–42.
- [22] P. Cannarsa and H. Frankowska, Some characterizations of optimal trajectories in control theory. *SIAM J. Control Optim.* **29** (1991) 1322–1347.
- [23] P. Cannarsa and H. Frankowska, Value function and optimality conditions for semilinear control problems. *Appl. Math. Optim.* **26** (1992) 139–169.
- [24] P. Cannarsa and H. Frankowska, Local regularity of the value function in optimal control. *Syst. Control Lett.* **62** (2013) 791–794.
- [25] P. Cannarsa and H. Frankowska, From pointwise to local regularity for solutions of Hamilton–Jacobi equations. *Calc. Var. Partial Differ. Equ.* **49** (2014) 1061–1074.
- [26] T. Breiten, K. Kunisch and L. Pfeiffer, Feedback stabilization of the two-dimensional Navier–Stokes equations by value function approximation. *Appl. Math. Optim.* **80** (2019) 599–641.

- [27] T. Breiten, K. Kunisch and L. Pfeiffer, Taylor expansions of the value function associated with a bilinear optimal control problem. *Ann. Inst. Henri Poincaré (C) Anal. Non Linéaire* **36** (2019) 1361–1399.
- [28] K. Malanowski, Application of the classical implicit function theorem in sensitivity analysis of parametric optimal control. *Control Cybern.* **27** (1998) 335–352.
- [29] K. Malanowski and H. Maurer, Sensitivity analysis for state constrained optimal control problems. *Discrete Contin. Dyn. Syst.* **4** (1998) 241–272.
- [30] K. Malanowski and H. Maurer, Sensitivity analysis of optimal control problems subject to higher order state constraints. *Ann. Oper. Res.* **101** (2001) 43–73.
- [31] K. Ito and K. Kunisch, Lagrange Multiplier Approach to Variational Problems and Applications, 1st edn. SIAM, Philadelphia, PA (2008).
- [32] J.T. Marti, Evaluation of the least constant in Sobolev’s inequality for $H^1(0, s)$. *SIAM J. Numer. Anal.* **20** (1983) 1239–1242.
- [33] M. Hinze, R. Pinnau, M. Ulbrich and S. Ulbrich, Optimization with PDE Constraints. Vol. 23 of *Mathematical Modelling: Theory and Applications*. Springer (2009).
- [34] E. Zeidler Applied Functional Analysis: Main Principles and Their Applications, 1st edn. Springer, New York (1995).
- [35] J.M. Holtzman, Short notes: explicit ϵ and δ for the implicit function theorem. *SIAM Rev.* **12** (1970) 284–286.
- [36] O. Forster, Analysis 1, Differential- und Integralrechnung einer Veränderlichen, Springer Fachmedien, Wiesbaden, 11th edn (2013).
- [37] L. Dieci and T. Eirola, Positive definiteness in the numerical solution of Riccati differential equations. *Numer. Math.* **67** (1994) 303–313.
- [38] E. Zeidler, Applied Functional Analysis: Applications to Mathematical Physics, 1st edn. Springer, New York (1995).
- [39] W. Rudin, Principles of Mathematical Analysis, 3rd edn. McGraw-Hill Kogakusha, Tokyo (1976).



Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.