

Strict dissipativity analysis for classes of optimal control problems involving probability density functions

Lars Grüne
Chair of Applied Mathematics
University of Bayreuth
95440 Bayreuth, Germany
`lars.gruene@uni-bayreuth.de`

Arthur Fleig
Chair of Serious Games
University of Bayreuth
95440 Bayreuth, Germany
`arthur.fleig@uni-bayreuth.de`

July 9, 2019

Abstract

Motivated by the stability and performance analysis of model predictive control schemes, we investigate strict dissipativity for a class of optimal control problems involving probability density functions. The dynamics are governed by a Fokker-Planck partial differential equation. However, for the particular classes under investigation involving linear dynamics, linear feedback laws, and Gaussian probability density functions, we are able to significantly simplify these dynamics. This enables us to perform an in-depth analysis of strict dissipativity for different cost functions.

1 Introduction

Strict dissipativity of optimal control problems is a pivotal property for the rigorous stability and performance analysis of general (often also termed economic) Model Predictive Control (MPC) schemes. This fact was revealed in a series of recent papers, see, e.g., [6, 1, 12] or the monographs and survey papers [20, 14, 7], and has triggered a renewed interest in this classical systems theoretic property that goes back to [24].

In this paper we investigate strict dissipativity for a class of optimal control problems for probability density functions (PDFs). Such problems, in which the entire distribution of an Itô-stochastic control system is shaped via a suitable optimization objective, provide an interesting alternative to classic stochastic optimal approaches that optimize the mean or higher moments. The dynamics of the optimal control problem to be solved in this setting is determined by the Fokker-Planck partial differential equation (FP-PDE) [21], and the approach became particularly popular due to a series of papers in which an MPC-like scheme was applied to this problem [2, 3]. Since then, this approach was used in different contexts, e.g. in [4, 22].

For the simplest case, in which the optimization objective penalizes the distance of the state-control pair to a desired equilibrium PDF and the corresponding control input, the stability and performance of MPC for this problem was analyzed in [9, 10]. However, the setting analyzed in these references assumes that the control input corresponding to the desired equilibrium PDF is known and used in the optimization objective. Both may not be realistic: the computation of the control may be difficult — as it is determined by an inverse problem involving the FP-PDE — and it may not be desirable to penalize the distance of the control to this control but rather an economically more meaningful quantity, such as the overall control effort. When proceeding this way, one ends up at a more general MPC problem, a so-called unreachable setpoint problem [19]. This problem falls into the more general class of economic MPC schemes for which strict dissipativity is the main ingredient for guaranteeing stability and near optimal performance.

Motivated by these facts, we decided to analyze strict dissipativity for certain classes of this optimal control problem. This analysis was started in [8], where objectives involving the L^2 -norm for penalizing the distance to the desired equilibrium PDF was investigated - the results from this paper will be briefly

summarized below. Here we extend this analysis to alternative cost functions including the Wasserstein distance W^2 . While both the L^2 - and the W^2 -cost are perfectly suited for being used in a nonlinear setting for general PDFs — the resulting control scheme performs excellently in numerical tests — it turned out that for a rigorous mathematical analysis the problem must be simplified. Hence, as in [8], we perform our analysis for linear SDE dynamics governed by the Ornstein-Uhlenbeck process, linear feedback controllers, and Gaussian PDFs. While this is clearly a restricted setting, we believe that the insights from this analysis are nevertheless very valuable for the general nonlinear setting, in particular those results that clarify that certain storage functions are not appropriate for the strict dissipativity analysis. Clearly, if certain approaches do (provably) not work in the linear Gaussian setting, they will inevitably also fail in more general settings. Moreover, the linear Gaussian setting allows us to compare our results with general purpose cost functions to results for a cost function that is particularly tailored to the linear Gaussian setting. This cost function combines the 2-norm for the mean and the Frobenius norm for the covariance matrix of the Gaussian PDF and is thus termed 2F-cost. Despite its similarity with the W^2 -cost, the results on strict dissipativity are strikingly different for these two cost functions, which is another important result of this paper.

The paper is organized as follows. Section 2 introduces the problem and the cost function under consideration. In Section 3 we introduce strict dissipativity and briefly summarize the main results for MPC schemes that can be derived from this property, in order to motivate our subsequent analysis. Section 4 collects a few auxiliary results and summarizes the main results from [8] for the L^2 -cost before we present our new results for the W^2 -cost and the 2F-cost in Section 5. We end the paper with concluding remarks in Section 6.

2 Problem Setting

In this paper we study the optimal control of probability density functions (PDFs) associated with linear continuous-time stochastic processes such as

$$dX_t = AX_t dt + Bu(t)dt + DdW_t, \quad t \in (0, T_E) \quad (1)$$

with an (almost surely) initial condition $\mathring{X} \in \mathbb{R}^d$ and where $A \in \mathbb{R}^{d \times d}$, $B \in \mathbb{R}^{d \times l}$, $D \in \mathbb{R}^{d \times m}$ are given matrices, $W_t \in \mathbb{R}^m$ is an m -dimensional Wiener process and the control $u(t)$ is defined by

$$u(t) := -K(t)X_t + c(t) \quad (2)$$

with functions $K: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^{l \times d}$ and $c: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^l$. Since the control $u(t)$ exhibits this special structure, whenever beneficial, we identify with u the pair (K, c) . Plugging (2) into (1) leads to

$$dX_t = (A - BK(t))X_t dt + Bc(t)dt + DdW_t, \quad t \in (0, T_E). \quad (3)$$

The initial condition $\mathring{X} \in \mathbb{R}^d$ is assumed to be normally distributed, i.e., $\mathring{X} \sim \mathcal{N}(\mathring{\mu}, \mathring{\Sigma})$ with mean $\mathring{\mu} \in \mathbb{R}^d$ and covariance matrix $\mathring{\Sigma} \in \mathbb{R}^{d \times d}$, which is symmetric and positive definite.

In this linear setting, $X_t \in \mathbb{R}^d$ is normally distributed for all $t \geq 0$ and the corresponding PDF ρ reads

$$\rho(x, t) := |2\pi\Sigma(t)|^{-1/2} \exp\left(-\frac{1}{2}(x - \mu(t))^T \Sigma(t)^{-1}(x - \mu(t))\right), \quad (4)$$

where for matrices $A \in \mathbb{R}^{d \times d}$, throughout the paper, we write $|A| := \det(A)$. The evolution of the PDF (associated with the stochastic differential equation (SDE) (1) or (3)) is described by the Fokker-Planck equation, a parabolic linear partial differential equation:

$$\begin{aligned} \partial_t \rho - \sum_{i,j=1}^d \partial_{ij}^2 (\alpha_{ij} \rho) + \sum_{i=1}^d \partial_i (b_i(u) \rho) &= 0 \text{ in } Q, \\ \rho(\cdot, 0) &= \mathring{\rho} \text{ in } \Omega, \end{aligned} \quad (5)$$

where $Q := \Omega \times (0, T_E)$, $\Omega := \mathbb{R}^d$, $\alpha_{ij} := \sum_k D_{ik} D_{jk} / 2$, and $b(X_t, t; u) := (A - BK(t))X_t + Bc(t)$. This approach works for more general Markov processes and is not limited to normal distributions, cf. [21, 17, 18]. As mentioned in the introduction, in order to enable the analysis in this paper, we limit

ourselves to the case of Gaussian distributions with mean $\mu(t) \in \mathbb{R}^n$ and covariance matrix $\Sigma(t) \in \mathbb{R}^{n \times n}$. In this case, we can replace the Fokker-Planck equation by the following system of ODEs for μ and Σ

$$\begin{aligned}\dot{\mu}(t) &= (A - BK(t))\mu(t) + Bc(t), & \mu(0) &= \dot{\mu}, \\ \dot{\Sigma}(t) &= (A - BK(t))\Sigma(t) + \Sigma(t)(A - BK(t))^T + DD^T, & \Sigma(0) &= \dot{\Sigma}.\end{aligned}\tag{6}$$

Using this ODE system will enable us to analyze strict dissipativity for the optimal control problem we will introduce soon. Particularly, we will carry out the analysis in this paper for the Ornstein-Uhlenbeck process.

Example 1 (Ornstein-Uhlenbeck). *The controlled Ornstein-Uhlenbeck process is defined by*

$$dX_t = -(\theta + K(t))X_t dt + c(t)dt + \varsigma dW_t, \quad t \in (0, T_E)\tag{7}$$

with initial condition $\dot{X} \sim \mathcal{N}(\dot{\mu}, \dot{\Sigma})$, parameters $\theta, \varsigma > 0$ and control constraints $K(t) > -\theta$, i.e.,

$$0 < \theta + K(t) =: K_\theta(t).\tag{8}$$

Plugging $A - BK(t) = -K_\theta(t) \in \mathbb{R}_{>0}$ and $D = \varsigma \in \mathbb{R}_{>0}$ into (6) results in

$$\begin{aligned}\dot{\mu}(t) &= -K_\theta(t)\mu(t) + c(t), & \mu(0) &= \dot{\mu}, \\ \dot{\Sigma}(t) &= -2K_\theta(t)\Sigma(t) + \varsigma^2, & \Sigma(0) &= \dot{\Sigma}.\end{aligned}\tag{9}$$

Our aim is to steer the probability density function (PDF) ρ to a desired Gaussian PDF

$$\bar{\rho}(x) := |2\pi\bar{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2}(x - \bar{\mu})^T \bar{\Sigma}^{-1}(x - \bar{\mu})\right),\tag{10}$$

starting from an initial (Gaussian) PDF $\dot{\rho}$. In continuous time, this can be formulated as the following optimal control problem:

$$J_\infty^c(\dot{\rho}, u) := \int_0^\infty \ell(\rho(x, t), u(t)) dt \rightarrow \min_{(u, \rho)}! \quad \text{subject to (6),}\tag{11}$$

where

$$\ell(\rho, u) = \frac{1}{2} \|\rho - \bar{\rho}\|^2 + \frac{\gamma}{2} \|u\|^2\tag{12}$$

for some norm $\|\cdot\|$ and some weight $\gamma \geq 0$ and where we use (4) to calculate the PDF from the solution of (6).

The choice of ℓ , i.e., the choice of the norms in (12) is important. For the control cost, we identify with u the pair (K, c) and suggest the Frobenius norm for K and the Euclidean norm for c , which fit well together. For the state cost, we have three options. The first possibility is to use the L^2 norm

$$\ell_{L^2}(\rho, u) := \frac{1}{2} \|\rho - \bar{\rho}\|_{L^2(\mathbb{R}^d)}^2 + \frac{\gamma}{2} \|K\|_F^2 + \frac{\gamma}{2} \|c\|_2^2,\tag{13}$$

which is the standard norm used in costs for optimal control problems governed by parabolic PDEs [23]. This choice of the cost was analyzed in [8]. We can express the L^2 norm in terms of Σ and μ , which proves useful when focusing on the ODE system (6) instead:

$$\begin{aligned}\|\rho(\cdot, t) - \bar{\rho}(\cdot)\|_{L^2(\mathbb{R}^d)}^2 &= 2^{-d} \pi^{-\frac{d}{2}} \left[|\Sigma(t)|^{-\frac{1}{2}} + |\bar{\Sigma}|^{-\frac{1}{2}} \right. \\ &\quad \left. - 2 \left| \frac{1}{2}(\Sigma(t) + \bar{\Sigma}) \right|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mu(t) - \bar{\mu})^T (\Sigma(t) + \bar{\Sigma})^{-1} (\mu(t) - \bar{\mu})\right) \right].\end{aligned}\tag{14}$$

Looking at the cost from the ODE perspective, the L^2 penalization (13) does not seem standard or intuitive at all. One alternative is to use the Wasserstein metric, which is specifically designed to measure the distance between two PDFs. For the general definition of this metric we refer to [11]. Here we only

use the formula for the Wasserstein metric for normal distributions derived in [11]. In case Σ and $\bar{\Sigma}$ commute¹ this formula yields the following stage cost:

$$\ell_{W^2}^\mu(\mu, \Sigma, K, c) := \frac{1}{2} \|\mu - \bar{\mu}\|_2^2 + \frac{1}{2} \|\Sigma^{1/2} - \bar{\Sigma}^{1/2}\|_F^2 + \frac{\gamma}{2} \|K\|_F^2 + \frac{\gamma}{2} \|c\|_2^2. \quad (15)$$

The third option we discuss in this paper is very similar to the Wasserstein distance from (15). The only difference is to consider Σ and $\bar{\Sigma}$ instead of $\Sigma^{1/2}$ and $\bar{\Sigma}^{1/2}$, respectively. Thus, we end up with

$$\ell_{2F}^\mu(\mu, \Sigma, K, c) := \frac{1}{2} \|\mu - \bar{\mu}\|_2^2 + \frac{1}{2} \|\Sigma - \bar{\Sigma}\|_F^2 + \frac{\gamma}{2} \|K\|_F^2 + \frac{\gamma}{2} \|c\|_2^2. \quad (16)$$

This form of the cost function is commonly used in optimization of systems governed by ODE systems. The index used in the notation for this cost, $2F$, indicates the combination of Euclidean and Frobenius norm. In the special case $\bar{\Sigma} = I$ we have that $\ell_{W^2}^\mu(\mu, \Sigma^2, K, c) = \ell_{2F}^\mu(\mu, \Sigma, K, c)$, i.e., considering the squared covariance matrix Σ^2 instead of Σ in the W^2 cost leads to the $2F$ cost.

To summarize, we consider infinite horizon optimal control problems for probability density functions governed by different types of cost functions. For these problems we investigate strict dissipativity in this paper. The motivation for this analysis is given by one of the most popular computational approaches to such optimal control problems, namely model predictive control (MPC). Before we turn to the dissipativity analysis in Sections 4 and 5, we explain this motivation and define strict dissipativity in the next section.

3 Model Predictive Control

In this section we briefly introduce the concept of (nonlinear) MPC. More details can be found in the monographs [14] and [20].

Suppose we have a process whose state $z(k)$ is measured at discrete times t_k , $k \in \mathbb{N}_0$. Furthermore, suppose we can control it on the time interval $[t_k, t_{k+1})$ via a control signal $u(k)$. Then we can consider nonlinear discrete time control systems

$$z^+ := z(k+1) = f(z(k), u(k)), \quad z(0) = z_0, \quad (17)$$

with state $z(k) \in \mathbb{X} \subset Z$ and control $u(k) \in \mathbb{U} \subset U$, where Z and U are metric spaces and state and control constraint sets are given by \mathbb{X} and \mathbb{U} , respectively. Continuous time models such as the one presented in Section 2 can be considered in the discrete-time setting by sampling, using a (constant) sampling time $T > 0$, i.e., $t_k = t_0 + kT$. Given an initial state z_0 and a control sequence $(u(k))_{k \in \mathbb{N}_0}$, the solution trajectory is denoted by $z_u(\cdot; z_0)$. Note that we do not require the control $u(k)$ to be constant on $[t_k, t_{k+1})$ – in general, each $u(k)$ can be a time-dependent function on $[t_k, t_{k+1})$.

As we have seen in Section 2, stabilization and tracking problems can be recast as infinite horizon optimal control problems (11). However, solving these is in general computationally hard. The idea behind MPC is to circumvent this issue by iteratively solving optimal control problems on a shorter, finite time horizon and use the resulting optimal control values to construct a feedback law $\mathcal{F}: \mathbb{X} \rightarrow \mathbb{U}$ for the closed loop system

$$z_{\mathcal{F}}(k+1) = f(z_{\mathcal{F}}(k), \mathcal{F}(z_{\mathcal{F}}(k))). \quad (18)$$

In the discrete-time setting, the infinite horizon functional (11) translates to

$$J_\infty(z_0, u) := \sum_{k=0}^{\infty} \ell(z_u(k; z_0), u(k)). \quad (19)$$

Since infinite horizon problems are in general very difficult to solve, we construct the feedback law \mathcal{F} through the following MPC scheme:

0. Given an initial value $z_{\mathcal{F}}(0) \in \mathbb{X}$, fix the length of the receding horizon $N \geq 2$ and set $n = 0$.

¹This does not limit our analysis since w.l.o.g. we can restrict ourselves to $\bar{\Sigma} = I$, see Section 4.

1. Initialize the state $z_0 = z_{\mathcal{F}}(n)$ and solve the following optimal control problem:

$$\begin{aligned} J_N(z_0, u) &:= \sum_{k=0}^{N-1} \ell(z_u(k; z_0), u(k)) \rightarrow \min_{u \in \mathbb{U}^N} ! \\ \text{s.t. } z_u(k+1; z_0) &= f(z_u(k; z_0), u(k)), \quad z_u(0; z_0) = z_0. \end{aligned} \quad (\text{OCP}_N)$$

Apply the first value of the resulting optimal control sequence denoted by $u^* \in \mathbb{U}^N$, i.e., set $\mathcal{F}(z_{\mathcal{F}}(n)) := u^*(0)$.

2. Evaluate $z_{\mathcal{F}}(n+1)$ according to relation (18), set $n := n+1$ and go to step 1.

Whenever we want to point out the importance of N , we will denote the feedback by \mathcal{F}_N instead of \mathcal{F} .

When passing from the infinite horizon formulation to the MPC scheme, a priori it is not clear, at all, whether we will obtain approximately optimal trajectories. In fact, it is not even clear whether the closed loop system is asymptotically stable.

A key difference for the analysis of MPC schemes is whether ℓ is positive definite with respect to some given equilibrium pair (z^e, u^e) of (17), i.e., $\ell(z^e, u^e) = 0$ and $\ell(z, u) > 0$ for $(z, u) \neq (z^e, u^e)$, where $f(z^e, u^e) = z^e$. A prime example is the stage cost

$$\ell(z(k), u(k)) = \frac{1}{2} \|z(k) - z^e\|^2 + \frac{\gamma}{2} \|u(k) - u^e\|^2, \quad (20)$$

for some norm $\|\cdot\|$ and some $\gamma > 0$. For this case, called *stabilizing MPC*, in [10] we have answered the question of minimal stabilizing horizon lengths for a class of linear stochastic processes.

However, the above cost function may be difficult to compute because one needs to know the corresponding u^e for a desired z^e beforehand, which may be cumbersome to compute. A stage cost that is less complicated to design and thus easier to implement is

$$\ell(z(k), u(k)) = \frac{1}{2} \|z(k) - z^e\|^2 + \frac{\gamma}{2} \|u(k)\|^2. \quad (21)$$

This function is also more common in optimal control literature and structurally similar to the costs (13), (15), and (16). For $u^e \neq 0$, the new stage cost ℓ is not positive definite w.r.t. (z^e, u^e) since $\ell(z^e, u^e) \neq 0$.² The specific stage cost (21) models a so-called *unreachable setpoint problem* [19], which is a particular instance of an *economic MPC* problem.

The conceptual difference between stabilizing and economic MPC is that, instead of stabilizing a prescribed equilibrium pair (z^e, u^e) via a stage cost that is positive definite w.r.t. that pair, in economic MPC the interplay of these stage cost and dynamics determines the optimal (long-term) behavior. As such, equilibria stay equally important, but the definition of the decisive optimal equilibrium changes.

Definition 2 (Optimal Equilibrium). *An equilibrium pair $(z^e, u^e) \in \mathbb{X} \times \mathbb{U}$ is called optimal $:\Leftrightarrow \forall (z, u) \in \mathbb{X} \times \mathbb{U}$ with $f(z, u) = z : \ell(\bar{z}, \bar{u}) \leq \ell(z, u)$.*

Assuming such an equilibrium pair (z^e, u^e) exists (which is the case for the dynamics considered in Section 2) and if f and ℓ are continuous and $\mathbb{X} \times \mathbb{U}$ is compact, then an optimal equilibrium exists, see, e.g., [14, Lemma 8.4]. The next natural question is under which circumstances – if at all – it is asymptotically stable for the MPC closed loop. In [1, 15] it was shown that *strict dissipativity* is the decisive property. In order to define it, we use the notation

$$|z_1|_{z_2} := d_Z(z_1, z_2) \quad (22)$$

for the distance from z_1 to z_2 and recall the notion of comparison functions, introduced by Hahn in [16].

Definition 3 (Comparison functions). *Let $\alpha: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ be a continuous function. Then*

- $\alpha \in \mathcal{K} \quad :\Leftrightarrow \alpha$ is strictly increasing and $\alpha(0) = 0$,
- $\alpha \in \mathcal{K}_{\infty} \quad :\Leftrightarrow \alpha \in \mathcal{K}$ and α is unbounded,
- $\alpha \in \mathcal{L} \quad :\Leftrightarrow \alpha$ is strictly decreasing and $\lim_{t \rightarrow \infty} \alpha(t) = 0$.

A continuous function $\beta: \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is called a \mathcal{KL} function $:\Leftrightarrow \forall t \geq 0 : \beta(\cdot, t) \in \mathcal{K}$ and $\forall r > 0 : \beta(r, \cdot) \in \mathcal{L}$.

²Redefining $\ell_2(z, u) := \ell(z, u) - \ell(z^e, u^e)$ usually does not help as it may lead to $\ell_2(z, u) < 0$ for some (z, u) .

Definition 4 ((Strict) Dissipativity, Storage Function, Modified Cost). (a) The optimal control problem (OCP_N) with stage cost ℓ is called strictly dissipative at an equilibrium pair $(z^e, u^e) \in \mathbb{X} \times \mathbb{U}$ if there exist a function $\lambda: \mathbb{X} \rightarrow \mathbb{R}$ that is bounded from below and satisfies $\lambda(z^e) = 0$ and a function $\varrho \in \mathcal{K}_\infty$ such that for all $(z, u) \in \mathbb{X} \times \mathbb{U}$:

$$\ell(z, u) - \ell(z^e, u^e) + \lambda(z) - \lambda(f(z, u)) \geq \varrho(|z|_{z^e}). \quad (23)$$

(b) If (a) holds with $\varrho \equiv 0$ then the optimal control problem is called dissipative.

(c) The function λ in (a) is called storage function.

(d) The left-hand-side of (23), i.e.,

$$\tilde{\ell}(z, u) := \ell(z, u) - \ell(z^e, u^e) + \lambda(z) - \lambda(f(z, u)), \quad (24)$$

is called modified cost or rotated cost.

Note that the requirement $\lambda(z^e) = 0$ in Definition 4(a) can always be satisfied by a constant translation of λ without influencing the inequality (23).

In a classical interpretation of (23), $\lambda(z)$ serves as a quantifier for the amount of energy stored at state z , $\ell(z, u) - \ell(z^e, u^e)$ can be viewed as a supply rate that tracks the amount of energy supplied to or withdrawn from the system via the control u , and $\varrho(|z|_{z^e})$ is the amount of energy the system releases (or dissipates) to the environment in each step. Note, however, that in the optimal control problems we discuss here there is not necessarily a notion of “energy” in a physical sense.

For strictly dissipative optimal control problems satisfying appropriate continuity properties³, the following statements hold.

- The optimal equilibrium x^e is practically asymptotically stable for the MPC closed loop, where the neighborhood around x^e to which the closed-loop trajectory converges shrinks down to x^e as the horizon $N \rightarrow \infty$.
- The MPC closed-loop trajectories are approximately averaged optimal with approximation error tending to 0 as $N \rightarrow \infty$.
- On any finite horizon K , the MPC closed-loop trajectories are approximately optimal among all other trajectories converging to x^e , with an approximation error that grows linearly in K and tends to 0 as $N \rightarrow \infty$.
- For suitable terminal constraints and costs these properties can be improved to exact (as opposed to practical) asymptotic stability, exact averaged optimality and to finite-horizon optimality with an approximation error that is independent of K .

In summary, strict dissipativity is the decisive structural property that makes MPC work. This is the main motivation why we analyze it in this paper. Before turning to this analysis, we briefly discuss its relation to another important property of optimal control problems, the so-called *turnpike property*. This property demands that there exists $\sigma \in \mathcal{L}$ such that for all $N, P \in \mathbb{N}$, $z \in \mathbb{X}$ and the optimal trajectories $z^*(k, z)$ with horizon N the set

$$\mathcal{Q}(z, u, P, N) := \{k \in \{0, \dots, N-1\} \mid |z^*(k, z)|_{z^e} \geq \sigma_\delta(P)\}$$

has at most P elements. In words, most of the time the finite-horizon optimal trajectories stay close to the optimal equilibrium x^e .

Under a boundedness condition on the optimal value function (known as cheap reachability), it can be shown that strict dissipativity implies the turnpike property and under a controllability condition, these two properties are even equivalent [13]. Hence, the turnpike property is often a good indicator for strict dissipativity. In contrast to strict dissipativity, the turnpike property is more difficult to check analytically, because it involves the knowledge of optimal trajectories. On the other hand, the turnpike property is more easily checked numerically by means of simulating optimal trajectories. Hence, these two properties complement each other in a nice way when analyzing strict dissipativity of optimal control problems.

³For a precise definition as well as for the precise mathematical formulations of the subsequent statements we refer to [14, Chapter 8].

4 Auxiliary Results Regarding Dissipativity

After the introduction of MPC, now we turn our attention to dissipativity. More precisely, we analyze whether the optimal control problems under consideration are (strictly) dissipative in the sense of Definition 4. For this, we first rephrase our objective to steer a PDF to a target PDF in discrete time. To this end, we have to specify the dynamics at hand.

As mentioned in Section 2, we will carry out our analysis for the Ornstein-Uhlenbeck Process from Example 1. The reason for this is its simple, but bilinear structure. For the sake of better comparability to [6, 5], in which dissipativity of linear discrete time dynamics was considered, we would like to keep the bilinear structure in the discrete time setting. Moreover, in any numerical implementation of MPC the dynamics must be approximated by a numerical scheme. In order to allow for a fast computation of the optimal open-loop trajectories, in MPC implementations simple but less accurate schemes are often preferred to more expensive high-order methods. For these reasons, as in [8], we perform our analysis for the forward Euler approximation of the ODE system (9). This discretization both maintains the bilinear structure and defines a scheme that is frequently used in practice. It is given by

$$\mu^+ = \mu(k) + T(-K_\theta(k)\mu(k) + c(k)), \quad \mu(0) = \mathring{\mu}, \quad (25a)$$

$$\Sigma^+ = \Sigma(k) + T(-2K_\theta(k)\Sigma(k) + \varsigma^2), \quad \Sigma(0) = \mathring{\Sigma}. \quad (25b)$$

In contrast to [8], in which strict dissipativity for the stage cost $\ell_{L^2}(\rho, u)$ from (13) was analyzed, the stage cost we consider in this paper is given either by $\ell_{W^2}^\mu(\mu, \Sigma, K, c)$ from (15) or by $\ell_{2F}^\mu(\mu, \Sigma, K, c)$ from (16).

Remark 5. *Note that the state constraint $\Sigma > 0$ automatically holds for (6) and (9). However, when switching to the Euler approximation (25), we have to explicitly require $\Sigma(k) > 0$ for all $k \in \mathbb{N}_0$. In conjunction with $K_\theta(k) > 0$, cf. (8), this can be incorporated as control constraints*

$$0 < K_\theta(k) < (\Sigma(k) + T\varsigma^2)/(2T\Sigma(k)). \quad (26)$$

The optimal control problem that is solved in the MPC algorithm, cf. (OCP_N), then is

$$J_N^\mu((\mathring{\mu}, \mathring{\Sigma}), (K, c)) := \sum_{k=0}^{N-1} \ell((\mu(k), \Sigma(k)), (K(k), c(k))) \rightarrow \min! \quad (27)$$

subject to (25), (26)

with ℓ given by either (15) or (16). To prove that (27) is strictly dissipative, we need to find a suitable storage function λ for which the inequality (23) in Definition 4 holds. In general, it is not easy to find such a function. However, for OCPs with linear discrete-time dynamics

$$z(k+1) = Az(k) + Bu(k) + c =: f^l(z(k), u(k)), \quad (28)$$

a convex constraint set and strictly convex stage cost ℓ , it is known [6] that the linear function

$$\lambda^l(z) := \bar{\lambda}^T z \quad (29)$$

is a suitable storage function; for a proof, see, e.g., [5].⁴ Here, $\bar{\lambda}$ is the Lagrange multiplier in the optimization problem consisting of finding the optimal equilibrium (z^e, u^e) :

$$\min_{(z,u)} \ell(z, u) \quad \text{s.t. } z = f^l(z, u). \quad (30)$$

The reason for this is the close connection between the Lagrange function $L(z, u, \lambda)$ associated to (30) and the resulting modified cost $\tilde{\ell}$:

$$\begin{aligned} \tilde{\ell}(z, u) &= \ell(z, u) - \ell(z^e, u^e) + \lambda^l(z) - \lambda^l(f^l(z, u)) \\ &= \ell(z, u) - \ell(z^e, u^e) + \bar{\lambda}^T (z - f^l(z, u)) \\ &= L(z, u, \bar{\lambda}) - \ell(z^e, u^e). \end{aligned} \quad (31)$$

⁴One can ensure the boundedness from below that is required in Definition 4 by state constraints.

In this particular form of dissipativity, also known as strict duality in optimization theory, the (strict) convexity of ℓ carries over to L and therefore to $\bar{\ell}$, with the global minimum being attained at (z^e, u^e) . In the final step, due to $L(z^e, u^e, \bar{\lambda}) = \ell(z^e, u^e)$, we have that $\bar{\ell}$ is positive definite with respect to the optimal equilibrium (z^e, u^e) , which allows to conclude (23), i.e., strict dissipativity.

Although this is, in general, not true for nonlinear $f(z, u)$, in the following, we analyze how far the approach of a linear storage function can be successfully extended to bilinear OCPs, such as (27) with stage cost ℓ given by (15) or (16). To this end, in the rest of this section, we state some auxiliary results. These were presented in [8, Lemmas 4,5,6] for the L^2 cost (13). Since they trivially extend to the stage costs (15) and (16), we omit the proofs.

The first result characterizes equilibria. We recall that the imposed constraints ensure $\bar{K}_\theta = \theta + \bar{K} > 0$, cf. (26).

Lemma 6. *Let $\bar{K}_\theta := \theta + \bar{K}$. The set of equilibria is identical for (9) and (25) and is given by*

$$\mathcal{E} := \left\{ (\bar{\mu}, \bar{\Sigma}, \bar{K}, \bar{c}) \mid \bar{\mu} = \frac{\bar{c}}{\bar{K}_\theta}, \bar{\Sigma} = \frac{\varsigma^2}{2\bar{K}_\theta} \right\}. \quad (32)$$

Without loss of generality, we assume that $(\bar{\mu}, \bar{\Sigma}) = (0, 1)$. Otherwise we can introduce a new random variable $Y_t := \bar{\Sigma}^{-1/2}(X_t - \bar{\mu})$ and get a similar ODE system to (9). With this assumption, due to (32), we have $\bar{c} = 0$, which allows us to further simplify the dynamics under consideration for the chosen cost criteria.

Lemma 7. *Assume that $(\bar{\mu}, \bar{\Sigma}) = (0, 1)$. Then the OCP (27) with ℓ given by either (15) or (16) is strictly dissipative at an equilibrium $(0, \bar{\Sigma}, \bar{K}, 0)$ if and only if the OCP*

$$J_N(\bar{\Sigma}, K) := \sum_{k=0}^{N-1} \ell((0, \Sigma(k)), (K(k), 0)) \rightarrow \min! \quad (33)$$

subject to (25b), (26),

with the same ℓ , is strictly dissipative at the equilibrium $(\bar{\Sigma}, \bar{K})$.

Thus, in the following, we will focus on whether (33) is strictly dissipative. Since we assume $(\bar{\mu}, \bar{\Sigma}) = (0, 1)$, the two different stage cost functions under consideration – (15) and (16) – can be simplified to

$$\ell_{W^2}(\Sigma, K) := \ell_{W^2}^\mu(0, \Sigma, K, 0) = \frac{1}{2} \left(\sqrt{\Sigma} - 1 \right)^2 + \frac{\gamma}{2} K^2, \quad (34)$$

$$\ell_{2F}(\Sigma, K) := \ell_{2F}^\mu(0, \Sigma, K, 0) = \frac{1}{2} (\Sigma - 1)^2 + \frac{\gamma}{2} K^2, \quad (35)$$

respectively. For these cost functions we make preliminary statements about the optimal equilibria, which will be important later on.

Lemma 8. *Let (Σ^e, K^e) be an optimal equilibrium for one of the stage cost functions $\ell_{W^2}(\Sigma, K)$ or $\ell_{2F}(\Sigma, K)$. Then*

$$\begin{cases} K^e \in [0, \frac{\varsigma^2}{2} - \theta] \text{ and } \Sigma^e \in [1, \frac{\varsigma^2}{2\theta}], & \text{if } \varsigma^2/2 - \theta > 0, \\ K^e \in [\frac{\varsigma^2}{2} - \theta, 0] \text{ and } \Sigma^e \in [\frac{\varsigma^2}{2\theta}, 1], & \text{if } \varsigma^2/2 - \theta < 0, \\ K^e = 0 \text{ and } \Sigma^e = 1, & \text{if } \varsigma^2/2 - \theta = 0. \end{cases} \quad (36)$$

Before we turn to present our new results for the W^2 stage cost (15) and the 2F stage cost (16), we briefly recall the results from [8] for the L^2 stage cost (13). The different cases in Lemma 8 were decisive for this analysis:

- For $\varsigma^2/2 - \theta > 0$, strict dissipativity cannot hold with a linear storage function.
- In the case $\varsigma^2/2 - \theta < 0$, for certain values of θ and ς strict dissipativity holds with a linear storage function.
- For both $\varsigma^2/2 - \theta > 0$ and $\varsigma^2/2 - \theta < 0$, we have constructed a nonlinear storage function for which strict dissipativity holds for certain values of θ and ς .
- The case $\varsigma^2/2 - \theta = 0$ is of minor interest since it corresponds to the stabilizing MPC case.

As we will see, these cases will also play a role in the analysis in the next section.

5 Results on strict dissipativity

In the previous section we simplified the OCP under consideration, (27), by finding an equivalent formulation (33), which is sufficient for analyzing dissipativity. This section is dedicated to the dissipativity analysis of the OCP (33) for the W^2 cost (34) and the 2F cost (35). We begin with the latter.

5.1 2F cost

In this section we consider (33) with the 2F stage cost (35). In the one-dimensional case, this amounts to penalizing the quadratic deviation of the variance in addition to the control effort. Overall, the optimization problem in this section is given by

$$\begin{aligned} J_N(\bar{\Sigma}, K) &:= \sum_{k=0}^{N-1} \ell_{2F}(\Sigma(k), K(k)) = \sum_{k=0}^{N-1} \left[\frac{1}{2}(\Sigma(k) - 1)^2 + \frac{\gamma}{2}K(k)^2 \right] \rightarrow \min! \\ \text{subject to } \Sigma^+ &= \Sigma(k) + T(-2K_\theta(k)\Sigma(k) + \varsigma^2) =: f(\Sigma(k), K(k)), \quad \Sigma(0) = \bar{\Sigma}, \\ 0 &< K_\theta(k) < (\Sigma(k) + T\varsigma^2)/(2T\Sigma(k)). \end{aligned} \quad (37)$$

With the linear storage function $\lambda^l(z)$, the corresponding modified cost $\tilde{\ell}_{2F}(\Sigma, K)$, cf. (24), is given by

$$\tilde{\ell}_{2F}(\Sigma, K) := \frac{1}{2}(\Sigma - 1)^2 + \frac{\gamma}{2}K^2 - \ell_{2F}(\Sigma^e, K^e) - \bar{\lambda}T(-2(\theta + K)\Sigma + \varsigma^2). \quad (38)$$

Throughout this section, the pair (Σ^e, K^e) denotes an optimal equilibrium, i.e., a solution of

$$\min_{(\Sigma, K)} \frac{1}{2}(\Sigma - 1)^2 + \frac{\gamma}{2}K^2 \quad \text{s.t. } \Sigma - f(\Sigma, K) = 0. \quad (39)$$

The unique⁵ Lagrange multiplier $\bar{\lambda} \in \mathbb{R}$ is obtained from the associated Lagrange function

$$L_{2F}(\Sigma, K, \lambda) := \frac{1}{2}(\Sigma - 1)^2 + \frac{\gamma}{2}K^2 + \lambda[-T(-2(\theta + K)\Sigma + \varsigma^2)]. \quad (40)$$

Note that we have not included state or control constraints in the Lagrange function. This is to keep the close connection to the modified cost $\tilde{\ell}$, cf. (31). From Lemma 8 we know that these constraints are always satisfied for optimal equilibria. However, away from the equilibria, we have to enforce them.

Remark 9. One could argue that, due to the forward Euler approximation (25), the dynamics are effectively scaled by the sampling time T , and this scaling should also be applied to the stage cost ℓ . In this case, T can be factored out of the Lagrange function, arriving at

$$L_{2F}^c(\Sigma, K, \lambda_c) := \frac{1}{2}(\Sigma - 1)^2 + \frac{\gamma}{2}K^2 - \lambda_c(-2(\theta + K)\Sigma + \varsigma^2), \quad (41)$$

which is the Lagrange function for the (unscaled) stage cost ℓ with continuous dynamics, cf. (9). The Lagrange multiplier $\bar{\lambda}_c$ is unique and independent of the sampling time T . The connection to (40) is easily established via

$$\lambda_c = \bar{\lambda}T.$$

Thus, while the Lagrange multiplier $\bar{\lambda}$ from (40) changes with T , the product $\bar{\lambda}T$ and the optimal solutions are independent of T . Since in the following only the product $\bar{\lambda}T$ is of relevance, we avoid scaling the stage cost ℓ .

With

$$Z := 2\bar{\lambda}T, \quad (42)$$

the gradient of $\tilde{\ell}_{2F}$ is given by

$$\nabla \tilde{\ell}_{2F}(\Sigma, K) = \begin{pmatrix} \Sigma - 1 \\ \gamma K \end{pmatrix} + Z \begin{pmatrix} \theta + K \\ \Sigma \end{pmatrix}, \quad (43)$$

⁵due to $\nabla(\Sigma - f(\Sigma, K)) \neq 0$

and it holds

$$\nabla \tilde{\ell}_{2F}(\Sigma, K) = \nabla_{\Sigma, K} L_{2F}(\Sigma, K, \bar{\lambda}). \quad (44)$$

The Hessian of $\tilde{\ell}_{2F}$ is given by

$$\nabla^2 \tilde{\ell}_{2F}(\Sigma, K) = \begin{pmatrix} 1 & Z \\ Z & \gamma \end{pmatrix}. \quad (45)$$

Due to the bilinear terms in f , the off-diagonal entries in the Hessian are non-zero in general. This again demonstrates that convexity of the stage cost ℓ_{2F} does not necessarily carry over to the modified cost $\tilde{\ell}_{2F}$, which is in contrast to the case with linear constraints. This is mitigated by the fact that, in general, (strict) convexity of $\tilde{\ell}_{2F}$ is only sufficient for (strict) dissipativity. A necessary condition, however, is that (Σ^e, K^e) is a (strict) global minimum of $\tilde{\ell}_{2F}$. To this end, as a first step, we characterize the stationary points of $\tilde{\ell}_{2F}$ for a fixed $\bar{\lambda}$, i.e., a fixed Z .

Lemma 10. *For a fixed $\bar{\lambda} \in \mathbb{R}$ and thus Z , the stationary points of $\tilde{\ell}_{2F}$ are given by either*

$$\Sigma = -\frac{\gamma(Z\theta - 1)}{\gamma - Z^2}, \quad K = \frac{Z(Z\theta - 1)}{\gamma - Z^2} \quad (46)$$

if $\gamma - Z^2 \neq 0$ or by

$$\Sigma = -\frac{K}{\theta} \quad (47)$$

for arbitrary K in case $\gamma - Z^2 = 0$.

Proof. Solving $\partial_{\Sigma} \tilde{\ell}_{2F}(\Sigma, K) = 0$ for Σ yields

$$\Sigma = 1 - Z(\theta + K), \quad (48)$$

cf. (43). Plugging this into $\partial_K \tilde{\ell}_{2F}(\Sigma, K) = 0$ results in

$$0 = \gamma K + Z(1 - Z(\theta + K)) = (\gamma - Z^2)K + Z(1 - Z\theta). \quad (49)$$

Assuming that $\gamma - Z^2 \neq 0$, one can solve for K , which results in the equation for K in (46). Plugging this K into (48) gives the equation for Σ in (46).

If $\gamma - Z^2 = 0$, then $Z \neq 0$ since $\gamma > 0$. Since (Σ^e, K^e) is always a stationary point due to (44), we infer from (49) that $1 - Z\theta = 0$, i.e., $Z = 1/\theta$. In this case, from (48) we get (47) for arbitrary K . \square

Remark 11. (a) *The set of possible stationary points in Lemma 10 is restricted by the constraints $K > -\theta$ and $\Sigma > 0$. More importantly, however, in case of $\gamma - Z^2 \neq 0$, the stationary point is unique and coincides with (Σ^e, K^e) , whereas there can be infinitely many stationary points if $\gamma - Z^2 = 0$.*

(b) *From the proof of Lemma 10 we see that $\gamma - Z^2 = 0$ can only occur if $\gamma = 1/\theta^2$.*

The sign of $\gamma - Z^2$ is indeed crucial for the rest of this section: Since the Hessian $\nabla^2 \tilde{\ell}_{2F}(\Sigma, K)$ is constant, the requirement that the optimal equilibrium is a (strict) global minimum of the modified cost $\tilde{\ell}_{2F}$ is equivalent to $\gamma - Z^2 > 0$. Hence, strict dissipativity with the linear storage function $\lambda^l(z)$ is equivalent to strong convexity of the modified cost $\tilde{\ell}_{2F}$. The decisive factor is the sign of $\gamma - Z^2$. Thus, in the following, we focus on finding sets of parameters for which a certain sign of $\gamma - Z^2$ can be guaranteed.

Proposition 12. *If $\frac{\gamma}{2} - \theta > 0$, then (37) is strictly dissipative with the linear storage function $\lambda^l(z)$ from (29).*

Proof. The assertion follows from the fact that for $\frac{\gamma}{2} - \theta > 0$ the Hessian $\nabla^2 \tilde{\ell}_{2F}(\Sigma, K)$ is positive definite. Indeed, in this case the function $\tilde{\ell}_{2F}$ in (24) is strongly convex, which immediately implies the existence of a quadratic lower bound $\varrho \in \mathcal{K}_{\infty}$ in the dissipativity inequality (23).

It is thus sufficient to prove the Hessian is positive definite, which holds if and only if $\gamma - Z^2 > 0$. To prove this, we need some information about the Lagrange multiplier $\bar{\lambda}$, which we get by taking a closer look at the Lagrange function (40). Since (Σ^e, K^e) is an optimal equilibrium, $\nabla L_{2F}(\Sigma, K, \bar{\lambda}) = 0$. In particular, we can use the results of Lemma 10 due to (44).

First, we show that $\gamma - Z^2 \neq 0$: If we assume the opposite, then from Lemma 10 we see that the optimal equilibrium (Σ^e, K^e) satisfies (47), i.e., $\Sigma^e = -K^e/\theta$ for some K^e . However, from Lemma 8 we know that $K^e \in [0, \frac{\zeta^2}{2} - \theta]$ and $\Sigma^e \in [1, \frac{\zeta^2}{2\theta}]$. In particular, $\Sigma^e \geq 1$ and $K^e \geq 0$, which contradicts (47).

Knowing that $\gamma - Z^2 \neq 0$, we now show that $\gamma - Z^2 > 0$: Since $K^e + \theta > 0$ and $\Sigma^e \in [1, \frac{\zeta^2}{2\theta}]$, from (48) we conclude that $Z \leq 0$. If $Z = 0$ then the assertion follows. Thus, we consider $Z < 0$. Since $\Sigma^e > 0$, (46) can only be satisfied in the following two cases:

$$\begin{aligned} \text{Case 1: } Z\theta - 1 < 0 \quad \wedge \quad \gamma - Z^2 > 0, \\ \text{Case 2: } Z\theta - 1 > 0 \quad \wedge \quad \gamma - Z^2 < 0. \end{aligned} \tag{50}$$

Due to $Z < 0$ and $\theta > 0$ case 2 can be excluded, which concludes the proof. \square

The case $\zeta^2/2 - \theta = 0$ is of no particular interest as it corresponds to the case of stabilizing MPC, cf. Lemma 8. Therefore, the natural follow-up question is what happens in case of $\zeta^2/2 - \theta < 0$. Indeed, although one can show that $Z \geq 0$ in this case, we can prove strong convexity of $\tilde{\ell}_{2F}$ also for $\frac{\zeta^2}{2} - \theta < 0$, by adjusting the regularization parameter γ .

Proposition 13. *Let $\zeta^2/2 - \theta < 0$ and $\gamma > 1/(4\zeta^4)$. Then (37) is strictly dissipative with the linear storage function $\lambda^l(z)$.*

Proof. From (36) we know that $\Sigma^e \leq 1$. Then from (48) and $\theta + K^e > 0$ we conclude that $Z \geq 0$. If $Z = 0$ then the assertion follows ($\gamma - Z^2 = \gamma > 0$). Thus, we consider $Z > 0$. It holds that

$$\Sigma^e = 1 - Z(\theta + K^e) \stackrel{!}{=} \frac{\zeta^2}{2(\theta + K^e)} \Leftrightarrow K^e + \theta = \frac{1}{2Z}(1 \pm \sqrt{1 - 2Z\zeta^2}). \tag{51}$$

In particular, $1 - 2Z\zeta^2 \geq 0$, which, due to $Z, \zeta^2 > 0$, is equivalent to $Z^2 \leq \frac{1}{4\zeta^4}$. Thus, for $\gamma > \frac{1}{4\zeta^4}$, we have

$$Z^2 \leq \frac{1}{4\zeta^4} < \gamma, \tag{52}$$

i.e., $\gamma - Z^2 > 0$, which concludes the proof. \square

Without the restriction on γ , there is one problematic case, in which we indeed lose strict dissipativity due to $\gamma - Z^2 = 0$. According to Remark 11(b), for this to happen it is necessary that $\gamma = 1/\theta^2$. The following proposition deals with this special case.

Proposition 14. *Let $\gamma = 1/\theta^2$.*

- (a) *If $2\zeta^2 - \theta < 0$, then the optimal equilibrium pair (Σ^e, K^e) is not unique. In particular, (37) is not strictly dissipative (irrespective of the storage function λ), but dissipative with the linear storage function $\lambda^l(z)$.*
- (b) *If $2\zeta^2 - \theta = 0$, then (37) is dissipative with $\lambda^l(z)$ but not strictly dissipative.*
- (c) *If $2\zeta^2 - \theta > 0$, then (37) is strictly dissipative with $\lambda^l(z)$.*

Proof. We first calculate the stationary points that are equilibria. To this end, we use

$$0 = \Sigma - f(\Sigma, K) \Leftrightarrow \Sigma = \frac{\zeta^2}{2(\theta + K)} \tag{53}$$

and plug this state into the cost function ℓ_{2F} , i.e.,

$$\ell_{2F}\left(\frac{\zeta^2}{2(\theta + K)}, K\right) = \frac{1}{2} \left[\left(\frac{c^2}{2(\theta + K)} - 1 \right)^2 + \gamma K^2 \right] =: \hat{\ell}_{2F}(K). \tag{54}$$

Then we compute the stationary points of the *reduced cost function* $\hat{\ell}_{2F}(K)$ in the special case $\gamma = \frac{1}{\theta^2}$:

$$\hat{\ell}'_{2F}(K) = -\frac{2\zeta^2}{(2(\theta + K))^2} \left(\frac{\zeta^2}{2(\theta + K)} - 1 \right) + \frac{K}{\theta^2} = 0 \Leftrightarrow K = K_i, \quad i = 1, \dots, 4 \tag{55}$$

with

$$K_{1/2} := -\frac{\theta}{2} \pm \frac{\sqrt{\theta}\sqrt{\theta-2\zeta^2}}{2} \quad \text{and} \quad K_{3/4} := -\theta \pm \frac{\sqrt{2\theta\zeta^2}}{2}. \quad (56)$$

Since $K_4 = -\theta - \frac{\sqrt{2\theta\zeta^2}}{2}$ violates the constraint $K > -\theta$, we ignore this solution. Moreover, we only care about real solutions. Therefore, we have three distinct solutions if and only if $2\zeta^2 - \theta < 0$.⁶ Now we consider the three different cases in the Proposition.

Let $2\zeta^2 - \theta < 0$. Then the controls K_1 , K_2 , and K_3 satisfy (26) with Σ as in (53). The respective cost is given by

$$\hat{\ell}_{2F}(K_1) = \frac{(\zeta^2 - \theta)(\zeta^2 - \theta - \sqrt{-2\zeta^2\theta + \theta^2})}{(\theta + \sqrt{-2\zeta^2\theta + \theta^2})^2} = \frac{\theta - \zeta^2}{2\theta} = \hat{\ell}_{2F}(K_2) \quad (57)$$

and

$$\hat{\ell}_{2F}(K_3) = \frac{\zeta^2 - 2\sqrt{2\zeta^2\theta} + \theta}{2\theta}. \quad (58)$$

We can exclude a minimum of $\hat{\ell}_{2F}(K)$ on the boundary since $\hat{\ell}_{2F}(K) \rightarrow \infty$ for $K \searrow -\theta$ and for $K \rightarrow \infty$. Since

$$\hat{\ell}_{2F}(K_3) - \hat{\ell}_{2F}(K_1) = \frac{2\zeta^2 - 2\sqrt{2\zeta^2\theta} + \theta}{2\theta} = \frac{(\sqrt{2\zeta^2} - \sqrt{\theta})^2}{2\theta} > 0, \quad (59)$$

there are two optimal equilibria, characterized by K_1 and K_2 . Thus, strict dissipativity is out of the question. However, we argue that dissipativity with $\lambda^l(z)$ does hold. For this, we show that $\gamma - Z^2 = 0$, i.e., that $\tilde{\ell}_{2F}(\Sigma, K)$ is convex but not strongly convex. With the corresponding states

$$\Sigma_1 = \frac{\zeta^2}{\theta + \sqrt{\theta}\sqrt{\theta-2\zeta^2}} \quad \text{and} \quad \Sigma_2 = \frac{\zeta^2}{\theta - \sqrt{\theta}\sqrt{\theta-2\zeta^2}}, \quad (60)$$

a short calculation using

$$0 = \partial_K L_{2F}(\Sigma^e, K^e, \bar{\lambda}) = \gamma K^e + Z\Sigma^e \quad (61)$$

yields the associated Lagrange multipliers $Z_1 = \frac{1}{\theta} = Z_2$. In particular, we have

$$\gamma - Z_1^2 = 0 = \gamma - Z_2^2. \quad (62)$$

For $2\zeta^2 - \theta = 0$, we get the same result, i.e., dissipativity but not strict dissipativity.

Lastly, if $2\zeta^2 - \theta > 0$, then (Σ_3, K_3) with $\Sigma_3 = \sqrt{\frac{\zeta^2}{2\theta}}$ is the unique optimal equilibrium and an analogous calculation reveals that $\gamma - Z_3^2 > 0$, i.e., strong convexity of $\tilde{\ell}_{2F}$ and thus strict dissipativity. \square

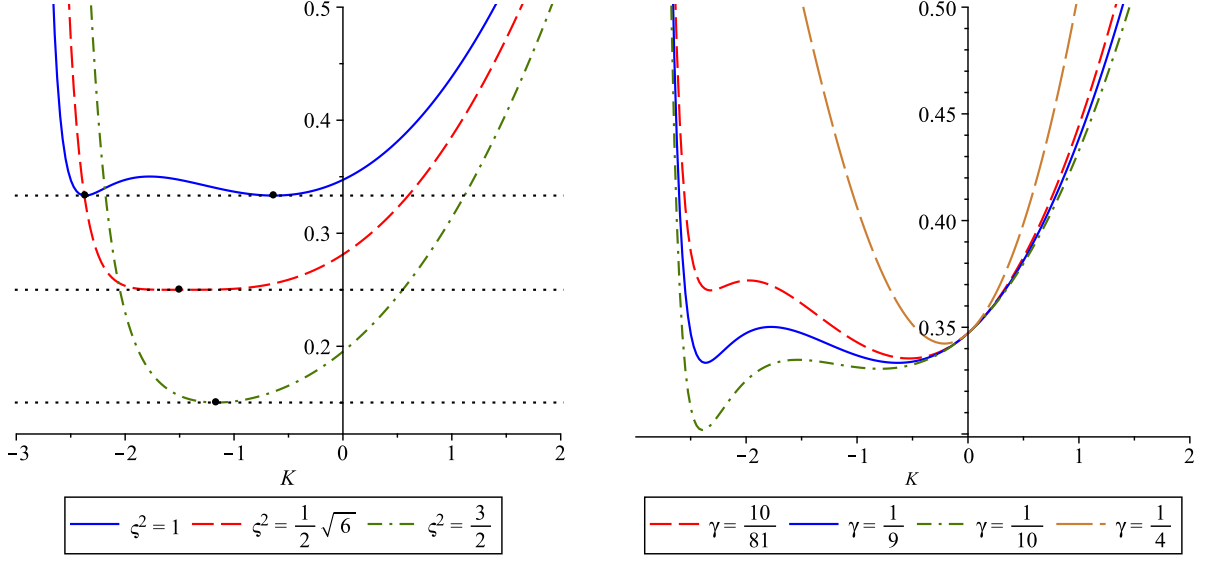
The three cases of Proposition 14 are exemplarily illustrated in Figure 1(a).

Remark 15. *Coinciding with the requirement on γ in Proposition 13, the reduced cost $\hat{\ell}_{2F}(K)$ from (54) is convex if and only if $\gamma \geq 1/(4\zeta^4)$, cf. Figure 1(b). In this case, there is a unique optimal equilibrium. However, as we will see in the subsequent section, in general, convexity of the reduced cost $\hat{\ell}_{2F}(K)$ does not transfer to the modified cost $\tilde{\ell}_{2F}(\Sigma, K)$.*

To summarize, we have shown strict dissipativity of (37) for $\zeta^2/2 - \theta > 0$ and arbitrary $\gamma > 0$, and for $\zeta^2/2 - \theta < 0$ provided that $\gamma > 1/(4\zeta^4)$. Furthermore, we identified cases in which strict dissipativity does not hold due to the existence of two optimal equilibria, which can only happen if $\gamma = 1/\theta^2$. Even for $\zeta^2/2 - \theta < 0$ and $\gamma \leq 1/(4\zeta^4)$, as long as $\gamma \neq 1/\theta^2$, our numerous simulations indicate that $\gamma - Z^2 > 0$. Thus, we conjecture that strict dissipativity (with a linear storage function) always holds. To prove this rigorously, one could solve $\nabla L_{2F}(\Sigma, K, \bar{\lambda}) = 0$ for arbitrary $\gamma > 0$. Ultimately, as (54) indicates, this requires finding the roots of a fourth-order polynomial. We avoid from carrying out this computation here for the sake of brevity.

A characteristic that can make the analysis more complicated when switching from linear to bilinear systems is the discrepancy between the control term K^2 in the stage cost and the bilinear term $(\theta + K)\Sigma$ in the dynamics. In fact, the analysis is simplified considerably if we replace the term penalizing the control effort, K^2 , with $(\theta + K)^2$ in $\ell_{2F}(\Sigma, K)$. In finding an optimal equilibrium, this is equivalent to setting θ to zero.

⁶For $2\zeta^2 - \theta = 0$ we have that $K_1 = K_2 = K_3$, i.e., only one stationary point.



(a) $\hat{\ell}_{2F}(K)$ for $\theta = 3$, $\gamma = 1/\theta^2$ and various values of ζ^2 with the respective minima

(b) $\hat{\ell}_{2F}(K)$ for $\zeta = 1$, $\theta = 3$, and various values of γ

Figure 1: (Non-)Convexity of the reduced cost $\hat{\ell}_{2F}(\Sigma, K)$ depending on ζ^2 (left) and on γ (right).

Proposition 16. *If, instead of $\ell_{2F}(\Sigma, K)$, the stage cost (35) in the OCP (37) is defined by*

$$\ell_{2F,\theta}(\Sigma, K) := \frac{1}{2}(\Sigma - 1)^2 + \frac{\gamma}{2}(K + \theta)^2, \quad (63)$$

then (37) is strictly dissipative with the linear storage function $\lambda^l(z)$.

Proof. To conclude strict dissipativity, we prove that $\tilde{\ell}_{2F,\theta}(\Sigma, K)$, defined analogously to (38) is strongly convex. To this end, we define $L_{2F,\theta}(\Sigma, K, \bar{\lambda})$ analogously to (40). Then

$$\partial_{\Sigma} L_{2F,\theta}(\Sigma, K, \bar{\lambda}) = \Sigma - 1 + 2\bar{\lambda}T(\theta + K) \quad (64)$$

and

$$\partial_K L_{2F,\theta}(\Sigma, K, \bar{\lambda}) = \gamma(\theta + K) + 2\bar{\lambda}T\Sigma. \quad (65)$$

With $Z = 2\bar{\lambda}T$, solving $\partial_{\Sigma} L_{2F,\theta}(\Sigma, K, \bar{\lambda}) = 0$ for Σ yields

$$\Sigma = 1 - Z(\theta + K). \quad (66)$$

Plugging this into $\partial_K L_{2F,\theta}(\Sigma, K, \bar{\lambda}) = 0$ results in

$$0 = \gamma(\theta + K) + Z(1 - Z(\theta + K)) = (\gamma - Z^2)(\theta + K) + Z. \quad (67)$$

From (67) we can exclude the case $\gamma - Z^2 = 0$ since $\gamma > 0$ and we know that at least one optimal equilibrium exists, i.e., (67) has at least one admissible solution. Thus, $\gamma - Z^2 \neq 0$, in which case

$$\theta + K = -\frac{Z}{\gamma - Z^2} \quad (68)$$

and therefore, according to (66),

$$\Sigma = 1 + \frac{Z^2}{\gamma - Z^2} = \frac{\gamma}{\gamma - Z^2}. \quad (69)$$

Since $\Sigma > 0$ and $\gamma > 0$, from (69) we infer that $\gamma - Z^2 > 0$, i.e., $\tilde{\ell}_{2F,\theta}(\Sigma, K)$ is strongly convex. \square

Note that (68)-(69) coincides with (46) in the case $\theta = 0$. For $\theta = 0$ the requirements of Proposition 12 are met and thus the result of Proposition 16 is not surprising. Although the stage cost $\ell_{2F,\theta}(\Sigma, K)$ is much easier to handle, the price to pay is the loss of optimal equilibria with $\Sigma^e \in (0, 1)$: we can see from (69) that $\Sigma^e = 1 + \frac{Z^2}{\gamma - Z^2} > 1$ since $\gamma - Z^2 > 0$.

We summarize our results for the 2F-cost in a similar form as for the L^2 -cost at the end of Section 4:

- For $\varsigma^2/2 - \theta > 0$, strict dissipativity holds with a linear storage function.
- For $\varsigma^2/2 - \theta < 0$ and $\gamma > 1/(4\varsigma^2)$, strict dissipativity holds with a linear storage function.
- For $\varsigma^2/2 - \theta < 0$ and $\gamma \leq 1/(4\varsigma^2)$, strict dissipativity fails to hold for some parameter values if $\gamma = 1/\theta^2$. Numerical evidence suggests that strict dissipativity always holds if $\gamma \neq 1/\theta^2$.
- If ℓ_{2F} is replaced by $\ell_{2F,\theta}$ from (63), then strict dissipativity holds for all parameter values.
- The case $\varsigma^2/2 - \theta = 0$ is again of minor interest since it corresponds to the stabilizing MPC case.

We emphasize once more that for the 2F stage cost considered in this section, proving strict dissipativity with a linear storage function is equivalent to proving strong convexity of $\tilde{\ell}_{2F}(\Sigma, K)$. This is in contrast to the L^2 -cost (13) considered in [8], where the modified cost was never convex, but for some parameters the OCP was nevertheless strictly dissipative with a linear storage function, cf. [8, Example 10]. In this sense, the W^2 -cost considered in the following section is more similar to the L^2 -cost as to the 2F-cost.

5.2 W^2 cost

The W^2 cost is designed to measure the distance between two PDFs. In our case, it differs only slightly from the cost in the previous section: Instead of $(\Sigma - 1)^2$, the square root of the current and the desired state is taken and a quadratic cost is inflicted on the distance thereof, i.e., $(\sqrt{\Sigma} - 1)^2$. In this one-dimensional case, this amounts to penalizing the difference in the standard deviation instead of in the variance. Surprisingly, this small difference changes the dissipativity analysis considerably.

Overall, the optimization problem in this section is given by

$$\begin{aligned} J_N(\mathring{\Sigma}, K) &:= \sum_{k=0}^{N-1} \ell_{W^2}(\Sigma(k), K(k)) = \sum_{k=0}^{N-1} \left[\frac{1}{2} \left(\sqrt{\Sigma(k)} - 1 \right)^2 + \frac{\gamma}{2} K(k)^2 \right] \rightarrow \min! \\ \text{subject to } \Sigma^+ &= \Sigma(k) + T(-2K_\theta(k)\Sigma(k) + \varsigma^2) = f(\Sigma(k), K(k)), \quad \Sigma(0) = \mathring{\Sigma}, \\ 0 &< K_\theta(k) < (\Sigma(k) + T\varsigma^2)/(2T\Sigma(k)). \end{aligned} \quad (70)$$

As before, (Σ^e, K^e) denotes an optimal equilibrium, i.e., a solution of

$$\min_{(\Sigma, K)} \frac{1}{2} \left(\sqrt{\Sigma} - 1 \right)^2 + \frac{\gamma}{2} K^2 \quad \text{s.t. } \Sigma - f(\Sigma, K) = 0. \quad (71)$$

The change in the cost function greatly influences the gradient and Hessian of the modified cost

$$\tilde{\ell}_{W^2}(\Sigma, K) = \frac{1}{2} \left(\sqrt{\Sigma} - 1 \right)^2 + \frac{\gamma}{2} K^2 - \ell_{W^2}(\Sigma^e, K^e) + \bar{\lambda} \left(-T(-2(\theta + K)\Sigma + \varsigma^2) \right). \quad (72)$$

With $Z = 2\bar{\lambda}T$, the gradient and the Hessian of $\tilde{\ell}_{W^2}(\Sigma, K)$ are given by

$$\nabla \tilde{\ell}_{W^2}(\Sigma, K) = \begin{pmatrix} \frac{\sqrt{\Sigma}-1}{2\sqrt{\Sigma}} \\ \gamma K \end{pmatrix} + Z \begin{pmatrix} \theta + K \\ \Sigma \end{pmatrix} \quad (73)$$

and

$$\nabla^2 \tilde{\ell}_{W^2}(\Sigma, K) = \begin{pmatrix} \frac{1}{4\Sigma^{3/2}} & Z \\ Z & \gamma \end{pmatrix}, \quad (74)$$

respectively. In particular, for any fixed Z , $\tilde{\ell}_{W^2}$ is not convex for sufficiently large Σ . This is in contrast to the previous section, where the Hessian was constant. While convexity is only a sufficient condition for strict dissipativity, the following example shows that the latter may also be lost.

Example 17. Consider (70) with the parameters

$$\varsigma = 10, \quad \theta = 2, \quad \gamma = 1/4, \quad \text{and} \quad T = 1/100. \quad (75)$$

We want to construct the modified cost $\tilde{\ell}_{W^2}(\Sigma, K)$. First, we determine the optimal equilibrium (Σ^e, K^e) and the corresponding Lagrange multiplier $\bar{\lambda}$. We formulate the Lagrange function associated to (71) and

solve the problem numerically. Note from (73) and (74) that the interest is in $Z = 2\bar{\lambda}T$ rather than in $\bar{\lambda}$. In particular, the optimal equilibrium is independent of the sampling time T . We get:

$$\Sigma^e \approx 10.2393012, \quad K^e \approx 2.8831457, \quad Z \approx -0.070394104. \quad (76)$$

With this, we can construct the modified cost $\tilde{\ell}_{W^2}(\Sigma, K)$, which is depicted in Figure 2. All pairs (Σ, K) illustrated in this figure satisfy the constraints (26). The white area depicts negative values, i.e., pairs (Σ, K) in which (23) is violated. Thus, (strict) dissipativity does not hold with a linear storage function.

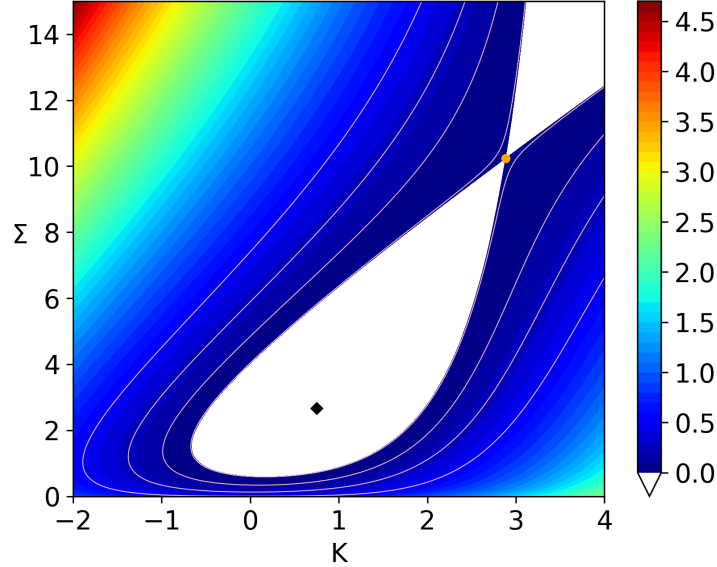


Figure 2: Modified cost $\tilde{\ell}_{W^2}(\Sigma, K)$ for Example 17. The optimal equilibrium (Σ^e, K^e) is illustrated by the orange circle. The white area represents negative values; the black diamond marks the minimum of the depicted area.

Note that in the above example, $\varsigma^2/2 - \theta > 0$. Moreover, γ is such that the reduced cost

$$\hat{\ell}_{W^2}(K) := \ell_{W^2}\left(\frac{\varsigma^2}{2(\theta + K)}, K\right) = \frac{1}{2} \left[\left(\sqrt{\frac{c^2}{2(\theta + K)}} - 1 \right)^2 + \gamma K^2 \right] \quad (77)$$

is convex.⁷ In short, the properties that were used in the previous section (see Propositions 12 and 13 and Remark 15) to guarantee strict dissipativity of (37) are not appropriate to prove strict dissipativity of (70). The following proposition and its proof give one of two reasons why this is the case.

Proposition 18. *If $\frac{\varsigma^2}{2} - \theta > 0$, then for sufficiently low sampling times $T > 0$, (70) is not dissipative with a linear storage function $\lambda(z) := \bar{\lambda}z$.*

Proof. The idea of the proof is to show that the modified cost $\tilde{\ell}_{W^2}$ can assume negative values, which violates (23). To this end, we first note that

$$\lim_{\Sigma \rightarrow \infty} \tilde{\ell}_{W^2}(\Sigma, K) = \text{sgn} \left((K + \theta)Z + \frac{1}{2} \right) \cdot \infty. \quad (78)$$

Next, we show that $Z < 0$. From the Lagrange function

$$L_{W^2}(\Sigma, K, \lambda) := \frac{1}{2} \left(\sqrt{\Sigma} - 1 \right)^2 + \frac{\gamma}{2} K^2 + \lambda \left[-T(-2(\theta + K)\Sigma + \varsigma^2) \right] \quad (79)$$

⁷One can show that $\hat{\ell}_{W^2}$ is convex for $\gamma > \frac{5^5}{2^{16}\varsigma^4}$. However, as this fact is not crucial for the subsequent statements we refrain from giving a rigorous proof.

we deduce

$$\partial_K L_{W^2}(\Sigma, K, \bar{\lambda}) = \partial_K \tilde{\ell}_{W^2}(\Sigma, K) = \gamma K + Z\Sigma \quad (80)$$

and

$$\partial_K L_{W^2}(\Sigma, K, \bar{\lambda}) = 0 \quad \Leftrightarrow \quad \begin{cases} \Sigma = -\frac{\gamma K}{Z}, & Z \neq 0 \\ K = 0, & Z = 0 \end{cases}, \quad (81)$$

recalling that $Z = 2\bar{\lambda}T$. Due to $\partial_K L_{W^2}(\Sigma^e, K^e, \bar{\lambda}) = 0$, we can exclude $Z = 0$: If $Z = 0$, then $K^e = 0$ and thus $\Sigma^e = 1$, cf. (73). But this contradicts (32) since $\frac{\varsigma^2}{2} - \theta > 0$, i.e., $\frac{\varsigma^2}{2\theta} > 1$. Thus, we have $\Sigma^e = -\gamma K^e/Z$ and $K^e \neq 0$, which, together with Lemma 8, results in $K^e > 0$. Then due to $\gamma > 0$ and $\Sigma^e > 0$ we arrive at $Z < 0$.⁸

Due to $Z < 0$, the term $(K + \theta)Z$ from (78) decreases as K increases. Taking into account the control constraint (26), we consider the limiting case of

$$K \nearrow \frac{\Sigma + T\varsigma^2}{2T\Sigma} - \theta, \quad (82)$$

which, due to $\Sigma \rightarrow \infty$, cf. (78), results in

$$K \nearrow \frac{1}{2T} - \theta. \quad (83)$$

Hence,

$$(K + \theta)Z + \frac{1}{2} \searrow \frac{Z}{2T} + \frac{1}{2} \quad \text{as} \quad K \rightarrow \frac{1}{2T} - \theta. \quad (84)$$

Thus, if $\frac{Z}{2T} + \frac{1}{2} < 0$, then $\text{sgn}((K + \theta)Z + \frac{1}{2}) = -1$ for large enough admissible K . In this case, (Σ^e, K^e) cannot be a global minimum, contradicting dissipativity. As in the previous section, cf. Remark 9, the product $\bar{\lambda}T$ and thus Z is constant in T . Hence, due to $Z < 0$, one can always achieve $\frac{Z}{2T} + \frac{1}{2} < 0$ for small enough $T > 0$. \square

Remark 19. The result of Proposition 18 is very similar to the L^2 case, see the end of Section 4. The only difference is that, for the W^2 cost, the statement depends on the sampling time $T > 0$. In this instance, the scaling of the stage cost mentioned in Remark 9 would in fact influence the result, such that the problems with using linear storage functions in the case of $\frac{\varsigma^2}{2} - \theta > 0$ potentially dissipate. For instance, one checks that the OCP (70) with parameters

$$\varsigma = 5, \quad \theta = 2, \quad \gamma = 1/4, \quad \text{and} \quad T = 1 \quad (85)$$

is strictly dissipative with storage function $\lambda^l(z)$.

Another difference to the previous section is hinted in Figure 2: In addition to the optimal equilibrium (Σ^e, K^e) , there exists a second stationary state of $\tilde{\ell}_{W^2}$ at $(\Sigma, K) \approx (2.6621866, 0.749609)$. In contrast, apart from the degenerate case $\gamma = 1/\theta^2$, $\tilde{\ell}_{2F}$ exhibits a unique stationary point for a fixed Z , cf. Lemma 10. The following proposition and the proof thereof shows that for $\tilde{\ell}_{W^2}$, we usually have two admissible stationary points.

Proposition 20. The modified cost $\tilde{\ell}_{W^2}(\Sigma, K)$ has at most two admissible stationary points.

Proof. From the gradient (73) we infer that $K = -Z\Sigma/\gamma$ and therefore,

$$0 = \frac{\sqrt{\Sigma} - 1}{2\sqrt{\Sigma}} + Z(\theta + K) = \frac{\sqrt{\Sigma} - 1}{2\sqrt{\Sigma}} + Z\theta - \frac{Z^2\Sigma}{\gamma} =: h(\Sigma) \quad (86)$$

If $h(\Sigma)$ has a unique admissible stationary point, then at most two admissible solutions for (86) can exist, i.e., the assertion follows. To this end, we look at the first derivative of $h(\Sigma)$:

$$h'(\Sigma) = \frac{1}{4\Sigma^{3/2}} - \frac{Z^2}{\gamma} = 0 \quad \Leftrightarrow \quad \Sigma = \left(\frac{\gamma}{4Z^2}\right)^{2/3} =: \Sigma_{W^2}^s \quad (87)$$

for $Z \neq 0$, and $\Sigma_{W^2}^s$ is admissible. If $Z = 0$, then only one admissible stationary point of $\tilde{\ell}_{W^2}(\Sigma, K)$ exists and it is given by $(\Sigma^e, K^e) = (1, 0)$. \square

⁸Analogously, one can prove that $Z > 0$ if $\frac{\varsigma^2}{2} - \theta < 0$.

In addition to the asymptotic behavior and the stationary points, we need to take the boundary of the control constraint set into account.

Example 21. Consider (70) with the parameters

$$\varsigma = 3/4, \quad \theta = 3/2, \quad \gamma = 1/5, \quad \text{and} \quad T = 1/10. \quad (88)$$

As in Example 17, we determine the optimal equilibrium (Σ^e, K^e) and the associated Z numerically:

$$\Sigma^e \approx 0.4679159, \quad K^e \approx -0.8989304, \quad Z \approx 0.3842274. \quad (89)$$

The reduced cost $\tilde{\ell}_{W^2}$, cf. (77), is convex, since $\frac{5^5}{2^{16}\varsigma^4} = \frac{5^5}{2^{8}3^4} < \frac{1}{5} = \gamma$. Furthermore, the Hessian of the modified cost $\tilde{\ell}_{W^2}$ evaluated at (Σ^e, K^e) is positive definite:

$$\nabla^2 \tilde{\ell}_{W^2}(\Sigma^e, K^e) \approx \begin{pmatrix} 0.7810671 & Z \\ Z & \gamma \end{pmatrix} \Rightarrow \left| \nabla^2 \tilde{\ell}_{W^2}(\Sigma^e, K^e) \right| \approx 0.00858275 > 0. \quad (90)$$

Moreover, the second stationary point of $\tilde{\ell}_{W^2}$ at approximately

$$(0.5044150447, -0.9690503190) =: (\Sigma^s, K^s) \quad (91)$$

is not an issue, since $\tilde{\ell}_{W^2}(\Sigma^s, K^s) \approx 9.2315 \cdot 10^{-6} > 0$. However, we face problems when looking at the boundary $K = -\theta$ respective $\Sigma = 0$:

$$\tilde{\ell}_{W^2}(0, K) = \frac{1}{2} + \frac{K^2}{2} - \ell_{W^2}(\Sigma^e, K^e) - \frac{Z\varsigma^2}{2}, \quad (92)$$

which is minimal at $K = 0$ with

$$\tilde{\ell}_{W^2}(0, 0) = \frac{1}{2} - \ell_{W^2}(\Sigma^e, K^e) - \frac{Z\varsigma^2}{2}. \quad (93)$$

Analogously, at the boundary $K = -\theta$, we have:

$$\tilde{\ell}_{W^2}(\Sigma, -\theta) = \frac{1}{2} \left(\sqrt{\Sigma} - 1 \right)^2 + \frac{\gamma}{2} \theta^2 - \ell_{W^2}(\Sigma^e, K^e) - \frac{Z\varsigma^2}{2}, \quad (94)$$

which is minimal at $\Sigma = 1$ with

$$\tilde{\ell}_{W^2}(1, -\theta) = \frac{\gamma}{2} \theta^2 - \ell_{W^2}(\Sigma^e, K^e) - \frac{Z\varsigma^2}{2}. \quad (95)$$

In total, we require that

$$\min \left\{ \frac{1}{2}, \frac{\gamma}{2} \theta^2 \right\} - \ell_{W^2}(\Sigma^e, K^e) - \frac{Z\varsigma^2}{2} \geq 0. \quad (96)$$

Otherwise, due to continuity of $\tilde{\ell}_{W^2}$, strict dissipativity with this storage function does not hold. Indeed, in this example, we have

$$\min \left\{ \frac{1}{2}, \frac{\gamma}{2} \theta^2 \right\} - \ell_{W^2}(\Sigma^e, K^e) - \frac{Z\varsigma^2}{2} \approx -0.0137857 < 0, \quad (97)$$

see Figure 3, and thus, no strict dissipativity with $\lambda^l(z)$.

The above examples show that the W^2 cost is more difficult to manage than the stage cost from the previous section. The most striking difference is that positive definiteness of the Hessian $\nabla^2 \tilde{\ell}_{W^2}(\Sigma^e, K^e)$ is not sufficient since $\nabla^2 \tilde{\ell}_{W^2}(\Sigma, K)$ is not constant. Although this criterion can be used to conclude local convexity in a neighborhood of (Σ^e, K^e) (which implies strict dissipativity if state and control are constrained to that region), in general it will not yield global convexity. Another difficulty arises due to the second stationary state of ℓ_{W^2} (see Proposition 20), for which, on top of that, there is no analytic formula, as opposed to the previous section, cf. Lemma 10. Moreover, one now needs to take into account the boundary, see Example 21. In Section 5.1, nothing was said about the boundary. This is because for constant Hessians, checking the (sign of the) determinant is sufficient/equivalent.

This does not mean, however, that strict dissipativity with $\lambda^l(z)$ never holds, as the following example shows.

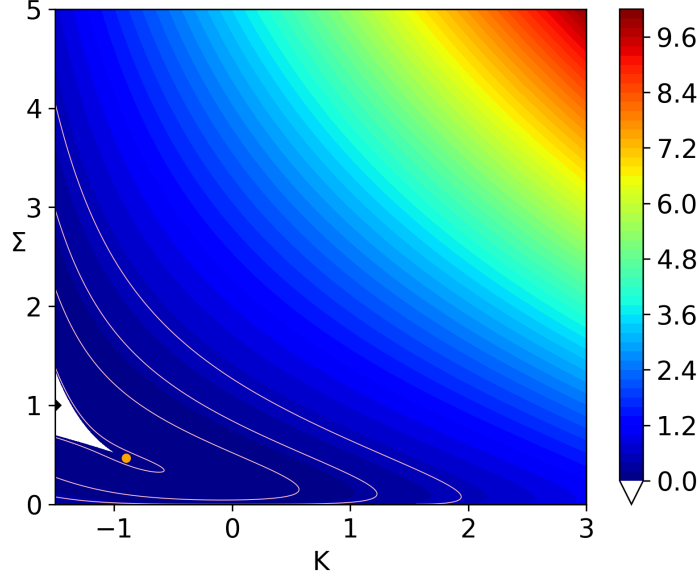


Figure 3: Modified cost $\tilde{\ell}_{L^2}(\Sigma, K)$ for Example 21. The optimal equilibrium (Σ^e, K^e) is illustrated by the orange circle. The white area represents negative values; the black diamond marks the minimum of the depicted area.

Example 22. Consider (70) with the parameters

$$\varsigma = 2/3, \quad \theta = 3/2, \quad \gamma = 1/3, \quad \text{and} \quad T = 1/10. \quad (98)$$

Numerically, we identify the optimal equilibrium and the corresponding value for Z :

$$\Sigma^e \approx 0.1865912, \quad K^e \approx -0.3090422, \quad Z \approx 0.5520844. \quad (99)$$

We also determine the second stationary state of $\tilde{\ell}_{W^2}$ numerically:

$$(0.8642951, -1.4314914) =: (\Sigma^s, K^s). \quad (100)$$

We get $\tilde{\ell}_{W^2}(\Sigma^s, K^s) \approx 0.0767537 > 0$ and $\min \left\{ \frac{1}{2}, \frac{\gamma}{2} \theta^2 \right\} - \ell_{W^2}(\Sigma^e, K^e) - \frac{Z\varsigma^2}{2} \approx 0.07506316 > 0$, i.e., both the second stationary point and lowest possible boundary value are positive. Due to $Z > 0$,

$$\begin{aligned} \tilde{\ell}_{W^2}(\Sigma, K) &= \frac{1}{2} \left(\sqrt{\Sigma} - 1 \right)^2 + \frac{\gamma}{2} K^2 - \ell_{W^2}(\Sigma^e, K^e) + \bar{\lambda} \left(-T(-2(\theta + K)\Sigma + \varsigma^2) \right) \\ &= \frac{1}{2} \left(\sqrt{\Sigma} - 1 \right)^2 + \frac{\gamma}{2} K^2 - \ell_{W^2}(\Sigma^e, K^e) + Z \left((\theta + K)\Sigma - \varsigma^2/2 \right) \rightarrow \infty \end{aligned} \quad (101)$$

for $\Sigma \rightarrow \infty$ or $K \rightarrow \infty$. Since no other stationary point exists, we can find a function $\varrho \in \mathcal{K}_\infty$ such that the dissipativity inequality (23) holds with $\lambda^l(z)$. Figure 4 depicts the corresponding modified cost $\tilde{\ell}_{W^2}(\Sigma, K)$.

Therefore, we cannot conclude that looking for a linear storage function is always in vain, either. Moreover, modifying the cost function ℓ_{W^2} by penalizing $(\theta + K)^2$ instead of K^2 as in the previous section (see Proposition 16) does not (always) help: the modified cost function yields the same optimal equilibria as considering $\theta = 0$. In particular, $\varsigma^2/2 - \theta > 0$ holds. However, in contrast to the previous section, this property does not guarantee strict dissipativity, cf. Example 17.⁹

Despite the similarity of the two cost functions ℓ_{W^2} and ℓ_{2F} , the results are very different. In fact, regarding dissipativity with the linear storage function $\lambda^l(z)$, the Wasserstein cost ℓ_{W^2} has more in common with the L^2 cost considered in [8]. This includes that, when running numerical simulations, the MPC closed loop converges to the optimal equilibrium (Σ^e, K^e) – even for the parameters in Examples 17 and 21, see Figures 5 and 6. These figures indicate that the turnpike property holds even in cases where the linear storage function fails.

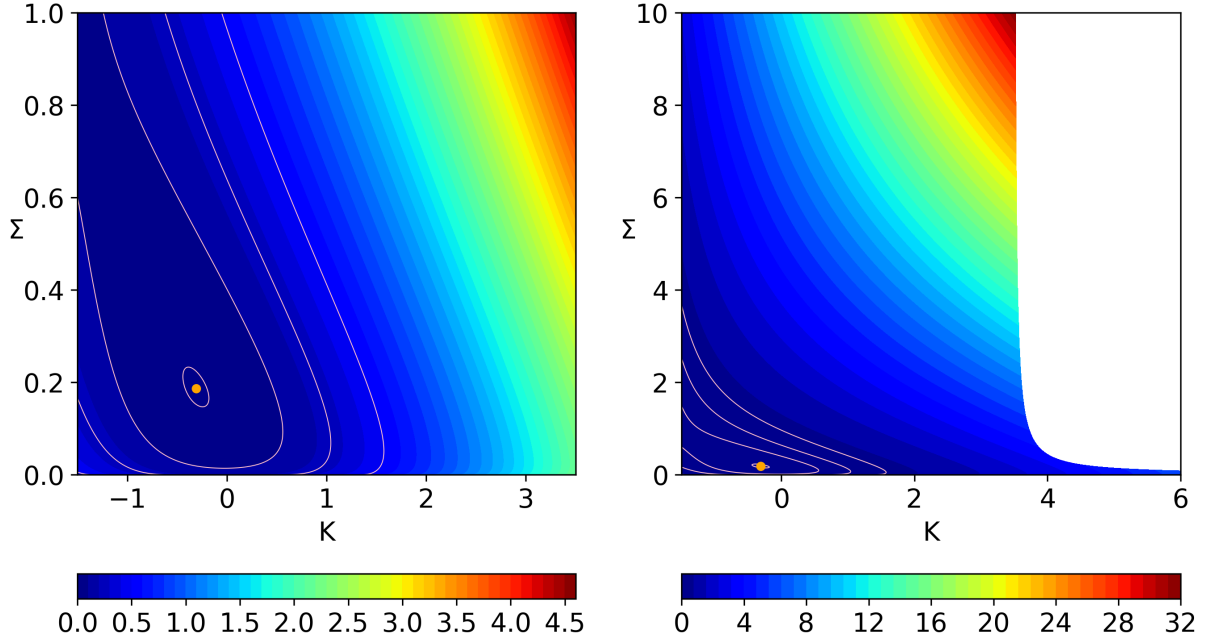


Figure 4: Modified cost $\tilde{\ell}_{W^2}(\Sigma, K)$ for Example 22 zoomed in (left) and zoomed out (right). The optimal equilibrium (Σ^e, K^e) is illustrated by the orange circle. The white area on the right plot is due to control constraints (26).

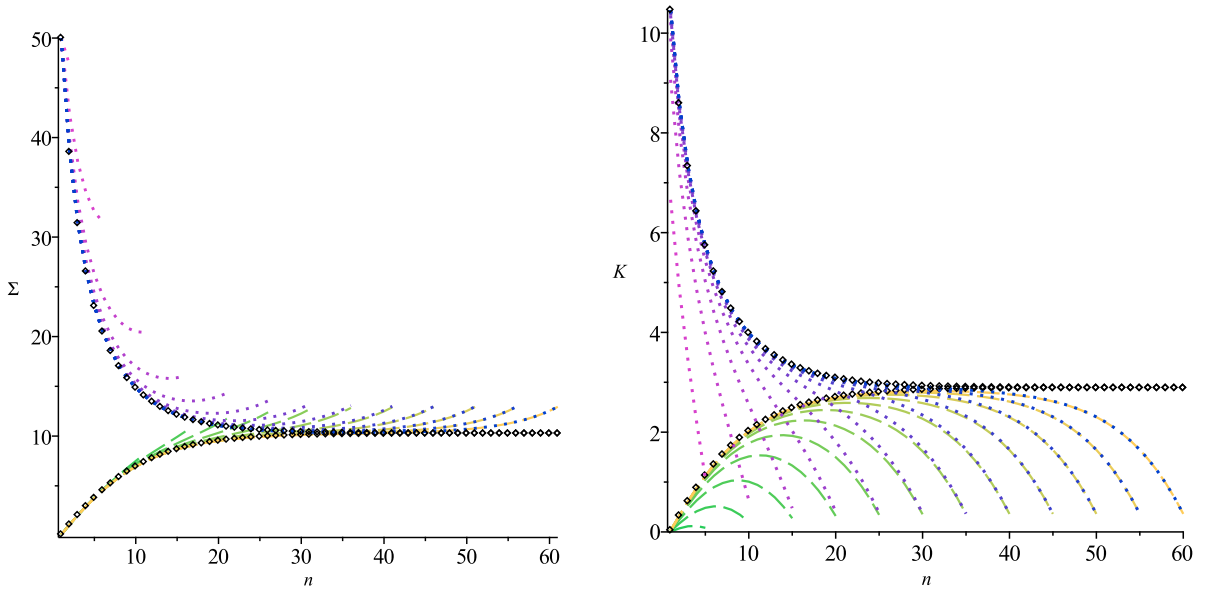


Figure 5: Open loop optimal trajectories for various horizons N between 1 and 60 and MPC closed loop trajectories for two different initial conditions, indicating turnpike behavior in Example 17; state Σ (left) and control K (right).

Due to the close relationship between dissipativity and the turnpike property, see Section 3, this strongly suggests that strict dissipativity does indeed hold, but with a nonlinear storage function. Thus, in the rest of this section, we revisit these examples with the nonlinear storage function

$$\lambda^s(z) := \alpha\sqrt{z+1}. \quad (102)$$

⁹Neither does $\varsigma^2/2 - \theta < 0$, see Example 21.

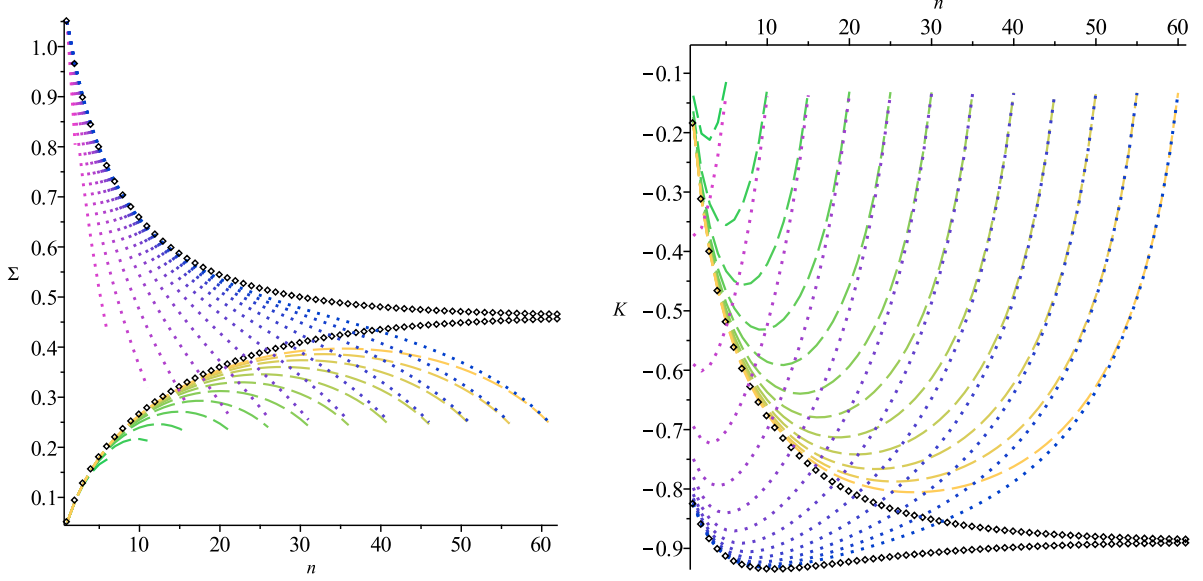


Figure 6: Open loop optimal trajectories for various horizons N between 1 and 60 and MPC closed loop trajectories for two different initial conditions, indicating turnpike behavior in Example 21; state Σ (left) and control K (right).

The parameter $\alpha \in \mathbb{R}$ is chosen such that the optimal equilibrium (Σ^e, K^e) is a stationary point of the new modified cost

$$\tilde{\ell}_{W^2}^s(\Sigma, K) := \ell_{W^2}(\Sigma, K) - \ell_{W^2}(\Sigma^e, K^e) + \lambda^s(\Sigma) - \lambda^s(\Sigma^+). \quad (103)$$

One notable advantage of $\lambda^s(z)$ over $\lambda^l(z)$ is the asymptotic behavior of the modified cost: While

$$\lim_{\Sigma \rightarrow \infty} \tilde{\ell}_{W^2}(\Sigma, K) = \text{sgn} \left((K + \theta)Z + \frac{1}{2} \right) \cdot \infty \quad (104)$$

depends on the value of Z (one can easily prove that Z is negative if and only if $\zeta^2/2 - \theta > 0$ by using the gradient (73) and Lemma 8 for computing the sign of K), with the nonlinear storage function $\lambda^s(z)$ we get $\tilde{\ell}_{W^2}^s(\Sigma, K) \rightarrow \infty$ for $\Sigma \rightarrow \infty$ or $K \rightarrow \infty$ irrespective of the value of α . Thus, when looking for a suitable/promising storage function $\lambda(z)$, the asymptotic behavior of $\lambda(z)$ should be compared to that of the cost $\ell(\Sigma, K)$.

Ideally, the storage function can be chosen such that the Hessian $\nabla^2 \tilde{\ell}(\Sigma, K)$ is constant. Then one can avoid checking everything by foot, i.e., the boundary values and the stationary points of the modified cost function. Unfortunately, the Hessian $\nabla^2 \tilde{\ell}_{W^2}^s(\Sigma, K)$ is not constant. However, the level sets in Figure 7 clearly suggest that strict dissipativity holds for both Example 17 (left) and 21.

We first consider Example 17. Our numerical calculations yield $\alpha \approx -23.5996705$ and three stationary states of $\tilde{\ell}_{W^2}^s(\Sigma, K)$ in total, of which one is not admissible due to $K > -\theta$. The remaining two are (Σ^e, K^e) and $(\Sigma^s, K^s) \approx (265.4413283, 41.51437144)$. The second one is admissible but not a problem since $\tilde{\ell}_{W^2}^s(\Sigma^s, K^s) \approx 86.1249768 > 0$. At the boundary $\Sigma = 0$, the minimum is attained at $K = 0$, with a value of $\tilde{\ell}_{W^2}^s(0, 0) \approx 6.816477628 > 0$. For $K = -\theta$, a minimum of approximately $2.40236824 > 0$ is attained at $\Sigma \approx 5.897079388$. Thus, strict dissipativity holds with $\lambda^s(z)$.

Next, we look at Example 21. In this case, from numerical calculations we get $\alpha \approx 4.6552057$. In addition to (Σ^e, K^e) , the new modified cost $\tilde{\ell}_{W^2}^s(\Sigma, K)$ exhibits a second admissible stationary point at $(\Sigma^s, K^s) \approx (0.8398851754, -1.424465947)$ with $\tilde{\ell}_{W^2}^s(\Sigma^s, K^s) \approx 0.00136419 > 0$.¹⁰ At the boundary $\Sigma = 0$, the minimum is attained at $K = 0$, with a value of $\tilde{\ell}_{W^2}^s(0, 0) \approx 0.2401417173 > 0$. Lastly, for $K = -\theta$, the minimum value is attained at approximately $\Sigma_\theta := 0.9095436678$, with a value of $\tilde{\ell}_{W^2}^s(\Sigma_\theta, -\theta) \approx 0.005474005283 > 0$. Thus, again, strict dissipativity holds with $\lambda^s(z)$.

As before, we end our analysis by summarizing our main results in short form:

¹⁰In total, three stationary states exist, but the third one is not admissible due to the control constraint $K > -\theta$.

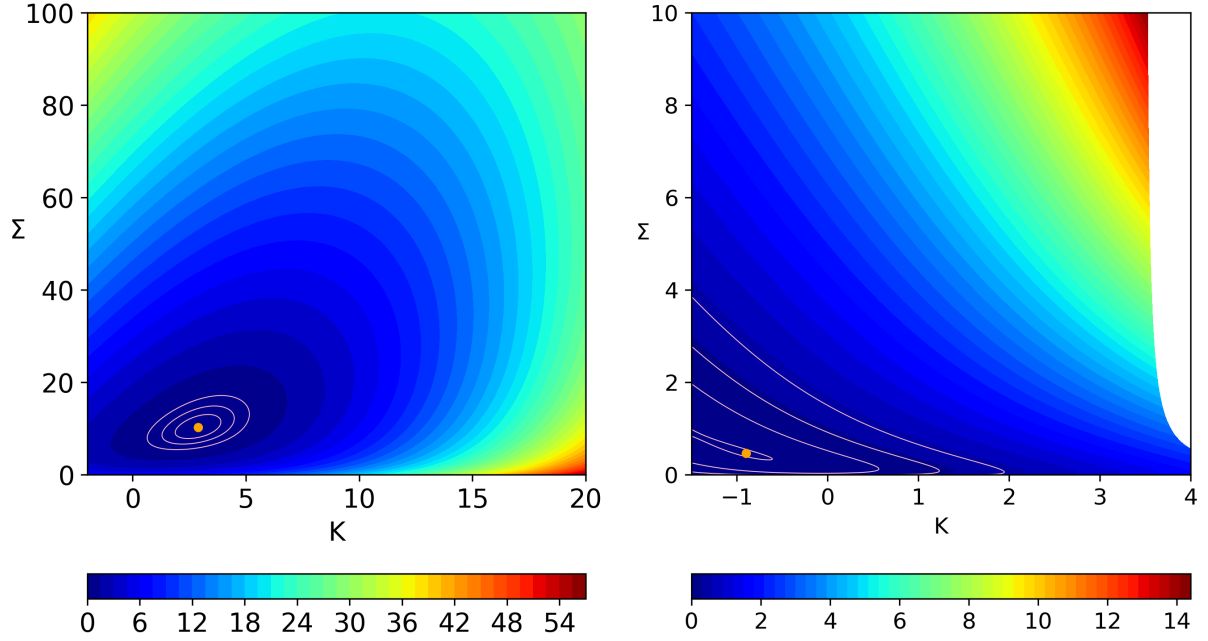


Figure 7: New modified cost $\tilde{\ell}_{W^2}^s(\Sigma, K)$ for Examples 17 (left) and 21 (right). The optimal equilibrium (Σ^e, K^e) is illustrated by the orange circle. The white area on the right plot is due to the control constraints (26).

- For $\varsigma^2/2 - \theta > 0$ and small enough sampling times $T > 0$, strict dissipativity cannot hold with a linear storage function. For large enough $T > 0$ strict dissipativity may hold, but has to be checked on a case-by-case basis.
- In the case $\varsigma^2/2 - \theta < 0$, independent of the sampling time T , for certain values of θ and ς strict dissipativity holds with a linear storage function, for other values it does not.
- For various values of θ and ς strict dissipativity holds with the nonlinear storage function (102). However, the verification is tedious and must be done on a case-by-case basis.
- Numerical verification of the turnpike property suggests that strict dissipativity holds for many parameters for which the analytical verification is not (yet) possible.

Concluding, the Wasserstein metric, which is in many aspects very suitable for measuring distances of PDFs, does not allow for a simple analysis of strict dissipativity, although our results give strong indication that strict dissipativity holds for many parameter values.

6 Conclusion

In this work we have analyzed whether a particular optimal control problem with bilinear dynamics connected to the Fokker-Planck equation is strictly dissipative. To this end, we have considered two cost functions: an often suggested Wasserstein cost, W^2 , and a quadratic cost function commonly used in tracking objectives, $2F$. We have found that for the latter cost, a linear storage function can be used to prove strict dissipativity for a large parameter set (if not for all parameters). The linear storage function is convenient due to its similarity with the Lagrange function. However, we have demonstrated that it is unsuitable if the W^2 cost is used. To show that the optimal control problems are strictly dissipative in the W^2 case, we have introduced a class of nonlinear storage functions.

Regarding MPC, our results suggest that all considered costs (including the L^2 -cost considered in [8]) work well, although a mathematically rigorous proof for large sets of parameters could only be achieved for the $2F$ -cost. Unfortunately, this is the only cost that is not derived from a metric for general PDFs, and thus it is only applicable to the Gaussian setting. It will be an interesting question for further

research to see whether it is possible to extend this cost and the associated strict dissipativity results beyond the Gaussian case.

References

- [1] D. Angeli, R. Amrit, and J.B. Rawlings. On average performance and stability of economic model predictive control. *IEEE Trans. Autom. Control*, 57(7):1615–1626, 2012.
- [2] M. Annunziato and A. Borzì. Optimal control of probability density functions of stochastic processes. *Math. Model. Anal.*, 15(4):393–407, 2010.
- [3] M. Annunziato and A. Borzì. A Fokker-Planck control framework for multidimensional stochastic processes. *J. Comput. Appl. Math.*, 237(1):487–507, 2013.
- [4] T. Breiten, K. Kunisch, and L. Pfeiffer. Control strategies for the Fokker-Planck equation. *ESAIM: COCV*, 24(2):741–763, 2018.
- [5] T. Damm, L. Grüne, M. Stieler, and K. Worthmann. An exponential turnpike theorem for dissipative discrete time optimal control problems. *SIAM J. Control Optim.*, 52(3):1935–1957, 2014.
- [6] M. Diehl, R. Amrit, and J.B. Rawlings. A Lyapunov function for economic optimizing model predictive control. *IEEE Trans. Autom. Control*, 56:703–707, 2011.
- [7] T. Faulwasser, L. Grüne, and M.A. Müller. Economic nonlinear model predictive control. *Foundations and Trends® in Systems and Control*, 5(1):1–98, 2018.
- [8] A. Fleig and L. Grüne. On dissipativity of the Fokker-Planck equation for the Ornstein-Uhlenbeck process. *IFAC-PapersOnLine*. to appear.
- [9] A. Fleig and L. Grüne. Estimates on the minimal stabilizing horizon length in model predictive control for the Fokker-Planck equation. *IFAC-PapersOnLine*, 49(8):260–265, 2016.
- [10] A. Fleig and L. Grüne. L^2 -tracking of Gaussian distributions via model predictive control for the Fokker-Planck equation. *Vietnam J. Math.*, 46(4):915–948, 2018.
- [11] C.R. Givens and R.M. Shortt. A class of wasserstein metrics for probability distributions. *Michigan Math. J.*, 31(2):231–240, 1984.
- [12] L. Grüne. Economic receding horizon control without terminal constraints. *Automatica*, 49:725–734, 2013.
- [13] L. Grüne and M.A. Müller. On the relation between strict dissipativity and the turnpike property. *Syst. Contr. Lett.*, 90:45–53, 2016.
- [14] L. Grüne and J. Pannek. *Nonlinear Model Predictive Control. Theory and Algorithms*. Springer, London, 2nd edition, 2017.
- [15] L. Grüne and M. Stieler. Asymptotic stability and transient optimality of economic MPC without terminal conditions. *J. Proc. Control*, 24(8):1187–1196, 2014.
- [16] W. Hahn. *Stability of motion*, volume 138. Springer, 1967.
- [17] S. Primak, V. Kontorovich, and V. Lyandres. *Stochastic methods and their applications to communications*. John Wiley & Sons, Inc., Hoboken, NJ, 2004.
- [18] P.E. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005.
- [19] J.B. Rawlings, D. Bonné, J.B. Jorgensen, A.N. Venkat, and S.B. Jorgensen. Unreachable setpoints in model predictive control. *IEEE Transactions on Automatic Control*, 53(9):2209–2215, 2008.
- [20] J.B. Rawlings, D.Q. Mayne, and M.M. Diehl. *Model Predictive Control: Theory and Design*. Nob Hill Publishing, 2nd edition, 2017.

- [21] H. Risken. *The Fokker-Planck Equation*, volume 18 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, 2nd edition, 1989.
- [22] S. Roy, M. Annunziato, and A. Borzi. A Fokker-Planck feedback control-constrained approach for modelling crowd motion. *J. Comput. Theor. Transp.*, 45(6):442–458, 2016.
- [23] F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010.
- [24] J.C. Willems. Dissipative dynamical systems. I. General theory. *Arch. Rational Mech. Anal.*, 45:321–351, 1972.