

Optimal Controller and Actuator Design for Nonlinear Parabolic Systems

M. Sajjad Edalatzadeh, *Member, IEEE*, and Kirsten A. Morris, *Senior Member, IEEE*,

Abstract—Many physical systems are modeled by nonlinear parabolic differential equations, such as the Kuramoto-Sivashinsky (KS) equation. In this paper, the existence of a concurrent optimal controller and actuator design is established for semilinear systems. Optimality equations are provided. The results are shown to apply to optimal controller/actuator design for the Kuramoto-Sivashinsky equation and also nonlinear diffusion.

I. INTRODUCTION

The best actuator design can improve performance and significantly reduce the cost of the control in distributed parameter systems; see for example [1]. The optimal actuator design problem of linear systems has been reviewed in various contexts, see [2], [3]. For linear partial differential equations (PDEs), the existence of an optimal actuator location has been proven in the literature. In [4], it is proven that an optimal actuator location exists for a linear system with quadratic cost function if the input operator is compact and continuously depends on actuator locations. Further conditions on operators and cost functions are needed to guarantee the convergence in numerical schemes [4]. Similar results have been obtained for H_2 and H_∞ controller design objectives [5], [6].

Nonlinearities can have a significant effect on dynamics, and such systems cannot be accurately modelled by linear differential equations. Control of systems modelled by nonlinear partial differential equations (PDE's) has been studied for a number of applications, including wastewater treatment systems [7], steel cooling plants [8], oil extraction through a reservoir [9], solidification models in metallic alloys [10], thermistors [11], Schlögl model [12], [13], FitzHugh–Nagumo system [13], micro-beam model [14], static elastoplasticity [15], type-II superconductivity [16], Fokker-Planck equation [17], Schrödinger equation with bilinear control [18], Cahn-Hilliard-Navier-Stokes system [19], wine fermentation process [20], time-dependent Kohn-Sham model [21], elastic crane-trolley-load system [22], and railway track model [23]. A review of PDE-constrained optimization theory can be found in the books [24], [25], [26]. State-constrained optimal control of PDEs has also been studied. In [27], the authors investigated the structure of Lagrange multipliers for state constrained optimal control problem of linear elliptic PDEs. Research on optimal control of PDEs, such as [28], [29], has focused on partial differential equations with certain structures. Optimal

control of differential equations in abstract spaces has rarely been discussed [30]. This paper extends previous results to abstract differential equations without an assumption of stability.

Few studies have discussed optimal control for general classes of nonlinear distributed parameter systems; and even less have looked into actuator design problem of such systems. Using a finite dimensional approximation of the original partial differential equation model, optimal actuator location has been addressed for some applications. Antoniadis and Christofides [31] investigated the optimal actuator and sensor location problem for a transport-reaction process using a finite-dimensional model. Similarly, Lou and Christofides [32] studied the optimal actuator and sensor location of Kuramoto-Sivashinsky equation using a finite-dimensional approximation. Other research concerned with optimal actuator location for nonlinear distributed parameter systems can be found in [33], [34], [35]. To our knowledge, there are no theoretical results on optimal actuator design of nonlinear distributed parameter systems.

The results of this paper apply to the Kuramoto-Sivashinsky (KS) equation. This equation was derived by Kuramoto to model angular phase turbulence in reaction-diffusion systems [36], and by Sivashinsky for modeling plane flame propagation [37]. It also models film layer flow on an inclined plane [38], directional solidification of dilute binary alloys [39], growth and saturation of the potential of dissipative trapped-ion [40], and terrace edge evolution during step-flow growth [41]. From system theoretic perspective, Christofides and Armaou studied the global stabilization of KS equation using distributed output feedback control [42]. Lou and Christofides investigated the optimal actuator/sensor placement for control of KS equation by approximating the model with a finite dimensional system [32]. Gomes et al. also studied the actuator placement problem for KS equation using numerical algorithms [43]. The feedback control as well as optimal actuator arrangement of multidimensional KS equation has been studied in [44]. Controllability of KS equation has also been studied [45], [46]. Optimal control of KS equation using maximum principle was studied in [47]. Optimal control of KS equation with pointwise state and mixed control-state constraints was studied in [48]. Liu and Krstic studied boundary control of KS equation in [49]. Al Jamal and Morris studied the relationship between stability and stabilization of linearized and nonlinear KS equation [50].

The paper is organized as follows. Section 2 is a short section containing notation and definitions. Section 3 discusses the existence of an optimal input together with an optimal actuator design to nonlinear parabolic systems. In section 4,

M. Sajjad Edalatzadeh is with Department of Applied Mathematics, University of Waterloo, Waterloo, Canada e-mail: msedalatzadeh@uwaterloo.ca

Kirsten A. Morris is with Faculty of Mathematics, University of Waterloo, Waterloo, Canada e-mail: kmorris@uwaterloo.ca

Manuscript received

the worst initial condition is discussed. In section 5 and 6, Kuramoto-Sivashinsky equation and nonlinear heat equation are discussed, respectively.

II. NOTATION AND DEFINITIONS

Let \mathbb{X} be a reflexive Banach space. The notation $\mathbb{X}_1 \hookrightarrow \mathbb{X}_2$ means that the space \mathbb{X}_1 is densely and continuously embedded in \mathbb{X}_2 . Also, letting $I \subset \mathbb{R}$ to be a possibly unbounded interval, the Banach space $C^s(I; \mathbb{X})$ consists of all Hölder continuous \mathbb{X} -valued functions with exponent s equipped with norm

$$\|\mathbf{x}\|_{C^s(I; \mathbb{X})} = \|\mathbf{x}\|_{C(I; \mathbb{X})} + \sup_{t, s \in I} \frac{\|\mathbf{x}(t) - \mathbf{x}(s)\|}{|t - s|^s}. \quad (1)$$

The Banach space $c^s(I, \mathbb{X})$ is the space of little-Hölder continuous functions with exponent s defined as all $\mathbf{x} \in C^s(I; \mathbb{X})$ such that

$$\lim_{\delta \rightarrow 0} \sup_{t, s \in I, |t-s| \leq \delta} \frac{\|\mathbf{x}(t) - \mathbf{x}(s)\|}{|t - s|^s} = 0. \quad (2)$$

Also, $W^{m,p}(I; \mathbb{X})$ is the space of all strongly measurable functions $\mathbf{x} : I \rightarrow \mathbb{X}$ for which $\|\mathbf{x}(t)\|_{\mathbb{X}}$ is in $W^{m,p}(I, \mathbb{R})$. For simplicity of notation, when I is an interval, the corresponding space will be indicated without the braces; for example $C([0, \tau]; \mathbb{X})$ will be indicated by $C(0, \tau; \mathbb{X})$.

Let \mathcal{A} be the generator of an analytic semigroup e^{At} on \mathbb{X} . For every $p \in [1, \infty]$ and $\alpha \in (0, 1)$, the interpolation space $D_{\mathcal{A}}(\alpha, p)$ is defined as the set of all $\mathbf{x}_0 \in \mathbb{X}$ such that the function

$$t \mapsto v(t) := \left\| t^{1-\alpha-1/p} \mathcal{A} e^{t\mathcal{A}} \mathbf{x}_0 \right\| \quad (3)$$

belongs to $L^p(0, 1)$ [51, Section 2.2.1]. The norm on this space is

$$\|\mathbf{x}_0\|_{D_{\mathcal{A}}(\alpha, p)} = \|\mathbf{x}_0\| + \|v\|_{L^p(0,1)}.$$

The Banach space $\mathbb{W}(0, \tau)$ is the set of all $\mathbf{x}(\cdot) \in W^{1,p}(0, \tau; \mathbb{X}) \cap L^p(0, \tau; D(\mathcal{A}))$ with norm [52, Section II.2]

$$\|\mathbf{x}\|_{\mathbb{W}(0, \tau)} = \|\dot{\mathbf{x}}\|_{L^p(0, \tau; \mathbb{X})} + \|\mathcal{A}\mathbf{x}\|_{L^p(0, \tau; \mathbb{X})}.$$

Definition 1. *The operator $\mathcal{A} : D(\mathcal{A}) \rightarrow \mathbb{X}$ is said to have maximal L^p regularity if for every $\mathbf{f} \in L^p(0, \tau; \mathbb{X})$, $1 < p < \infty$, the equation*

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathcal{A}\mathbf{x}(t) + \mathbf{f}(t), & t > 0, \\ \mathbf{x}(0) = \mathbf{x}_0, \end{cases} \quad (4)$$

admits a unique solution in $\mathbb{W}(0, \tau)$ that satisfies (4) almost everywhere on $[0, \tau]$.

Every generator of an analytic semigroup on a Hilbert space has maximal L^p regularity [53, Theorem 4.1].

III. NONLINEAR PARABOLIC SYSTEMS

Let $\mathbf{x}(t)$ and $\mathbf{u}(t)$ be the state and input taking values in reflexive Banach spaces \mathbb{X} and \mathbb{U} , respectively. Also, let \mathbf{r} denote the actuator design parameter that takes value in a compact set K_{ad} of a topological space \mathbb{K} . Consider the following initial value problem (IVP):

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathcal{A}\mathbf{x}(t) + \mathcal{F}(\mathbf{x}(t)) + \mathcal{B}(\mathbf{r})\mathbf{u}(t), & t > 0, \\ \mathbf{x}(0) = \mathbf{x}_0. \end{cases} \quad (5)$$

The linear operator $\mathcal{A} : D(\mathcal{A}) \rightarrow \mathbb{X}$ is assumed to have maximal L^p regularity. In particular, if \mathcal{A} is associated with a sesquilinear form that is bounded and coercive with respect to $\mathbb{V} \hookrightarrow \mathbb{X}$, it generates an analytic semigroup on \mathbb{X} [54, Lemma 36.5 and Theorem 36.6].

The nonlinear operator $\mathcal{F}(\cdot)$ maps a reflexive Banach space \mathbb{V} to \mathbb{X} where $D_{\mathcal{A}}(1/p, p) \hookrightarrow \mathbb{V} \hookrightarrow \mathbb{X}$. The operator $\mathcal{F}(\cdot)$ is locally Lipschitz continuous; that is, for every bounded set D in \mathbb{V} , there is a positive number $L_{\mathcal{F}}$ such that

$$\|\mathcal{F}(\mathbf{x}_2) - \mathcal{F}(\mathbf{x}_1)\|_{\mathbb{X}} \leq L_{\mathcal{F}} \|\mathbf{x}_2 - \mathbf{x}_1\|_{\mathbb{V}}, \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in D. \quad (6)$$

When there is no ambiguity, the norm on \mathbb{X} will not be explicitly indicated.

For each $\mathbf{r} \in \mathbb{K}$, the input operator $\mathcal{B}(\mathbf{r})$ is a linear bounded operator that maps the input space \mathbb{U} into the state space \mathbb{X} and it is continuous with respect to \mathbf{r} :

$$\lim_{\mathbf{r}_n \rightarrow \mathbf{r}_0} \|\mathcal{B}(\mathbf{r}_n) - \mathcal{B}(\mathbf{r}_0)\| = 0, \quad (7)$$

where the convergence $\mathbf{r}_n \rightarrow \mathbf{r}_0$ is with respect to the topology on \mathbb{K} .

For any positive numbers R_1 and R_2 , define the sets

$$B_{L^p(0, \tau; \mathbb{U})}(R_1) = \left\{ \mathbf{u} \in L^p(0, \tau; \mathbb{U}) : \|\mathbf{u}\|_p \leq R_1 \right\}, \quad (8)$$

$$B_{\mathbb{V}}(R_2) = \{ \mathbf{x}_0 \in \mathbb{V} : \|\mathbf{x}_0\|_{\mathbb{V}} \leq R_2 \}. \quad (9)$$

Definition 2. [52, Definition 3.1.i] *(strict solution) The function $\mathbf{x}(\cdot)$ is said to be a strict solution of (5) if $\mathbf{x}(0) = \mathbf{x}_0$, $\mathbf{x} \in \mathbb{W}(0, \tau)$, and $\mathbf{x}(t)$ satisfies (5) for almost every $t \in [0, \tau]$.*

Lemma 3. [55, Proposition 2.2 and Corollary 2.3] *Let $\tau_0 > \tau$ and $p \in (1, \infty)$ be given. If \mathcal{A} has maximal L^p regularity, then there exists a constant c_{τ_0} independent of τ such that for all $\tau \in (0, \tau_0]$ and $\mathbf{v} \in W^{1,p}(0, \tau; \mathbb{X}) \cap L^p(0, \tau; D(\mathcal{A}))$,*

$$\begin{aligned} & \|\dot{\mathbf{v}}\|_{L^2(0, \tau; \mathbb{X})} + \|\mathcal{A}\mathbf{v}\|_{L^2(0, \tau; \mathbb{X})} \\ & \leq M_{\tau_0} \left(\|\dot{\mathbf{v}} + \mathcal{A}\mathbf{v}\|_{L^2(0, \tau; \mathbb{X})} + \|\mathbf{v}(0)\|_{D_{\mathcal{A}}(1/p, p)} \right). \end{aligned}$$

Furthermore, if $\mathbf{v}(0) = 0$,

$$\|\mathbf{v}\|_{C(0, \tau; D_{\mathcal{A}}(1/p, p))} \leq M_{\tau_0} \left(\|\dot{\mathbf{v}}\|_{L^2(0, \tau; \mathbb{X})} + \|\mathcal{A}\mathbf{v}\|_{L^2(0, \tau; \mathbb{X})} \right).$$

Theorem 4. *For every pair $R_1 > 0$, $R_2 > 0$, there is $\tau > 0$ and $\delta > 0$ such that the IVP (5) admits a unique strict solution $\mathbf{x} \in \mathbb{W}(0, \tau)$, $\|\mathbf{x}\|_{\mathbb{W}(0, \tau)} \leq \delta$ for all $(\mathbf{u}, \mathbf{r}, \mathbf{x}_0) \in B_{L^p(0, \tau; \mathbb{U})}(R_1) \times K_{ad} \times B_{\mathbb{V}}(R_2)$.*

Proof. The proof of this theorem follows the same line as that of [55, Theorem 2.1] with some modifications. Let \mathbf{w} solve the linear equation

$$\begin{cases} \dot{\mathbf{w}}(t) = \mathcal{A}\mathbf{w}(t) + \mathcal{F}(\mathbf{x}_0) + \mathcal{B}(\mathbf{r})\mathbf{u}(t), & t \in (0, \tau], \\ \mathbf{w}(0) = \mathbf{x}_0. \end{cases} \quad (10)$$

Define for an arbitrary number $\rho > 0$ the set

$$\Sigma_{\rho, \tau} = \left\{ \mathbf{v} \in \mathbb{W}(0, \tau) : \mathbf{v}(0) = \mathbf{x}_0, \|\mathbf{v} - \mathbf{w}\|_{\mathbb{W}(0, \tau)} \leq \rho \right\}. \quad (11)$$

Because $\mathbf{w}(\cdot) \in \mathbb{W}(0, \tau)$, $\mathbf{w}(\cdot) \in C(0, \tau; \mathbb{V})$. Define $\phi(\tau; R_1, R_2) = \|\mathbf{w} - \mathbf{x}_0\|_{C(0, \tau; \mathbb{V})}$ where here \mathbf{x}_0 indicates the constant function in $C(0, \tau; \mathbb{V})$ that equals \mathbf{x}_0 . Note that

$$\lim_{\tau \rightarrow 0} \phi(\tau; R_1, R_2) = 0. \quad (12)$$

According to Lemma 3, there is a constant M independent of τ such that

$$\|\mathbf{v} - \mathbf{x}_0\|_{C(0,\tau;\mathbb{V})} \leq M\rho + \phi(\tau; R_1, R_2), \quad \forall \mathbf{v} \in \Sigma_{\rho,\tau}. \quad (13)$$

Consider the mapping $\gamma : \mathbb{W}(0, \tau) \rightarrow \mathbb{W}(0, \tau)$, $\mathbf{x}(\cdot) \mapsto \mathbf{v}(\cdot)$ defined by

$$\begin{cases} \dot{\mathbf{v}}(t) = \mathcal{A}\mathbf{v}(t) + \mathcal{F}(\mathbf{x}(t)) + \mathcal{B}(\mathbf{r})\mathbf{u}(t), & t \in (0, \tau], \\ \mathbf{v}(0) = \mathbf{x}_0. \end{cases} \quad (14)$$

It will now be shown that for some numbers ρ and τ the mapping γ defines a contraction on $\Sigma_{\rho,\tau}$ and hence has a unique fixed point.

Consider the linear equation

$$\begin{cases} \dot{\mathbf{v}}(t) - \dot{\mathbf{w}}(t) = \mathcal{A}(\mathbf{v}(t) - \mathbf{w}(t)) + \mathcal{F}(\mathbf{x}(t)), & t \in (0, \tau], \\ (\mathbf{v} - \mathbf{w})(0) = 0, \end{cases}$$

Use Lemma 3 together with Lipschitz continuity of \mathcal{F} , let $L_{\mathcal{F}}$ be the Lipschitz constant of \mathcal{F} over the ball $B(\mathbf{x}_0, M\rho + \phi(\tau; R_1, R_2))$. It follows that

$$\begin{aligned} \|\mathbf{v} - \mathbf{w}\|_{\mathbb{W}(0,\tau)} &\leq M \|\mathcal{F}(\mathbf{x}(t)) - \mathcal{F}(\mathbf{x}_0)\|_p \\ &\leq ML_{\mathcal{F}}\tau^{\frac{1}{p}} \|\mathbf{x} - \mathbf{x}_0\|_{C(0,\tau;\mathbb{V})} \\ &\leq M^2 L_{\mathcal{F}}\tau^{\frac{1}{p}} (M\rho + \phi(\tau; R_1, R_2)). \end{aligned} \quad (15)$$

Furthermore, for any $\mathbf{x}_1, \mathbf{x}_2 \in \Sigma_{\rho,\tau}$, define $\mathbf{v}_1 = \gamma(\mathbf{x}_1)$ and $\mathbf{v}_2 = \gamma(\mathbf{x}_2)$, then Lemma 3 yields

$$\begin{aligned} \|\mathbf{v}_2 - \mathbf{v}_1\|_{\mathbb{W}(0,\tau)} &\leq M \|\mathcal{F}(\mathbf{x}_2) - \mathcal{F}(\mathbf{x}_1)\|_p \\ &\leq ML_{\mathcal{F}}\tau^{\frac{1}{p}} \|\mathbf{x}_2 - \mathbf{x}_1\|_{C(0,\tau;\mathbb{V})} \\ &\leq M^2 L_{\mathcal{F}}\tau^{\frac{1}{p}} \|\mathbf{x}_2 - \mathbf{x}_1\|_{\mathbb{W}(0,\tau)}. \end{aligned} \quad (16)$$

Choose ρ and τ so that

$$\begin{aligned} M^2 L_{\mathcal{F}}\tau^{\frac{1}{p}} &< 1, \\ M^2 L_{\mathcal{F}}\tau^{\frac{1}{p}} (M\rho + \phi(\tau; R_1, R_2)) &\leq \rho. \end{aligned}$$

The Contraction Mapping Theorem ensures that the mapping γ has a unique fixed point in $\Sigma_{\rho,\tau}$. This fixed point is the unique solution \mathbf{x} to (5). Also, from the definition (11), every \mathbf{x} in $\Sigma_{\rho,\tau}$ satisfies

$$\|\mathbf{x}\|_{\mathbb{W}(0,\tau)} \leq \|\mathbf{w}\|_{\mathbb{W}(0,\tau)} + \rho. \quad (17)$$

Let $L_{\mathcal{F}}$ be the Lipschitz constant of \mathcal{F} over the ball $B(0, \|\mathbf{x}_0\|)$. Proposition 2.2 in [55] yields

$$\begin{aligned} \|\mathbf{w}\|_{\mathbb{W}(0,\tau)} &\leq M(\|\mathbf{x}_0\|_{\mathbb{V}} + \|\mathcal{F}(\mathbf{x}_0) + \mathcal{B}(\mathbf{r})\mathbf{u}(t)\|_p) \\ &\leq M(\|\mathbf{x}_0\|_{\mathbb{V}} + \tau^{\frac{1}{p}} L_{\mathcal{F}} \|\mathbf{x}_0\|_{\mathbb{V}} + \|\mathcal{B}(\mathbf{r})\|_{\mathcal{L}(\mathbb{X},\mathbb{U})} \|\mathbf{u}(t)\|_p) \\ &\leq M(R_2 + \tau^{\frac{1}{p}} L_{\mathcal{F}} R_2 + R_1 \underbrace{\max_{\mathbf{r} \in K_{ad}} \|\mathcal{B}(\mathbf{r})\|_{\mathcal{L}(\mathbb{X},\mathbb{U})}}_{\delta}). \end{aligned}$$

Defining

$$\delta = M(R_2 + \tau^{\frac{1}{p}} L_{\mathcal{F}} R_2 + R_1 \max_{\mathbf{r} \in K_{ad}} \|\mathcal{B}(\mathbf{r})\|_{\mathcal{L}(\mathbb{X},\mathbb{U})}),$$

yields the required upper-bound on $\|\mathbf{x}\|_{\mathbb{W}(0,\tau)}$. \square

Definition 5. Let $\mathbf{x}(t)$ be the strict solution to (5). The mapping $\mathcal{S}(\mathbf{u}, \mathbf{r}, \mathbf{x}_0) : B_{L^p(0,\tau;\mathbb{U})}(R_1) \times K_{ad} \times B_{\mathbb{V}}(R_2) \rightarrow \mathbb{W}(0, \tau)$, $(\mathbf{u}(t), \mathbf{r}, \mathbf{x}_0) \mapsto \mathbf{x}(t)$, is called the solution map.

An embedding $D(\mathcal{A}) \hookrightarrow \mathbb{X}$ where $D(\mathcal{A})$ is compact in \mathbb{X} ensures that the space $W^{1,p}(0, \tau; \mathbb{X}) \cap L^p(0, \tau, D(\mathcal{A}))$ is compactly embedded in $C^s(0, \tau; \mathbb{V})$, $0 \leq s < 1$ [56, Theorem 5.2]. Since $C^s(0, \tau; \mathbb{V}) \hookrightarrow C(0, \tau; \mathbb{V})$, it follows that the space $W^{1,p}(0, \tau; \mathbb{X}) \cap L^p(0, \tau, D(\mathcal{A}))$ is compactly embedded in $C(0, \tau; \mathbb{V})$.

Theorem 6. If the embedding $D(\mathcal{A}) \hookrightarrow \mathbb{X}$ is compact then the solution map is weakly continuous with respect to $(\mathbf{u}(t), \mathbf{r}, \mathbf{x}_0) \in L^p(0, \tau; \mathbb{U}) \times \mathbb{K} \times \mathbb{V}$.

Proof. The weak continuity of the solution map with respect to $\mathbf{u}(t)$ is shown in [30, Lemma 2.12]. Weak continuity with respect to $(\mathbf{u}(t), \mathbf{r}, \mathbf{x}_0)$ follows from a similar proof. Choose any weakly convergent sequences $\{\mathbf{u}_n(t)\} \subset L^p(0, \tau; \mathbb{U})$, $\{\mathbf{x}_0^n\} \subset \mathbb{V}$, and $\{\mathbf{r}_n\} \subset \mathbb{K}$. Since sets $B_{L^p(0,\tau;\mathbb{U})}(R_1)$ and $B_{\mathbb{V}}(R_2)$ are bounded, closed, convex subsets of Banach spaces $L^p(0, \tau; \mathbb{U})$ and \mathbb{V} , respectively; these sets are weakly closed [26, Theorem 2.11]. This implies that there are $\mathbf{u}^o \in B_{L^p(0,\tau;\mathbb{U})}(R_1)$ and $\mathbf{x}_0 \in B_{\mathbb{V}}(R_2)$ such that

$$\mathbf{u}_n \rightharpoonup \mathbf{u}^o \text{ in } B_{L^p(0,\tau;\mathbb{U})}(R_1), \quad (18)$$

$$\mathbf{x}_0^n \rightharpoonup \mathbf{x}_0 \text{ in } B_{\mathbb{V}}(R_2). \quad (19)$$

Since the set K_{ad} is a compact subset of \mathbb{K}

$$\mathbf{r}_n \rightarrow \mathbf{r}^o \text{ in } K_{ad}. \quad (20)$$

It will be shown that $\mathcal{B}_{\mathbf{r}_n} \mathbf{u}_n(t)$ converges weakly to $\mathcal{B}_{\mathbf{r}^o} \mathbf{u}^o(t)$ in $L^p(0, \tau; \mathbb{X})$. For every $\mathbf{z} \in L^q(0, \tau; \mathbb{X})$, $1/q = 1 - 1/p$,

$$\begin{aligned} I &:= \langle \mathbf{z}, \mathcal{B}_{\mathbf{r}_n} \mathbf{u}_n - \mathcal{B}_{\mathbf{r}^o} \mathbf{u}^o \rangle_{L^q(0,\tau;\mathbb{X}^*), L^p(0,\tau;\mathbb{X})} \\ &= \langle \mathbf{z}, \mathcal{B}_{\mathbf{r}_n} \mathbf{u}_n - \mathcal{B}_{\mathbf{r}^o} \mathbf{u}_n \rangle_{L^q(0,\tau;\mathbb{X}^*), L^p(0,\tau;\mathbb{X})} \\ &\quad + \langle \mathbf{z}, \mathcal{B}_{\mathbf{r}^o} \mathbf{u}_n - \mathcal{B}_{\mathbf{r}^o} \mathbf{u}^o \rangle_{L^q(0,\tau;\mathbb{X}^*), L^p(0,\tau;\mathbb{X})}. \end{aligned} \quad (21)$$

Taking the adjoint and norm yield

$$\begin{aligned} I &\leq \|\mathcal{B}_{\mathbf{r}_n} - \mathcal{B}_{\mathbf{r}^o}\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \int_0^\tau \|\mathbf{u}_n(t)\|_{\mathbb{U}} \|\mathbf{z}(t)\| dt \\ &\quad + \left| \int_0^\tau \langle \mathcal{B}_{\mathbf{r}^o}^* \mathbf{z}(t), \mathbf{u}_n(t) - \mathbf{u}^o(t) \rangle_{\mathbb{U}^*, \mathbb{U}} dt \right|. \end{aligned}$$

Use Hölder inequality and let $\mathbf{v}(t) = \mathcal{B}_{\mathbf{r}^o}^* \mathbf{z}(t)$, it follows that

$$\begin{aligned} I &\leq \|\mathcal{B}_{\mathbf{r}_n} - \mathcal{B}_{\mathbf{r}^o}\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \|\mathbf{u}_n\|_{L^p(0,\tau;\mathbb{U})} \|\mathbf{z}\|_{L^q(0,\tau;\mathbb{X})} \\ &\quad + \left| \langle \mathbf{u}_n - \mathbf{u}^o, \mathbf{v} \rangle_{L^p(0,\tau;\mathbb{U}), L^q(0,\tau;\mathbb{U}^*)} \right|. \end{aligned}$$

The convergence of the first term follows from (7). The second term converges to zero because $\mathbf{u}_n \rightharpoonup \mathbf{u}^o$ in $L^p(0, \tau; \mathbb{U})$. Combining these yields

$$\mathcal{B}_{\mathbf{r}_n} \mathbf{u}_n \rightharpoonup \mathcal{B}_{\mathbf{r}^o} \mathbf{u}^o \text{ in } L^p(0, \tau; \mathbb{X}). \quad (22)$$

Using Theorem 4, the corresponding solution $\mathbf{x}_n(t)$ is a bounded sequence in the reflexive Banach space $L^p(0, \tau; D(\mathcal{A})) \cap W^{1,p}(0, \tau; \mathbb{X})$. Thus, there is a subsequence of $\mathbf{x}_n(t)$ such that

$$\mathbf{x}_{n_k} \rightharpoonup \mathbf{x} \text{ in } \mathbb{W}(0, \tau). \quad (23)$$

This in turn implies that the sequence $\mathbf{x}_n(t)$ strongly converges to $\mathbf{x}(t)$ in $C(0, \tau; \mathbb{V})$. This together with Lipschitz continuity of $\mathcal{F}(\cdot)$ yields

$$\mathcal{F}(\mathbf{x}_{n_k}(t)) \rightarrow \mathcal{F}(\mathbf{x}(t)) \quad \text{in } L^p(0, \tau; \mathbb{X}). \quad (24)$$

This strong convergence also yields weak convergence in the same space, that is

$$\mathcal{F}(\mathbf{x}_{n_k}(t)) \rightharpoonup \mathcal{F}(\mathbf{x}(t)) \quad \text{in } L^p(0, \tau; \mathbb{X}). \quad (25)$$

Now apply (18), (22), (23), and (25) to the IVP (5); take the limit; notice that a solution to the IVP is unique; it follows that $\mathbf{x} = \mathcal{S}(\mathbf{u}, \mathbf{r}, \mathbf{x}_0)$. Deleting elements $\{\mathbf{x}_{n_k}(t)\}$ from $\{\mathbf{x}_n(t)\}$ and repeating the previous processing, knowing that a weak limit is unique, it follows that $\mathbf{x}_n(t) \rightharpoonup \mathbf{x}(t)$ in $\mathbb{W}(0, \tau)$. \square

IV. OPTIMAL ACTUATOR DESIGN

Consider a cost function $J(\mathbf{x}, \mathbf{u}, \mathbf{r}) : \mathbb{W}(0, \tau) \times L^p(0, \tau; \mathbb{U}) \times \mathbb{K} \rightarrow \mathbb{R}$ that is bounded below and weakly lower-semicontinuous with respect to \mathbf{x} , \mathbf{u} , and \mathbf{r} . For a fixed initial condition $\mathbf{x}_0 \in B_{\mathbb{V}}(R_2)$, consider the following optimization problem over the admissible input set U_{ad} and actuator design set K_{ad}

$$\begin{cases} \min & J(\mathbf{x}, \mathbf{u}, \mathbf{r}) \\ \text{s.t.} & \mathbf{x} = \mathcal{S}(\mathbf{u}, \mathbf{r}, \mathbf{x}_0), \\ & (\mathbf{u}, \mathbf{r}) \in U_{ad} \times K_{ad}. \end{cases} \quad (\text{P})$$

The set U_{ad} will be assumed a convex and closed set contained in the interior of $B_{L^p(0, \tau; \mathbb{U})}(R_1)$.

Theorem 7. *For every $\mathbf{x}_0 \in B_{\mathbb{V}}(R_2)$, there exists a control input $\mathbf{u}^\circ \in U_{ad}$ together with an actuator design $\mathbf{r}^\circ \in K_{ad}$ that solve the optimization problem (P).*

Proof. The proof of this theorem follows from standard analysis; see for example, [24, Theorem 1.45] and [57, Theorem 4.1] for a similar argument. Define

$$j(\mathbf{x}_0) := \inf_{(\mathbf{u}, \mathbf{r}) \in U_{ad} \times K_{ad}} J(\mathcal{S}(\mathbf{u}, \mathbf{r}, \mathbf{x}_0), \mathbf{u}, \mathbf{r}). \quad (26)$$

and let $(\mathbf{u}_n, \mathbf{r}_n)$ be the minimizing sequence:

$$\lim_{n \rightarrow \infty} J(\mathcal{S}(\mathbf{u}_n, \mathbf{r}_n, \mathbf{x}_0), \mathbf{u}_n, \mathbf{r}_n) = j(\mathbf{x}_0). \quad (27)$$

The set U_{ad} is closed and convex in the reflexive Banach space $L^p(0, \tau; \mathbb{U})$, so it is weakly closed. This implies that there is a subsequence of \mathbf{u}_n , denote it by the same symbol, that converges weakly to some elements \mathbf{u}° in U_{ad} . Because of compactness of K_{ad} , there is also a subsequence of \mathbf{r}_n , denote it by the same symbol, that strongly converges to \mathbf{r}° . Theorem 4 and Theorem 6 state that the solution map is bounded and weakly continuous in each variable. Thus, the corresponding state $\mathbf{x}_n = \mathcal{S}(\mathbf{u}_n, \mathbf{r}_n; \mathbf{x}_0)$ also weakly converges to $\mathbf{x}^\circ = \mathcal{S}(\mathbf{u}^\circ, \mathbf{r}^\circ; \mathbf{x}_0)$ in $\mathbb{W}(0, \tau)$. The cost function is weakly lower semi-continuous with respect to each \mathbf{x} , \mathbf{u} , and \mathbf{r} , this ensures that $(\mathbf{x}^\circ, \mathbf{u}^\circ, \mathbf{r}^\circ)$ minimizes the cost function. Therefore, $(\mathbf{u}^\circ, \mathbf{r}^\circ)$ is a solution to the optimization problem (P). \square

Definition 8. [24, Definition 1.29] *The operator $\mathcal{G} : \mathbb{X} \rightarrow \mathbb{Y}$ is said to be Gâteaux differentiable at $\mathbf{x} \in \mathbb{X}$ in the direction*

$\mathbf{p} \in \mathbb{X}$, if there is a linear bounded operator $\mathcal{G}'_{\mathbf{x}}$ such that for all real ϵ

$$\lim_{\epsilon \rightarrow 0} \|\mathcal{G}(\mathbf{x} + \epsilon \mathbf{p}) - \mathcal{G}(\mathbf{x}) - \epsilon \mathcal{G}'_{\mathbf{x}} \mathbf{p}\|_{\mathbb{Y}} = 0. \quad (28)$$

The optimality conditions are derived next after assuming that the problem has certain properties. Consider the assumptions:

- A1. The spaces \mathbb{X} and \mathbb{U} are Hilbert spaces and $p = 2$. The space \mathbb{K} is a Banach space.
- A2. Let $a : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{C}$ be a sesquilinear form (see [58, Chapter 4]), where $\mathbb{V} \hookrightarrow \mathbb{X}$, and let there be positive numbers α and β such that

$$\begin{aligned} |a(\mathbf{x}_1, \mathbf{x}_2)| &\leq \alpha \|\mathbf{x}_1\|_{\mathbb{V}} \|\mathbf{x}_2\|_{\mathbb{V}}, & \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{V}, \\ \operatorname{Re} a(\mathbf{x}, \mathbf{x}) &\geq \beta \|\mathbf{x}\|_{\mathbb{V}}^2, & \forall \mathbf{x} \in \mathbb{V}. \end{aligned}$$

The operator \mathcal{A} has an extension to $\bar{\mathcal{A}} \in \mathcal{L}(\mathbb{V}, \mathbb{V}^*)$ described by

$$\langle \bar{\mathcal{A}} \mathbf{v}, \mathbf{w} \rangle_{\mathbb{V}^*, \mathbb{V}} = a(\mathbf{v}, \mathbf{w}), \quad \forall \mathbf{v}, \mathbf{w} \in \mathbb{V}, \quad (29)$$

where \mathbb{V}^* denotes the dual of \mathbb{V} with respect to pivot space \mathbb{X} .

- A3. The cost function $J(\mathbf{x}, \mathbf{u}, \mathbf{r})$ is continuously Fréchet differentiable with respect to each variable.
- A4. The nonlinear operator $\mathcal{F}(\cdot)$ is Gâteaux differentiable. Indicate the Gâteaux derivative of $\mathcal{F}(\cdot)$ at \mathbf{x} in the direction \mathbf{p} by $\mathcal{F}'_{\mathbf{x}} \mathbf{p}$. Furthermore, the mapping $\mathbf{x} \mapsto \mathcal{F}'_{\mathbf{x}}$ is bounded; that is, bounded sets in \mathbb{V} are mapped to bounded sets in $\mathcal{L}(\mathbb{V}, \mathbb{X})$.
- A5. The control operator $\mathcal{B}(\mathbf{r})$ is Gâteaux differentiable with respect to \mathbf{r} from K_{ad} to $\mathcal{L}(\mathbb{U}, \mathbb{X})$. Indicate the Gâteaux derivative of $\mathcal{B}(\mathbf{r})$ at \mathbf{r}° in the direction \mathbf{r} by $\mathcal{B}'_{\mathbf{r}^\circ} \mathbf{r}$. Furthermore, the mapping $\mathbf{r}^\circ \mapsto \mathcal{B}'_{\mathbf{r}^\circ}$ is bounded; that is, bounded sets in \mathbb{K} are mapped to bounded sets in $\mathcal{L}(\mathbb{K}, \mathcal{L}(\mathbb{U}, \mathbb{X}))$.

Using these assumptions, the Gâteaux derivative of the solution map with respect to a trajectory $\mathbf{x}(t) = \mathcal{S}(\mathbf{u}(t), \mathbf{r}, \mathbf{x}_0)$ is calculated. The resulting map is a time-varying linear IVP. Let $\mathbf{g} \in L^p(0, \tau; \mathbb{X})$, consider the time-varying system

$$\begin{cases} \dot{\mathbf{h}}(t) = (\mathcal{A} + \mathcal{F}'_{\mathbf{x}(t)}) \mathbf{h}(t) + \mathbf{g}(t), \\ \mathbf{h}(0) = 0. \end{cases} \quad (30)$$

Lemma 9. [59, Corollary 5.2] *Let assumptions A1 and A2 hold. For any $\tau > 0$, let $\mathcal{P}(\cdot) : [0, \tau] \rightarrow \mathcal{L}(\mathbb{V}, \mathbb{X})$ be such that $\mathcal{P}(\cdot) \mathbf{x}$ is weakly measurable for all $\mathbf{x} \in \mathbb{V}$, and there exists an integrable function $h : [0, \tau] \rightarrow [0, \infty)$ such that $\|\mathcal{P}(t)\|_{\mathcal{L}(\mathbb{V}, \mathbb{X})} \leq h(t)$ for all $t \in [0, \tau]$. Then for every $\mathbf{x}_0 \in \mathbb{V}$ and $\mathbf{g} \in L^2(0, \tau; \mathbb{X})$, there exists a unique \mathbf{x} in $\mathbb{W}(0, \tau)$ such that*

$$\begin{cases} \dot{\mathbf{x}}(t) = (\mathcal{A} + \mathcal{P}(t)) \mathbf{x}(t) + \mathbf{g}(t), \\ \mathbf{x}(0) = \mathbf{x}_0. \end{cases} \quad (31)$$

Moreover, there exists a constant $c > 0$ independent of \mathbf{x}_0 and $\mathbf{g}(t)$ such that

$$\|\mathbf{x}\|_{\mathbb{W}(0, \tau)}^2 \leq c \left(\|\mathbf{g}\|_{L^2(0, \tau; \mathbb{X})}^2 + \|\mathbf{x}_0\|_{\mathbb{V}}^2 \right). \quad (32)$$

Since $\mathbb{W}(0, \tau)$ is embedded in $C(0, \tau; \mathbb{V})$, the state $\mathbf{x}(t)$ is bounded in \mathbb{V} for all $t \in [0, \tau]$. This together with Gâteaux differentiability of $\mathcal{F}(\cdot)$ ensures that there is a positive number $M_{\mathcal{F}}$ such that

$$\sup_{t \in [0, \tau]} \left\| \mathcal{F}'_{\mathbf{x}(t)} \right\|_{\mathcal{L}(\mathbb{V}, \mathbb{X})} \leq M_{\mathcal{F}}. \quad (33)$$

Thus, replacing the operator $\mathcal{P}(t)$ with $\mathcal{F}'_{\mathbf{x}(t)}$ and noting that

$$\|\mathcal{P}(t)\|_{\mathcal{L}(\mathbb{V}, \mathbb{X})} \leq M_{\mathcal{F}}, \quad (34)$$

shows that the conditions of Lemma 9 hold. Thus, there is a positive number c independent of \mathbf{g} such that

$$\|\mathbf{h}\|_{\mathbb{W}(0, \tau)} \leq c \|\mathbf{g}\|_{L^2(0, \tau; \mathbb{X})}. \quad (35)$$

Proposition 10. *Under assumptions A1-A5, the solution map $\mathcal{S}(\mathbf{u}(t), \mathbf{r}; \mathbf{x}_0)$ is Gâteaux differentiable with respect to each $\mathbf{u}(t)$ and \mathbf{r} in $U_{ad} \times K_{ad}$. Let $\mathbf{x}(t) = \mathcal{S}(\mathbf{u}(t), \mathbf{r}, \mathbf{x}_0)$.*

- a. *The Gâteaux derivative of $\mathcal{S}(\mathbf{u}(t), \mathbf{r}; \mathbf{x}_0)$ at \mathbf{r} in the direction $\tilde{\mathbf{r}}$ is the mapping $\mathcal{S}'_{\mathbf{r}} : \mathbb{K} \rightarrow L^2(0, \tau; D(\mathcal{A})) \cap W^{1,2}(0, \tau; \mathbb{X})$, $\tilde{\mathbf{r}} \mapsto \mathbf{z}(t)$, where $\mathbf{z}(t)$ is the strict solution to*

$$\begin{cases} \dot{\mathbf{z}}(t) = (\mathcal{A} + \mathcal{F}'_{\mathbf{x}(t)})\mathbf{z}(t) + (\mathcal{B}'_{\mathbf{r}}\tilde{\mathbf{r}})\mathbf{u}(t), \\ \mathbf{z}(0) = 0. \end{cases} \quad (36)$$

- b. *The Gâteaux derivative of $\mathcal{S}(\mathbf{u}(t), \mathbf{r}; \mathbf{x}_0)$ at $\mathbf{u}(t)$ in the direction $\tilde{\mathbf{u}}(t)$ is the mapping $\mathcal{S}'_{\mathbf{u}} : L^2(0, \tau; \mathbb{U}) \rightarrow L^2(0, \tau; D(\mathcal{A})) \cap W^{1,2}(0, \tau; \mathbb{X})$, $\tilde{\mathbf{u}}(t) \mapsto \mathbf{h}(t)$, where $\mathbf{h}(t)$ is the strict solution to*

$$\begin{cases} \dot{\mathbf{h}}(t) = (\mathcal{A} + \mathcal{F}'_{\mathbf{x}(t)})\mathbf{h}(t) + \mathcal{B}(\mathbf{r})\tilde{\mathbf{u}}(t), \\ \mathbf{h}(0) = 0. \end{cases} \quad (37)$$

Proof. a) Let ϵ be sufficiently small such that $\mathbf{r} + \epsilon\tilde{\mathbf{r}} \in K_{ad}$. Define $\mathbf{x}_{\epsilon}(t) = \mathcal{S}(\mathbf{u}(t), \mathbf{r} + \epsilon\tilde{\mathbf{r}}, \mathbf{x}_0)$, this state solves

$$\begin{cases} \dot{\mathbf{x}}_{\epsilon}(t) = \mathcal{A}\mathbf{x}_{\epsilon}(t) + \mathcal{F}(\mathbf{x}_{\epsilon}(t)) + \mathcal{B}(\mathbf{r} + \epsilon\tilde{\mathbf{r}})\mathbf{u}(t), & t > 0, \\ \mathbf{x}_{\epsilon}(0) = \mathbf{x}_0. \end{cases} \quad (38)$$

Similarly, $\mathbf{x}(t) = \mathcal{S}(\mathbf{u}(t), \mathbf{r}, \mathbf{x}_0)$ solves (53) with $\epsilon = 0$. Define $\mathbf{e}_{\mathcal{F}}(t)$ and $\mathbf{e}_{\mathcal{B}}$ as

$$\mathbf{e}_{\mathcal{F}}(t) := \frac{1}{\epsilon} \left(\mathcal{F}(\mathbf{x}(t)) - \mathcal{F}(\mathbf{x}_{\epsilon}(t)) - \mathcal{F}'_{\mathbf{x}(t)}(\mathbf{x}(t) - \mathbf{x}_{\epsilon}(t)) \right), \quad (39a)$$

$$\mathbf{e}_{\mathcal{B}} := \frac{1}{\epsilon} (\mathcal{B}(\mathbf{r} + \epsilon\tilde{\mathbf{r}}) - \mathcal{B}(\mathbf{r})) - \mathcal{B}'_{\mathbf{r}}\tilde{\mathbf{r}}. \quad (39b)$$

The state $\mathbf{e}(t) = (\mathbf{x}(t) - \mathbf{x}_{\epsilon}(t))/\epsilon - \mathbf{z}(t)$ satisfies

$$\begin{cases} \dot{\mathbf{e}}(t) = (\mathcal{A} + \mathcal{F}'_{\mathbf{x}(t)})\mathbf{e}(t) + \mathbf{e}_{\mathcal{F}}(t) + \mathbf{e}_{\mathcal{B}}\mathbf{u}(t), & t > 0, \\ \mathbf{e}(0) = 0. \end{cases} \quad (40)$$

Assumption A4 and A5 ensure that as $\epsilon \rightarrow 0$

$$\|\mathbf{e}_{\mathcal{F}}(t)\| \rightarrow 0, \quad \forall t \in [0, \tau], \quad (41a)$$

$$\|\mathbf{e}_{\mathcal{B}}\|_{\mathcal{L}(\mathbb{U}, \mathbb{X})} \rightarrow 0. \quad (41b)$$

It will be shown that $\lim_{\epsilon \rightarrow 0} \|\mathbf{e}\|_{\mathbb{W}(0, \tau)} = 0$. First, consider $\mathbf{x}(t) - \mathbf{x}_{\epsilon}(t)$, which satisfies

$$\begin{cases} \dot{\mathbf{x}}(t) - \dot{\mathbf{x}}_{\epsilon}(t) = \mathcal{A}(\mathbf{x}(t) - \mathbf{x}_{\epsilon}(t)) + \mathcal{F}(\mathbf{x}(t)) - \mathcal{F}(\mathbf{x}_{\epsilon}(t)) \\ \quad + (\mathcal{B}(\mathbf{r}) - \mathcal{B}(\mathbf{r} + \epsilon\tilde{\mathbf{r}}))\mathbf{u}(t), \\ \mathbf{x}(0) - \mathbf{x}_{\epsilon}(0) = 0. \end{cases}$$

Lemma 3 implies that there is a number c_{τ} depending only on τ such that for all $t \in [0, \tau]$

$$\|\mathbf{x}(t) - \mathbf{x}_{\epsilon}(t)\|_{\mathbb{V}} \quad (42)$$

$$\leq c_{\tau} \left(\|\dot{\mathbf{x}} - \dot{\mathbf{x}}_{\epsilon}\|_{L^2(0, t; \mathbb{X})} + \|\mathcal{A}(\mathbf{x} - \mathbf{x}_{\epsilon})\|_{L^2(0, t; \mathbb{X})} \right). \quad (43)$$

Also, use [55, Proposition 2.2], there is a number d_{τ} depending only on τ such that for all $t \in [0, \tau]$

$$\|\dot{\mathbf{x}} - \dot{\mathbf{x}}_{\epsilon}\|_{L^2(0, t; \mathbb{X})} + \|\mathcal{A}(\mathbf{x} - \mathbf{x}_{\epsilon})\|_{L^2(0, t; \mathbb{X})} \quad (44)$$

$$\leq d_{\tau} \|\dot{\mathbf{x}} - \dot{\mathbf{x}}_{\epsilon} - \mathcal{A}(\mathbf{x} - \mathbf{x}_{\epsilon})\|_{L^2(0, t; \mathbb{X})}.$$

Combine (43) and (44) to obtain

$$\|\mathbf{x}(t) - \mathbf{x}_{\epsilon}(t)\|_{\mathbb{V}} \leq c_{\tau} d_{\tau} \|\mathcal{F}(\mathbf{x}) - \mathcal{F}(\mathbf{x}_{\epsilon})\|_{L^2(0, t; \mathbb{X})} \quad (45)$$

$$+ c_{\tau} d_{\tau} \|(\mathcal{B}(\mathbf{r}) - \mathcal{B}(\mathbf{r} + \epsilon\tilde{\mathbf{r}}))\mathbf{u}\|_{L^2(0, t; \mathbb{X})}.$$

Theorem 4 implies that the states $\mathbf{x}(t)$ and $\mathbf{x}_{\epsilon}(t)$ belong to some bounded set in $\mathbb{W}(0, \tau)$ and so in $C(0, \tau, \mathbb{V})$. Let $D \subset \mathbb{V}$ be a bounded set that contains the trajectories $\mathbf{x}(t)$ and $\mathbf{x}_{\epsilon}(t)$. Let $L_{\mathcal{F}}$ be the Lipschitz constant of $\mathcal{F}(\cdot)$ on D . Since the set K_{ad} is compact and $\mathcal{B}(\mathbf{r})$ satisfies assumption A5, the number $L_{\mathcal{B}}$ defined as

$$L_{\mathcal{B}} = \sup_{\mathbf{r} \in K_{ad}} \|\mathcal{B}'_{\mathbf{r}}\|_{\mathcal{L}(\mathbb{K}, \mathcal{L}(\mathbb{U}, \mathbb{X}))}. \quad (46)$$

is finite. This together with [60, Theorem 12.1.1 and Corollary 3] yields

$$\|\mathcal{B}(\mathbf{r}) - \mathcal{B}(\mathbf{r}_{\epsilon})\|_{\mathcal{L}(\mathbb{U}, \mathbb{X})} \leq L_{\mathcal{B}} \|\mathbf{r} - \mathbf{r}_{\epsilon}\|_{\mathbb{K}} \leq L_{\mathcal{B}}\epsilon. \quad (47)$$

Use these to obtain the inequality

$$\begin{aligned} \|\mathbf{x}(t) - \mathbf{x}_{\epsilon}(t)\|_{\mathbb{V}}^2 &\leq 2c_{\tau}^2 d_{\tau}^2 L_{\mathcal{F}}^2 \int_0^t \|\mathbf{x}(s) - \mathbf{x}_{\epsilon}(s)\|_{\mathbb{V}}^2 ds \\ &\quad + 2c_{\tau}^2 d_{\tau}^2 L_{\mathcal{B}}^2 \epsilon^2 \|\mathbf{u}\|_{L^2(0, \tau; \mathbb{U})}^2 \end{aligned} \quad (48)$$

Applying Gronwall's lemma yields

$$\|\mathbf{x}(t) - \mathbf{x}_{\epsilon}(t)\|_{\mathbb{V}} \leq \sqrt{2} e^{c_{\tau}^2 d_{\tau}^2 L_{\mathcal{F}}^2} c_{\tau} d_{\tau} L_{\mathcal{B}} \epsilon \|\mathbf{u}\|_{L^2(0, \tau; \mathbb{U})}. \quad (49)$$

Define

$$M_{\mathcal{F}} := \sup_{t \in [0, \tau]} \|\mathcal{F}'_{\mathbf{x}(t)}\|_{\mathcal{L}(\mathbb{V}, \mathbb{X})}.$$

Assumption A4 ensures that $M_{\mathcal{F}}$ is finite. Take the norm of the right side of (39a) in \mathbb{X} . It follows that

$$\|\mathbf{e}_{\mathcal{F}}(t)\| \leq (L_{\mathcal{F}} + M_{\mathcal{F}} c_{\epsilon}) \sqrt{2} e^{c_{\tau}^2 d_{\tau}^2 L_{\mathcal{F}}^2} c_{\tau} d_{\tau} L_{\mathcal{B}} \|\mathbf{u}\|_{L^2(0, \tau; \mathbb{U})}. \quad (50)$$

This and (41a) together with the Bounded Convergence Theorem ensure that

$$\lim_{\epsilon \rightarrow 0} \int_0^{\tau} \|\mathbf{e}_{\mathcal{F}}(t)\|^2 dt = 0. \quad (51)$$

Statements (51) and (41b), and Lemma 9 can be applied to conclude

$$\lim_{\epsilon \rightarrow 0} \|\mathbf{e}\|_{\mathbb{W}(0, \tau)} = 0. \quad (52)$$

This shows that $\mathcal{S}(\mathbf{u}, \mathbf{r}, \mathbf{x}_0)$ is Gâteaux differentiable at \mathbf{r} in the direction $\tilde{\mathbf{r}}$ with derivative $z(t) = \mathcal{S}'_{\mathbf{r}} \tilde{\mathbf{r}}$.

b) This part is proven in [30, Theorem 3.4] assuming that $\partial_t + \mathcal{A}$ is invertible. However, the result is still true without assuming the invertibility of $\partial_t + \mathcal{A}$. Let ϵ be sufficiently small such that $\mathbf{u} + \epsilon \tilde{\mathbf{u}} \in U_{ad}$. Define $\mathbf{x}_\epsilon(t) = \mathcal{S}(\mathbf{u}(t) + \epsilon \tilde{\mathbf{u}}(t), \mathbf{r}, \mathbf{x}_0)$, this state solves

$$\begin{cases} \dot{\mathbf{x}}_\epsilon(t) = \mathcal{A}\mathbf{x}_\epsilon(t) + \mathcal{F}(\mathbf{x}_\epsilon(t)) + \mathcal{B}(\mathbf{r})(\mathbf{u}(t) + \epsilon \tilde{\mathbf{u}}(t)), & t > 0, \\ \mathbf{x}_\epsilon(0) = \mathbf{x}_0. \end{cases} \quad (53)$$

Let $e(t) = (\mathbf{x}(t) - \mathbf{x}_\epsilon(t))/\epsilon - \mathbf{h}(t)$. Following the same steps as in part (a) yields

$$\lim_{\epsilon \rightarrow 0} \|e\|_{\mathbb{W}(0, \tau)} = 0. \quad (54)$$

This means that $\mathcal{S}(\mathbf{u}, \mathbf{r}, \mathbf{x}_0)$ is Gâteaux differentiable at \mathbf{u} in the direction $\tilde{\mathbf{u}}$ with derivative $\mathbf{h}(t) = (\mathcal{S}'_{\mathbf{u}} \tilde{\mathbf{u}})(t)$. \square

Assumption A1 implies that the dual of each of \mathbb{X} and \mathbb{U} will be identified with the space itself. For each \mathbf{u} , the operator $(\mathcal{B}'_{\mathbf{r}^o} \mathbf{u})^* : \mathbb{X} \rightarrow \mathbb{K}^*$ is defined by

$$\langle (\mathcal{B}'_{\mathbf{r}^o} \mathbf{u})^* \mathbf{p}, \mathbf{r} \rangle_{\mathbb{K}^*, \mathbb{K}} = \langle \mathbf{p}, (\mathcal{B}'_{\mathbf{r}^o} \mathbf{r}) \mathbf{u} \rangle, \quad \forall (\mathbf{u}, \mathbf{p}, \mathbf{r}) \in \mathbb{U} \times \mathbb{X} \times \mathbb{K}.$$

Theorem 11. *Suppose assumptions A1-A5 hold, and writing the derivatives $J'_{\mathbf{x}}$, $J'_{\mathbf{u}}$, and $J'_{\mathbf{r}}$ by elements $\mathbf{j}_{\mathbf{x}} \in \mathbb{W}(0, \tau)^*$, $\mathbf{j}_{\mathbf{u}} \in L^2(0, \tau; \mathbb{U})$ and $\mathbf{j}_{\mathbf{r}} \in \mathbb{K}^*$, respectively. For any initial condition $\mathbf{x}_0 \in \mathbb{X}$, let the pair $(\mathbf{u}^o, \mathbf{r}^o) \in U_{ad} \times K_{ad}$ be a local minimizer of the optimization problem (P) with the optimal trajectory $\mathbf{x}^o = \mathcal{S}(\mathbf{u}^o; \mathbf{r}^o, \mathbf{x}_0)$ and let $\mathbf{p}^o(t)$ indicate the strict solution in $\mathbb{W}(0, \tau)^*$ of the final value problem*

$$\dot{\mathbf{p}}^o(t) = -(\mathcal{A}^* + \mathcal{F}'_{\mathbf{x}^o(t)})^* \mathbf{p}^o(t) - \mathbf{j}_{\mathbf{x}^o}(t), \quad \mathbf{p}^o(\tau) = 0. \quad (55)$$

Then $(\mathbf{u}^o, \mathbf{r}^o)$ satisfy

$$\begin{aligned} \langle \mathbf{j}_{\mathbf{u}^o} + \mathcal{B}^*(\mathbf{r}^o) \mathbf{p}^o, \mathbf{u} - \mathbf{u}^o \rangle_{L^2(0, \tau; \mathbb{U})} &\geq 0, \\ \left\langle \mathbf{j}_{\mathbf{r}^o} + \int_0^\tau (\mathcal{B}'_{\mathbf{r}^o} \mathbf{u}^o(t))^* \mathbf{p}^o(t) dt, \mathbf{r} - \mathbf{r}^o \right\rangle_{\mathbb{K}^*, \mathbb{K}} &\geq 0. \end{aligned}$$

for all $\mathbf{u} \in U_{ad}$ and $\mathbf{r} \in K_{ad}$.

Proof. Let

$$\mathcal{G}(\mathbf{u}, \mathbf{r}) = J(\mathcal{S}(\mathbf{u}, \mathbf{r}, \mathbf{x}_0), \mathbf{u}, \mathbf{r}).$$

The Gâteaux derivative of $\mathcal{G}(\mathbf{u}, \mathbf{r})$ with respect to \mathbf{u} has been obtained in the proof of [30, Proposition 4.13]. Using the chain rule to take the Gâteaux derivative of $\mathcal{G}(\mathbf{u}, \mathbf{r})$ at \mathbf{u}^o in the direction $\tilde{\mathbf{u}}$ yields

$$\mathcal{G}'_{\mathbf{u}^o} \tilde{\mathbf{u}} = J'_{\mathbf{u}^o} \tilde{\mathbf{u}} + J'_{\mathbf{x}^o} \mathcal{S}'_{\mathbf{u}^o} \tilde{\mathbf{u}}. \quad (57)$$

Identify the functionals $\mathcal{G}'_{\mathbf{u}^o} : L^2(0, \tau; \mathbb{U}) \rightarrow \mathbb{R}$ and $J'_{\mathbf{u}^o} : L^2(0, \tau; \mathbb{U}) \rightarrow \mathbb{R}$ with elements of $L^2(0, \tau; \mathbb{U})$. That is

$$\mathcal{G}'_{\mathbf{u}^o} \tilde{\mathbf{u}} = \langle \mathbf{g}_{\mathbf{u}^o}, \tilde{\mathbf{u}} \rangle_{L^2(0, \tau; \mathbb{U})}, \quad (58)$$

$$J'_{\mathbf{u}^o} \tilde{\mathbf{u}} = \langle \mathbf{j}_{\mathbf{u}^o}, \tilde{\mathbf{u}} \rangle_{L^2(0, \tau; \mathbb{U})}. \quad (59)$$

Also, identifying the functional $J'_{\mathbf{x}^o} : L^2(0, \tau; \mathbb{X}) \rightarrow \mathbb{R}$ with an element of $\mathbb{W}(0, \tau)^* = L^2(0, \tau; D(\mathcal{A}^*)) \cap W^{1,2}(0, \tau; \mathbb{X})$ yields

$$J'_{\mathbf{x}^o} \mathcal{S}'_{\mathbf{u}^o} \tilde{\mathbf{u}} = \langle \mathbf{j}_{\mathbf{x}^o}, \mathcal{S}'_{\mathbf{u}^o} \tilde{\mathbf{u}} \rangle_{L^2(0, \tau; \mathbb{X})}. \quad (60)$$

The adjoint operator $\mathcal{S}'_{\mathbf{u}^o}$ can be obtained as follows. Use (55) in the following inner product and let $\mathbf{h}(t) = \mathcal{S}'_{\mathbf{u}^o} \tilde{\mathbf{u}}$

$$\begin{aligned} \langle \mathbf{j}_{\mathbf{x}^o}, \mathcal{S}'_{\mathbf{u}^o} \tilde{\mathbf{u}} \rangle_{L^2(0, \tau; \mathbb{X})} \\ = \int_0^\tau \left\langle -\dot{\mathbf{p}}^o(t) - (\mathcal{A}^* + \mathcal{F}'_{\mathbf{x}^o(t)})^* \mathbf{p}^o(t), \mathbf{h}(t) \right\rangle dt. \end{aligned}$$

Taking the adjoint and integration by parts yield

$$\begin{aligned} \langle \mathbf{j}_{\mathbf{x}^o}, \mathcal{S}'_{\mathbf{u}^o} \tilde{\mathbf{u}} \rangle_{L^2(0, \tau; \mathbb{X})} \\ = \int_0^\tau \left\langle \mathbf{p}^o(t), \dot{\mathbf{h}}(t) - (\mathcal{A} + \mathcal{F}'_{\mathbf{x}^o(t)}) \mathbf{h}(t) \right\rangle dt \\ = \int_0^\tau \langle \mathbf{p}^o(t), \mathcal{B}(\mathbf{r}) \tilde{\mathbf{u}}(t) \rangle dt \\ = \int_0^\tau \langle \mathcal{B}^*(\mathbf{r}) \mathbf{p}^o(t), \tilde{\mathbf{u}}(t) \rangle_{\mathbb{U}} dt. \end{aligned}$$

This implies

$$\mathcal{S}'_{\mathbf{u}^o} \mathbf{j}_{\mathbf{x}^o} = \mathcal{B}^*(\mathbf{r}) \mathbf{p}^o(t). \quad (61)$$

Combine (58), (59), (60) and use (61), equation (57) is written using the functionals as

$$\langle \mathbf{g}_{\mathbf{u}}, \tilde{\mathbf{u}} \rangle_{L^2(0, \tau; \mathbb{U})} = \langle \mathbf{j}_{\mathbf{u}^o} + \mathcal{B}^*(\mathbf{r}^o) \mathbf{p}^o, \tilde{\mathbf{u}} \rangle_{L^2(0, \tau; \mathbb{U})}. \quad (62)$$

Applying [24, Theorem 1.46] and letting $\tilde{\mathbf{u}} = \mathbf{u} - \mathbf{u}^o$ for all $\mathbf{u} \in U_{ad}$ yields

$$\langle \mathbf{j}_{\mathbf{u}^o} + \mathcal{B}^*(\mathbf{r}^o) \mathbf{p}^o, \mathbf{u} - \mathbf{u}^o \rangle_{L^2(0, \tau; \mathbb{U})} \geq 0. \quad (63)$$

Using the chain rule to take the Gâteaux derivative of $\mathcal{G}(\mathbf{u}, \mathbf{r})$ at \mathbf{r}^o in the direction $\tilde{\mathbf{r}}$ yields

$$\mathcal{G}'_{\mathbf{r}^o} \tilde{\mathbf{r}} = J'_{\mathbf{r}^o} \tilde{\mathbf{r}} + J'_{\mathbf{x}^o} \mathcal{S}'_{\mathbf{r}^o} \tilde{\mathbf{r}}. \quad (64)$$

Write the functionals $\mathcal{G}'_{\mathbf{r}^o} : \mathbb{K} \rightarrow \mathbb{R}$ and $J'_{\mathbf{r}^o} : \mathbb{K} \rightarrow \mathbb{R}$ as elements of \mathbb{K}^* and $\mathbf{j}_{\mathbf{r}^o}$ in \mathbb{K}^* , respectively, and take the adjoint of $\mathcal{S}'_{\mathbf{r}^o}$. It follows that

$$\mathbf{g}_{\mathbf{r}^o} = \mathcal{S}'_{\mathbf{r}^o} \mathbf{j}_{\mathbf{x}^o}(t) + \mathbf{j}_{\mathbf{r}^o}. \quad (65)$$

An explicit representation of the adjoint operator $\mathcal{S}'_{\mathbf{r}^o}$ will be derived. Consider the inner product

$$\langle \mathbf{j}_{\mathbf{x}^o}, \mathcal{S}'_{\mathbf{r}^o} \tilde{\mathbf{r}} \rangle_{L^2(0, \tau; \mathbb{X})} = \int_0^\tau \langle \mathbf{j}_{\mathbf{x}^o}(t), \mathcal{S}'_{\mathbf{r}^o} \tilde{\mathbf{r}} \rangle dt.$$

Write $z(t) = \mathcal{S}'_{\mathbf{r}^o} \tilde{\mathbf{r}}$. Substitute for $\mathbf{j}_{\mathbf{x}^o}(t)$ from (55) into this integral. Perform integration by parts to obtain

$$\begin{aligned} \int_0^\tau \left\langle -\dot{\mathbf{p}}^o(t) - (\mathcal{A}^* + \mathcal{F}'_{\mathbf{x}^o(t)})^* \mathbf{p}^o(t), z(t) \right\rangle dt \\ = \int_0^\tau \left\langle \mathbf{p}^o(t), \dot{z}(t) - (\mathcal{A} + \mathcal{F}'_{\mathbf{x}^o(t)}) z(t) \right\rangle dt \\ = \int_0^\tau \langle \mathbf{p}^o(t), (\mathcal{B}'_{\mathbf{r}^o} \tilde{\mathbf{r}}) \mathbf{u}^o(t) \rangle dt \\ = \left\langle \int_0^\tau (\mathcal{B}'_{\mathbf{r}^o} \mathbf{u}^o(t))^* \mathbf{p}^o(t) dt, \tilde{\mathbf{r}} \right\rangle_{\mathbb{K}^*, \mathbb{K}}. \end{aligned} \quad (66)$$

Thus,

$$\mathcal{S}'_{\mathbf{r}^o} \mathbf{j}_{\mathbf{x}^o}(t) = \int_0^\tau (\mathcal{B}'_{\mathbf{r}^o} \mathbf{u}^o(t))^* \mathbf{p}^o(t) dt. \quad (67)$$

As a result, the Gâteaux derivative of $\mathcal{G}(\mathbf{u}, \mathbf{r})$ at \mathbf{r}° in the direction $\tilde{\mathbf{r}}$ is

$$\mathbf{g}'_{\mathbf{r}^\circ} = \int_0^\tau (\mathcal{B}'_{\mathbf{r}^\circ} \mathbf{u}^\circ(t))^* \mathbf{p}^\circ(t) dt + \mathbf{j}_{\mathbf{r}^\circ}. \quad (68)$$

The optimality conditions now follow by substituting the Gâteaux derivatives $\mathbf{g}'_{\mathbf{r}^\circ}$ in [24, Theorem 1.46]. \square

Corollary 12. *Let the cost $J(\mathbf{x}, \mathbf{u}, \mathbf{r})$ be*

$$J(\mathbf{x}, \mathbf{u}, \mathbf{r}) = \int_0^\tau \langle \mathcal{Q}\mathbf{x}(t), \mathbf{x}(t) \rangle + \langle \mathcal{R}\mathbf{u}(t), \mathbf{u}(t) \rangle_{\mathbb{U}} dt, \quad (69)$$

where \mathcal{Q} is a positive semi-definite, self-adjoint bounded linear operator on \mathbb{X} , and \mathcal{R} is a coercive, self-adjoint linear bounded operator on \mathbb{U} . If the minimizer $(\mathbf{u}^\circ, \mathbf{r}^\circ)$ is in the interior of $U_{ad} \times K_{ad}$, then the following set of equations characterizes $(\mathbf{x}^\circ, \mathbf{p}^\circ, \mathbf{u}^\circ, \mathbf{r}^\circ)$:

$$\begin{cases} \dot{\mathbf{x}}^\circ(t) = \mathcal{A}\mathbf{x}^\circ(t) + \mathcal{F}(\mathbf{x}^\circ(t)) + \mathcal{B}(\mathbf{r}^\circ)\mathbf{u}^\circ(t), & \mathbf{x}^\circ(0) = \mathbf{x}_0, \\ \dot{\mathbf{p}}^\circ(t) = -(\mathcal{A}^* + \mathcal{F}'_{\mathbf{x}^\circ(t)})^* \mathbf{p}^\circ(t) - \mathcal{Q}\mathbf{x}^\circ(t), & \mathbf{p}^\circ(\tau) = 0, \\ \mathbf{u}^\circ(t) = -\mathcal{R}^{-1}\mathcal{B}^*(\mathbf{r}^\circ)\mathbf{p}^\circ(t), \\ \int_0^\tau (\mathcal{B}'_{\mathbf{r}^\circ} \mathbf{u}^\circ(t))^* \mathbf{p}^\circ(t) dt = 0. \end{cases}$$

Proof. If the optimizer $(\mathbf{u}^\circ, \mathbf{r}^\circ)$ is in the interior of $U_{ad} \times K_{ad}$, then the optimality conditions of Theorem 11 hold if and only if

$$\mathbf{j}_{\mathbf{u}^\circ} + \mathcal{B}^*(\mathbf{r}^\circ)\mathbf{p}^\circ = 0, \quad (70)$$

$$\mathbf{j}_{\mathbf{r}^\circ} + \int_0^\tau (\mathcal{B}'_{\mathbf{r}^\circ} \mathbf{u}^\circ(t))^* \mathbf{p}^\circ(t) dt = 0. \quad (71)$$

The derivatives $J'_{\mathbf{x}^\circ}(t) : L^2(0, \tau; \mathbb{X}) \rightarrow \mathbb{R}$ and $J'_{\mathbf{u}^\circ}(t) : L^2(0, \tau; \mathbb{U}) \rightarrow \mathbb{R}$ are

$$J'_{\mathbf{x}^\circ} \tilde{\mathbf{x}} = \langle \mathcal{Q}\mathbf{x}^\circ, \tilde{\mathbf{x}} \rangle_{L^2(0, \tau; \mathbb{X})}, \quad (72)$$

$$J'_{\mathbf{u}^\circ} \tilde{\mathbf{u}} = \langle \mathcal{R}\mathbf{u}^\circ, \tilde{\mathbf{u}} \rangle_{L^2(0, \tau; \mathbb{U})}. \quad (73)$$

Identify these functionals with elements $\mathbf{j}_{\mathbf{x}^\circ} = \mathcal{Q}\mathbf{x}^\circ(t)$ and $\mathbf{j}_{\mathbf{u}^\circ} = \mathcal{R}\mathbf{u}^\circ(t)$, and notice that $\mathbf{j}_{\mathbf{r}^\circ} = 0$. Substituting the derivatives in (70) and (71) yields the optimality conditions. \square

For all \mathbf{x}_1 and \mathbf{x}_2 in $D(\mathcal{A})$ and $t \in (0, \tau)$, let $\Pi(t)$ be the solution to the differential Riccati equation

$$\begin{cases} \frac{d}{dt} \langle \mathbf{x}_2, \Pi(t)\mathbf{x}_1 \rangle = -\langle \mathbf{x}_2, \Pi(t)\mathcal{A}\mathbf{x}_1 \rangle - \langle \mathcal{A}\mathbf{x}_2, \Pi(t)\mathbf{x}_1 \rangle \\ \quad - \langle \mathcal{Q}\mathbf{x}_2, \mathbf{x}_1 \rangle + \langle \Pi(t)\mathcal{B}(\mathbf{r})\mathcal{R}^{-1}\mathcal{B}^*(\mathbf{r})\Pi(t)\mathbf{x}_2, \mathbf{x}_1 \rangle, \\ \Pi(\tau) = 0. \end{cases} \quad (74)$$

It is well-known, [61, Chapter 6] and [62, Chapter 1], that if the system is linear then the adjoint trajectory state $\mathbf{p}^\circ(t)$ satisfies

$$\mathbf{p}^\circ(t) = \Pi(t)\mathbf{x}^\circ(t). \quad (75)$$

As a result, the optimal input and actuator design satisfy in this case

$$\begin{cases} \mathbf{u}^\circ(t) = -\mathcal{R}^{-1}\mathcal{B}^*(\mathbf{r}^\circ)\Pi(t)\mathbf{x}^\circ(t), \\ \int_0^\tau (\mathcal{B}'_{\mathbf{r}^\circ} \mathbf{u}^\circ(t))^* \Pi(t)\mathbf{x}^\circ(t) dt = 0. \end{cases}$$

V. WORST INITIAL CONDITION

In this section, sets U_{ad} and K_{ad} and numbers τ and R_2 are the same sets and numbers as in the previous section.

The worst initial condition maximizes $J(\mathbf{x}, \mathbf{u}, \mathbf{r})$ over all choices of initial conditions in $B_{\mathbb{V}}(R_2)$ subject to IVP (5) for a fixed input $\mathbf{u} \in U_{ad}$ and fixed actuator design $\mathbf{r} \in K_{ad}$. Formally, define $\mathcal{G}(\cdot) : \mathbb{V} \rightarrow \mathbb{R}$ as

$$\mathcal{G}(\mathbf{x}_0) = J(\mathcal{S}(\mathbf{u}, \mathbf{r}; \mathbf{x}_0), \mathbf{u}, \mathbf{r}),$$

the worst initial condition over $B_{\mathbb{V}}(R_2)$ is the solution to

$$\begin{cases} \max & \mathcal{G}(\mathbf{x}_0) \\ \text{s.t.} & \mathbf{x}_0 \in B_{\mathbb{V}}(R_2). \end{cases} \quad (\text{P1})$$

Lemma 13. *For every $\mathbf{u} \in U_{ad}$ and $\mathbf{r} \in K_{ad}$, the optimization problem (P1) admits a maximizer.*

Proof. As in the proof of Theorem 7, define

$$j := \sup_{\mathbf{x}_0 \in B_{\mathbb{V}}(R_2)} \mathcal{G}(\mathbf{x}_0). \quad (76)$$

Extract a maximizing sequence \mathbf{x}_0^n in $B_{\mathbb{V}}(R_2)$. The set $B_{\mathbb{V}}(R_2)$ is closed and convex in the reflexive Banach space \mathbb{V} , it is therefore weakly closed. This implies that \mathbf{x}_0^n has a subsequence that converges weakly to some element $\tilde{\mathbf{x}}_0$ in $B_{\mathbb{V}}(R_2)$. Also, according to Theorem 4 and Theorem 6, the solution map is bounded and weakly continuous in \mathbf{x}_0 . The cost function is also convex and continuous in \mathbf{x}_0 , so it is weakly lower semi-continuous in \mathbf{x}_0 . These imply that $\tilde{\mathbf{x}}_0$ solves (P1). \square

Proposition 14. *Under assumptions A1-A4, the solution map $\mathcal{S}(\mathbf{u}(t), \mathbf{r}; \mathbf{x}_0)$ is Gâteaux differentiable with respect to $\mathbf{x}_0 \in B_{\mathbb{V}}(R_2)$. Let $\mathbf{x}(t) = \mathcal{S}(\mathbf{u}(t), \mathbf{r}, \mathbf{x}_0)$, the Gâteaux derivative of $\mathcal{S}(\mathbf{u}(t), \mathbf{r}; \mathbf{x}_0)$ at \mathbf{x}_0 in the interior of $B_{\mathbb{V}}(R_2)$ in the direction $\tilde{\mathbf{x}}_0$ is the mapping $\mathcal{S}'_{\mathbf{x}}(\mathbf{u}(t), \mathbf{r}; \cdot) : \mathbb{V} \rightarrow \mathbb{W}(0, \tau)$, $\tilde{\mathbf{x}}_0 \mapsto \mathbf{q}(t)$, where $\mathbf{q}(t)$ is the strict solution to*

$$\begin{cases} \dot{\mathbf{q}}(t) = (\mathcal{A} + \mathcal{F}'_{\mathbf{x}(t)})\mathbf{q}(t), \\ \mathbf{q}(0) = \tilde{\mathbf{x}}_0. \end{cases} \quad (77)$$

Proof. Let the number $\epsilon > 0$ be small enough such that $\mathbf{x}_0 + \epsilon\tilde{\mathbf{x}}_0 \in B_{\mathbb{V}}(R_2)$. Define $\mathbf{x}_\epsilon(t) := \mathcal{S}(\mathbf{u}(t), \mathbf{r}, \mathbf{x}_0 + \epsilon\tilde{\mathbf{x}}_0)$, it solves

$$\begin{cases} \dot{\mathbf{x}}_\epsilon(t) = \mathcal{A}\mathbf{x}_\epsilon(t) + \mathcal{F}(\mathbf{x}_\epsilon(t)) + \mathcal{B}(\mathbf{r})\mathbf{u}(t), & t > 0, \\ \mathbf{x}_\epsilon(0) = \mathbf{x}_0 + \epsilon\tilde{\mathbf{x}}_0. \end{cases} \quad (78)$$

Define $\mathbf{e}_{\mathcal{F}}(t)$ as

$$\mathbf{e}_{\mathcal{F}}(t) := \frac{1}{\epsilon} \left(\mathcal{F}(\mathbf{x}(t)) - \mathcal{F}(\mathbf{x}_\epsilon(t)) - \mathcal{F}'_{\mathbf{x}(t)}(\mathbf{x}(t) - \mathbf{x}_\epsilon(t)) \right).$$

Let $\mathbf{e}(t) = (\mathbf{x}(t) - \mathbf{x}_\epsilon(t))/\epsilon - \mathbf{q}(t)$, it satisfies

$$\begin{cases} \dot{\mathbf{e}}(t) = (\mathcal{A} + \mathcal{F}'_{\mathbf{x}(t)})\mathbf{e}(t) + \mathbf{e}_{\mathcal{F}}(t), \\ \mathbf{e}(0) = 0. \end{cases} \quad (79)$$

Assumption A4 ensures that as $\epsilon \rightarrow 0$

$$\|\mathbf{e}_{\mathcal{F}}(t)\| \rightarrow 0, \quad \forall t \in [0, \tau]. \quad (80)$$

The convergence in (80) is uniform; to show this, note that $\mathbf{x}(t) - \mathbf{x}_\epsilon(t)$ satisfies

$$\begin{cases} \dot{\mathbf{x}}(t) - \dot{\mathbf{x}}_\epsilon(t) = \mathcal{A}(\mathbf{x}(t) - \mathbf{x}_\epsilon(t)) + \mathcal{F}(\mathbf{x}(t)) - \mathcal{F}(\mathbf{x}_\epsilon(t)), \\ \mathbf{x}(0) - \mathbf{x}_\epsilon(0) = \epsilon \tilde{\mathbf{x}}_0. \end{cases}$$

According to [55, Proposition 2.2], there is d_τ depending only on τ such that for all $t \in [0, \tau]$

$$\begin{aligned} \|\dot{\mathbf{x}} - \dot{\mathbf{x}}_\epsilon\|_{L^2(0,t;\mathbb{X})} + \|\mathcal{A}(\mathbf{x} - \mathbf{x}_\epsilon)\|_{L^2(0,t;\mathbb{X})} \\ \leq d_\tau \left(\|\mathcal{F}(\mathbf{x}) - \mathcal{F}(\mathbf{x}_\epsilon)\|_{L^2(0,t;\mathbb{X})} + \epsilon \|\tilde{\mathbf{x}}_0\|_{\mathbb{V}} \right) \end{aligned} \quad (81)$$

Also, letting c_τ be the embedding constant of $\mathbb{W}(0, \tau) \hookrightarrow C(0, \tau; \mathbb{V})$, $\mathbf{x} - \mathbf{x}_\epsilon$ satisfies

$$\begin{aligned} \|\mathbf{x} - \mathbf{x}_\epsilon\|_{C(0,t;\mathbb{V})} \leq c_\tau \left(\|\dot{\mathbf{x}} - \dot{\mathbf{x}}_\epsilon\|_{L^2(0,t;\mathbb{X})} \right. \\ \left. + \|\mathcal{A}(\mathbf{x} - \mathbf{x}_\epsilon)\|_{L^2(0,t;\mathbb{X})} \right). \end{aligned} \quad (82)$$

Theorem 4 implies that the states $\mathbf{x}(t)$ and $\mathbf{x}_\epsilon(t)$ belong to some bounded set D ; so let $L_{\mathcal{F}}$ be the Lipschitz constant $\mathcal{F}(\cdot)$ on D . Combining this with inequalities (81) and (82) yield

$$\begin{aligned} \|\mathbf{x}(t) - \mathbf{x}_\epsilon(t)\|_{\mathbb{V}}^2 \leq 2c_\tau^2 d_\tau^2 L_{\mathcal{F}}^2 \int_0^t \|\mathbf{x}(s) - \mathbf{x}_\epsilon(s)\|_{\mathbb{V}}^2 ds \\ + 2\epsilon^2 \|\tilde{\mathbf{x}}_0\|_{\mathbb{V}}^2. \end{aligned} \quad (83)$$

Applying Gronwall's lemma to this inequality yields

$$\|\mathbf{x}(t) - \mathbf{x}_\epsilon(t)\|_{\mathbb{V}} \leq \sqrt{2} e^{c_\tau^2 d_\tau^2 L_{\mathcal{F}}^2 t} \epsilon \|\tilde{\mathbf{x}}_0\|_{\mathbb{V}}. \quad (84)$$

Take the norm of $e_{\mathcal{F}}(t)$ in \mathbb{X} , use (84), define

$$M_{\mathcal{F}} := \sup\{\|\mathcal{F}'_{\mathbf{x}(t)}\|_{\mathcal{L}(\mathbb{V}, \mathbb{X})} : t \in [0, \tau]\}.$$

It follows that

$$\|e_{\mathcal{F}}(t)\| \leq (L_{\mathcal{F}} + M_{\mathcal{F}}) \sqrt{2} e^{c_\tau^2 d_\tau^2 L_{\mathcal{F}}^2 t} \|\tilde{\mathbf{x}}_0\|_{\mathbb{V}} < \infty.$$

The Bounded Convergence Theorem now ensures that

$$\lim_{\epsilon \rightarrow 0} \int_0^\tau \|e_{\mathcal{F}}(t)\|_{\mathbb{V}}^2 dt = 0. \quad (85)$$

Lemma 9 together with (85) gives

$$\lim_{\epsilon \rightarrow 0} \|e\|_{\mathbb{W}(0,\tau)} = 0. \quad (86)$$

This shows that $\mathcal{S}(\mathbf{u}, \mathbf{r}, \mathbf{x}_0)$ is Gâteaux differentiable at \mathbf{x}_0 in the direction $\tilde{\mathbf{x}}_0$. \square

Theorem 15. *Suppose assumptions A1-A4 hold, and identify the derivative $J'_{\mathbf{x}}$ by element $\mathbf{j}_{\mathbf{x}} \in \mathbb{W}(0, \tau)^*$. Let $\mathbf{u} \in U_{ad}$, $\mathbf{r} \in K_{ad}$, and $\mathbf{x} = \mathcal{S}(\mathbf{u}, \mathbf{r}; \mathbf{x}_0)$. Also, let $\mathbf{p}(t)$, the adjoint trajectory state, satisfy*

$$\begin{cases} \dot{\mathbf{p}}(t) = -(\mathcal{A} + \mathcal{F}'_{\mathbf{x}(t)})\mathbf{p}(t) - \mathbf{j}_{\mathbf{x}}(t), & t \in [0, \tau], \\ \mathbf{p}(\tau) = 0. \end{cases} \quad (87)$$

If \mathbf{x}_0 is a worst initial condition over $B_{\mathbb{V}}(R_2)$, then, there is a non-negative number μ such that

$$\begin{cases} \mu (\|\mathbf{x}_0\| - R_2) = 0, \\ \mathbf{p}(0) + \mu \mathbf{x}_0 = 0. \end{cases} \quad (88)$$

Proof. Define $f(\mathbf{x}_0) := \frac{1}{2}(\|\mathbf{x}_0\|_{\mathbb{V}}^2 - R_2^2)$. Rewrite (P1) as

$$\begin{cases} \max & \mathcal{G}(\mathbf{x}_0) \\ \text{s.t.} & f(\mathbf{x}_0) \leq 0. \end{cases} \quad (89)$$

The constraint $f(\mathbf{x}_0) \leq 0$ satisfies Robinson's regularity condition [24, Section 1.7.3.2]. This allows one to apply [24, Theorem 1.56]. Let λ be a non-negative number, and define the Lagrangian

$$\mathcal{L}(\mathbf{x}_0, \lambda) := \mathcal{G}(\mathbf{x}_0, \lambda) + \mu f(\mathbf{x}_0). \quad (90)$$

Let $\mathcal{L}'_{\mathbf{x}_0} : \mathbb{V} \rightarrow \mathbb{R}$ be the Gâteaux derivative of $\mathcal{L}(\mathbf{x}_0, \lambda)$ at \mathbf{x}_0 . Identify $\mathcal{L}'_{\mathbf{x}_0}$ with an element $l_{\mathbf{x}_0} \in \mathbb{V}$. Theorem 1.56 of [24] ensures that the worst initial condition satisfies for all $\tilde{\mathbf{x}}_0 \in \mathbb{V}$ the conditions

$$f(\mathbf{x}_0) \leq 0, \quad (91a)$$

$$\mu \geq 0, \quad (91b)$$

$$\mu f(\mathbf{x}_0) = 0, \quad (91c)$$

$$\langle l_{\mathbf{x}_0}, \tilde{\mathbf{x}}_0 - \mathbf{x}_0 \rangle_{\mathbb{V}} \geq 0. \quad (91d)$$

In the following, an explicit expression for $l_{\mathbf{x}_0}$ will be derived. First, the Gâteaux derivative of $f(\mathbf{x}_0)$ at \mathbf{x}_0 along $\tilde{\mathbf{x}}_0$ is

$$f'_{\mathbf{x}_0} \tilde{\mathbf{x}}_0 = \langle \tilde{\mathbf{x}}_0, \mathbf{x}_0 \rangle_{\mathbb{V}}. \quad (92)$$

This implies that the functional $f'_{\mathbf{x}_0} : \mathbb{V} \rightarrow \mathbb{R}$ can be identified with the element \mathbf{x}_0 . The Gâteaux derivative of $\mathcal{G}(\mathbf{x}_0)$ at \mathbf{x}_0 along $\tilde{\mathbf{x}}_0$ is derived using the chain rule,

$$\mathcal{G}'_{\mathbf{x}_0} \tilde{\mathbf{x}}_0 = J'_{\mathbf{x}} \mathcal{S}'_{\mathbf{x}_0} \tilde{\mathbf{x}}_0. \quad (93)$$

The functionals $\mathcal{G}'_{\mathbf{x}_0} : \mathbb{V} \rightarrow \mathbb{R}$ and $J'_{\mathbf{x}} : L^2(0, \tau; \mathbb{X}) \rightarrow \mathbb{R}$ can be identified with some elements $\mathbf{g}_{\mathbf{x}_0} \in \mathbb{V}$ and $\mathbf{j}_{\mathbf{x}} \in L^2(0, \tau; \mathbb{X})$, respectively. Then, equality (93) implies that

$$\mathbf{g}_{\mathbf{x}_0} = \mathcal{S}'_{\mathbf{x}_0} \mathbf{j}_{\mathbf{x}}(t). \quad (94)$$

The adjoint operator $\mathcal{S}'_{\mathbf{x}_0}$ will be derived. Let $\mathcal{S}'_{\mathbf{x}_0} \tilde{\mathbf{x}}_0 = \mathbf{q}(t)$. Consider the inner-product

$$\begin{aligned} \langle \mathbf{j}_{\mathbf{x}}, \mathcal{S}'_{\mathbf{x}_0} \tilde{\mathbf{x}}_0 \rangle_{L^2(0,\tau;\mathbb{X})} &= \int_0^\tau \langle \mathbf{j}_{\mathbf{x}}(t), \mathbf{q}(t) \rangle dt \\ &= \int_0^\tau \langle -\dot{\mathbf{p}}(t) - (\mathcal{A} + \mathcal{F}'_{\mathbf{x}(t)})\mathbf{p}(t), \mathbf{q}(t) \rangle dt. \end{aligned} \quad (95)$$

Using Proposition 14 and applying integration by parts yield

$$\begin{aligned} \langle \mathbf{j}_{\mathbf{x}}, \mathcal{S}'_{\mathbf{x}_0} \tilde{\mathbf{x}}_0 \rangle_{L^2(0,\tau;\mathbb{X})} &= \langle \mathbf{p}(0), \mathbf{q}(0) \rangle_{\mathbb{V}} - \langle \mathbf{p}(\tau), \mathbf{q}(\tau) \rangle_{\mathbb{V}} \\ &+ \int_0^\tau \langle \mathbf{p}(t), \dot{\mathbf{q}}(t) - (\mathcal{A} + \mathcal{F}'_{\mathbf{x}(t)})\mathbf{q}(t) \rangle dt \\ &= \langle \mathbf{p}(0), \tilde{\mathbf{x}}_0 \rangle_{\mathbb{V}}. \end{aligned} \quad (96)$$

It follows that $\mathcal{S}'_{\mathbf{x}_0} \mathbf{j}_{\mathbf{x}} = \mathbf{p}(0)$, and so

$$\mathbf{g}_{\mathbf{x}_0} = \mathbf{p}(0). \quad (97)$$

Combining (92) and (97) yield

$$l_{\mathbf{x}_0} = \mathbf{p}(0) + \mu \mathbf{x}_0. \quad (98)$$

Substituting this in (91) yields

$$\langle \mathbf{p}(0) + \mu \mathbf{x}_0, \tilde{\mathbf{x}}_0 - \mathbf{x}_0 \rangle_{\mathbb{V}} \geq 0, \quad \forall \tilde{\mathbf{x}}_0 \in \mathbb{V}. \quad (99)$$

Since $\tilde{x}_0 \in \mathbb{V}$ is arbitrary, the inequality condition (99) becomes an equality condition. This together with (91c) yields (88). \square

For linear systems with quadratic cost, the adjoint trajectory state satisfies $\mathbf{p}^o(t) = \Pi(t)\mathbf{x}^o(t)$ where $\Pi(t)$ solves (74). Consequently, the optimality condition $\mathbf{p}^o(0) + \mu\mathbf{x}_0 = 0$ becomes

$$\Pi(0)\mathbf{x}_0 = -\mu\mathbf{x}_0. \quad (100)$$

This implies that the worst initial condition is an eigenfunction of the operator $\Pi(0)$.

VI. KURAMOTO–SIVASHINSKY EQUATION

For every actuator location $r \in (0, 1)$, let the function $b(\cdot; r)$ be in $C^1[0, 1]$. Consider the controlled Kuramoto–Sivashinsky equation with Dirichlet boundary conditions and initial condition $w_0(\xi)$ on $\xi \in [0, 1]$ and some number λ

$$\begin{cases} \frac{\partial w}{\partial t} + \frac{\partial^4 w}{\partial \xi^4} + \lambda \frac{\partial^2 w}{\partial \xi^2} + w \frac{\partial w}{\partial \xi} = b(\xi; r)u(t), & t > 0, \\ w(0, t) = w(1, t) = 0, & t \geq 0, \\ \frac{\partial w}{\partial \xi}(0, t) = \frac{\partial w}{\partial \xi}(1, t) = 0, & t \geq 0, \\ w(\xi, 0) = w_0(\xi), & \xi \in [0, 1]. \end{cases}$$

Define the state $\mathbf{x}(t) := w(\cdot, t)$, the state space $\mathbb{X} := L^2(0, 1)$. Let the state operator $\mathcal{A} : D(\mathcal{A}) \subset \mathbb{X} \rightarrow \mathbb{X}$ be

$$\begin{aligned} \mathcal{A}w &:= -w_{\xi\xi\xi\xi} - \lambda w_{\xi\xi}, \\ D(\mathcal{A}) &= H^4(0, 1) \cap H_0^2(0, 1). \end{aligned} \quad (101)$$

Also, the control space is $\mathbb{U} := \mathbb{R}$. The actuator design space is $\mathbb{K} := \mathbb{R}$. Define $\mathbb{V} := H_0^1(0, 1)$; the nonlinear operator $\mathcal{F}(\cdot) : \mathbb{V} \rightarrow \mathbb{X}$ and the input operator $\mathcal{B}(\cdot) : \mathbb{K} \rightarrow \mathcal{L}(\mathbb{U}, \mathbb{X})$ are defined as

$$\mathcal{F}(w) := -w w_{\xi}, \quad (102)$$

$$\mathcal{B}(r)u := b(\xi, r)u. \quad (103)$$

The state space representation of the model will then be (5).

The operator $\mathcal{A} : D(\mathcal{A}) \rightarrow \mathbb{X}$ is a self-adjoint operator, is bounded from below, and has compact resolvent. According to Theorem [54, Theorem 32.1], \mathcal{A} generates an analytic semigroup on \mathbb{X} . Since the operator \mathcal{A} is analytic on a Hilbert space, Theorem 4.1 in [53] ensures that this operator enjoys maximal parabolic regularity. Also, by Rellich-Kondrachov compact embedding theorem [63, Chapter 6], the space $D(\mathcal{A})$ is compactly embedded in \mathbb{X} . The operator \mathcal{A} is also associated with a form described in A2.

Lemma 16. *The nonlinear operator $\mathcal{F}(\cdot)$ is Gâteaux differentiable from \mathbb{V} to \mathbb{X} . The Gâteaux derivative of $\mathcal{F}(\cdot)$ at w in the direction f is $\mathcal{F}'_w f = -w f_{\xi} - w_{\xi} f$.*

Proof. The operator \mathcal{F}'_w , if exists, needs to satisfy

$$\lim_{\epsilon \rightarrow 0} \left\| \frac{1}{\epsilon} (\mathcal{F}(w + \epsilon f) - \mathcal{F}(w)) - \mathcal{F}'_w f \right\|_{L^2} = 0. \quad (104)$$

Substituting in (102), inside the limit becomes

$$\begin{aligned} & \left\| \frac{1}{\epsilon} (w w_{\xi} - (w + \epsilon f)(w_{\xi} + \epsilon f_{\xi})) - w f_{\xi} - w_{\xi} f \right\|_{L^2} \\ &= \|\epsilon f f_{\xi}\|_{L^2}. \end{aligned} \quad (105)$$

Note that $f \in H_0^1(0, 1)$. Embedding $H_0^1(0, 1) \hookrightarrow C[0, 1]$ means that f is a continuous function over $[0, 1]$. This implies that $f f_{\xi}$ is in $L^2(0, 1)$, thus

$$\lim_{\epsilon \rightarrow 0} \epsilon \|f f_{\xi}\|_{L^2} = 0. \quad (106)$$

The lemma now follows from the uniqueness of Gâteaux derivative. \square

Note that $D_{\mathcal{A}}(1/2, 2) = H_0^2(0, 1) \hookrightarrow \mathbb{V}$ (see [64, Corollary 4.10]). The operator $\mathcal{F}(\cdot) : \mathbb{V} \rightarrow \mathbb{X}$ is not however weakly continuous, and does not satisfy assumption B1 of [57].

For all functions f and w in $H_0^1(0, 1)$ and g in $H^1(0, 1)$, the adjoint of \mathcal{F}'_w satisfies

$$\langle f, \mathcal{F}'_w{}^* g \rangle_{L^2} = \langle \mathcal{F}'_w f, g \rangle_{L^2} = \int_0^1 (-w f_{\xi} - w_{\xi} f) g d\xi. \quad (107)$$

Performing integration by parts yields

$$\int_0^1 (-w f_{\xi} - w_{\xi} f) g d\xi = - \int_0^1 w g_{\xi} f d\xi. \quad (108)$$

The operator $\mathcal{F}'_w{}^*$ maps $D(\mathcal{F}'_w{}^*) = H^1(0, 1)$ to $L^2(0, 1)$ as follows

$$\mathcal{F}'_w{}^* g = -w g_{\xi}. \quad (109)$$

In addition,

$$\mathcal{B}^*(r)w = \int_0^1 b(\xi, r)w(\xi) d\xi, \quad \forall w \in \mathbb{V}, \quad (110)$$

$$(\mathcal{B}'_r u)^* f = u \int_0^1 b_r(\xi; r)f(\xi) d\xi, \quad \forall f \in \mathbb{V}. \quad (111)$$

Also, define

$$K_{ad} := \{r \in [a, b] : 0 < a < b < 1\}. \quad (112)$$

Global stability of an uncontrolled KS equation has been studied extensively, see e.g. [50], [49], [65], [66]. Theorem 2.1 of [49] proves that for $\lambda < 4\pi^2$, the uncontrolled KS equation is globally exponentially stable. Proof of this theorem can be modified to ensure that there is solution to the controlled KS equation over $[0, \tau]$ for all initial conditions in \mathbb{V} . The following lemma ensures that for some parameters λ there is a solution to the KS equation for all initial conditions and inputs over arbitrary time intervals.

Lemma 17. *Let $\lambda < 4\pi^2$ and $\sigma(\lambda)$ be the smallest eigenvalue of $-\mathcal{A}$. For all initial conditions $w_0 \in \mathbb{V}$ and inputs $u \in L^2(0, \tau)$, the strict solution to the KS system satisfies*

$$\|w(\tau)\|^2 \leq \|w_0\|^2 + \frac{1}{\sigma(\lambda)} \|u\|_{L^2(0, \tau)}^2 \max_{\xi \in [0, 1]} b^2(\xi; r).$$

Proof. Theorem 4 ensures that there is a solution $w \in \mathbb{W}(0, \tau)$ over $[0, \tau]$ to the KS system with initial condition $w_0 \in \mathbb{V}$ and input $u \in L^2(0, \tau)$. Consider the Lyapunov function

$$E(t) := \int_0^1 w^2(\xi, t) d\xi. \quad (113)$$

Since $w \in W^{1,2}(0, \tau; \mathbb{X})$, the function $E(t)$ is differentiable. Taking the derivative of $E(t)$ and applying [49, Lemma 3.1] yield

$$\dot{E}(t) \leq -2\sigma(\lambda)E(t) + 2 \int_0^1 w(\xi, t)b(\xi; r)u(t)d\xi. \quad (114)$$

Apply Young's inequality to the integral term, for every $\epsilon > 0$,

$$\dot{E}(t) \leq (-2\sigma(\lambda) + \epsilon)E(t) + \frac{1}{\epsilon} \int_0^1 b^2(\xi; r)u^2(t)d\xi. \quad (115)$$

Let $\epsilon = \sigma(\lambda)$. Taking an integral over $[0, \tau]$ yields the desired inequality in the lemma. \square

Since the KS system satisfies assumptions A1-A5, Corollary 12 can be applied to obtain the optimality conditions. The cost function to be optimized is

$$J(\mathbf{x}, \mathbf{u}, \mathbf{r}) = \int_0^\tau \int_0^1 w^2(\xi, t)d\xi dt + \int_0^\tau u^2(t)dt. \quad (116)$$

Letting $\mathbf{p}(t) = f(\cdot, t)$, the optimizer $(u^\circ, r^\circ, w^\circ, f^\circ)$ with initial condition $w_0(\xi) \in H_0^1(0, 1)$ satisfies

$$\begin{cases} \frac{\partial w^\circ}{\partial t} + \frac{\partial^4 w^\circ}{\partial \xi^4} + \lambda \frac{\partial^2 w^\circ}{\partial \xi^2} + w^\circ \frac{\partial w^\circ}{\partial \xi} = b(\xi; r^\circ)u^\circ(t), & t > 0 \\ w^\circ(0, t) = w^\circ(1, t) = 0, & t > 0 \\ \frac{\partial w^\circ}{\partial \xi}(0, t) = \frac{\partial w^\circ}{\partial \xi}(1, t) = 0, & t > 0 \\ w^\circ(\xi, 0) = w_0(\xi), \\ \frac{\partial f^\circ}{\partial t} - \frac{\partial^4 f^\circ}{\partial \xi^4} - \lambda \frac{\partial^2 f^\circ}{\partial \xi^2} - w^\circ \frac{\partial f^\circ}{\partial \xi} = -w^\circ(\xi, t), & t > 0 \\ f^\circ(0, t) = f^\circ(1, t) = 0, & t > 0 \\ \frac{\partial f^\circ}{\partial \xi}(0, t) = \frac{\partial f^\circ}{\partial \xi}(1, t) = 0, & t > 0 \\ f^\circ(\xi, \tau) = 0, \\ u^\circ(t) = - \int_0^1 b(\xi; r^\circ)f^\circ(\xi, t) d\xi, & t > 0, \\ \int_0^\tau \int_0^1 u^\circ(t)b_r(\xi; r^\circ)f^\circ(\xi, t) d\xi dt = 0. \end{cases}$$

The worst initial condition over a unit ball satisfies

$$\begin{cases} \mu \left(\|w_0\|_{H_0^1(0,1)} - 1 \right) = 0, \\ f^\circ(\xi, 0) + \mu w_0(\xi) = 0. \end{cases} \quad (117)$$

VII. NONLINEAR DIFFUSION

Consider the transfer of heat in a bounded, open, connected set $\Omega \subset \mathbb{R}^2$. It is assumed that Ω has a Lipschitz boundary separated into $\partial\Omega = \overline{\Gamma_0} \cup \overline{\Gamma_1}$ where $\Gamma_0 \cap \Gamma_1 = \emptyset$ and $\Gamma_0 \neq \emptyset$. Denote by ν the unit outward normal vector field on $\partial\Omega$. The class of nonlinear heat transfer models is, for actuator shape $r \in C^1(\overline{\Omega})$,

$$\begin{cases} \frac{\partial w}{\partial t}(\xi, t) = \\ \Delta w(\xi, t) + F(w(\xi, t)) + r(\xi)u(t), & (\xi, t) \in \Omega \times (0, \tau], \\ w(\xi, t) = 0, & (\xi, t) \in \Gamma_0 \times [0, \tau], \\ \frac{\partial w}{\partial \nu}(\xi, t) = 0, & (\xi, t) \in \Gamma_1 \times [0, \tau], \\ w(\xi, 0) = w_0(\xi), & \xi \in \Omega. \end{cases}$$

Defining $\mathbb{K} = L^2(\Omega)$, a set of admissible actuator shapes is

$$K_{ad} = \{r \in C^1(\overline{\Omega}) : \|r\|_{C^1(\overline{\Omega})} \leq 1\}.$$

The set K_{ad} is compact in \mathbb{K} with respect to the norm topology [63, Chapter 6].

Let $\mathbb{X} := L^2(\Omega)$, $\mathbb{U} := \mathbb{R}$, and the state $\mathbf{x}(t) := w(\cdot, t)$. The operator $\mathcal{A} : D(\mathcal{A}) \rightarrow \mathbb{X}$ is defined as

$$\mathcal{A}w = \Delta w, \quad (118a)$$

$$D(\mathcal{A}) = \left\{ w \in H^2(\Omega) \cap H_{\Gamma_0}^1 : \frac{\partial w}{\partial \nu} = 0 \text{ on } \Gamma_1 \right\}. \quad (118b)$$

The operator \mathcal{A} self-adjoint, non-negative and has compact resolvent. Thus, it generates an analytic semi-group on the Hilbert space $L^2(\Omega)$ [54, Theorem 32.1], and has maximal L^p regularity.

Define $\mathbb{V} = H_{\Gamma_0}^1(\Omega)$ and assume that the nonlinear operator $\mathcal{F}(\cdot) : \mathbb{V} \rightarrow \mathbb{X}$. The proof of the following lemma is the same as that of [57, Lemma 7.1.1].

Lemma 18. *Let $\mathbb{V} = H_{\Gamma_0}^1(\Omega)$. Assume that*

- 1) $F(\zeta)$ is twice continuously differentiable over \mathbb{R} ; denote its derivatives by $\mathcal{F}'(\zeta)$ and $\mathcal{F}''(\zeta)$;
- 2) there are numbers $a_0 > 0$ and $b > 1/2$ such that $|\mathcal{F}''(\zeta)| \leq a_0(1 + |\zeta|^b)$.

Then $\mathcal{F}(\cdot)$ is Gâteaux differentiable from \mathbb{V} to \mathbb{X} . The Gâteaux derivative of $\mathcal{F}(\cdot)$ at $w(\xi)$ in the direction $f(\xi)$ is $\mathcal{F}'_w f = \mathcal{F}'(w)f$.

It is straightforward to show that the operator $\mathcal{F}'_w : \mathbb{V}(\subset \mathbb{X}) \rightarrow \mathbb{X}$ is self-adjoint, i.e.,

$$\langle \mathcal{F}'_w^* g, f \rangle = \langle g, \mathcal{F}'_w f \rangle, \quad \forall f, g \in \mathbb{V}. \quad (119)$$

Define $\mathbb{U} = \mathbb{R}$ and the input operator $\mathcal{B}(r) \in \mathcal{L}(\mathbb{U}, \mathbb{X})$ maps u to $r(\xi)u$. Also, for all f in \mathbb{X}

$$\mathcal{B}^*(r)f = \int_\Omega r(\xi)f(\xi)d\xi, \quad (120)$$

$$(\mathcal{B}'_r u)^* f = uf. \quad (121)$$

For every initial condition in \mathbb{V} , a strict solution over $[0, \tau]$ to the nonlinear heat equation is not guaranteed. The following lemma states a condition under which there is a solution to the diffusion equation for all initial conditions and inputs over arbitrary time intervals.

Lemma 19. *If the function $F(\zeta)$ satisfies $\zeta F(\zeta) \leq 0$ for all $\zeta \in \mathbb{R}$, then there is $c_\Omega > 0$ such that the strict solution to the nonlinear heat equation satisfies*

$$\|w(\tau)\|^2 \leq \|w_0\|^2 + \frac{4}{c_\Omega} \|u\|_{L^2(0, \tau)}^2 \|r\|_{\mathbb{K}}^2.$$

Proof. Theorem 1 in [67] proves that the nonlinear equation in one spatial dimension is input-to-state stable. This lemma extends [67, Theorem 1] to two-spatial dimension. Using the same idea of proof, consider the Lyapunov function

$$E(t) := \int_\Omega w^2(\xi, t) d\xi. \quad (122)$$

The function $E(t)$ is differentiable since $w \in W^{1,2}(0, \tau; \mathbb{X})$. Take the derivative of this function, substitute for $\dot{w}(\xi, t)$ from the heat equation, and perform integration by parts as follows

$$\begin{aligned} \dot{E}(t) &= 2 \int_{\Omega} w(\xi, t) (\Delta w(\xi, t) + F(w(\xi, t)) + r(\xi)u(t)) d\xi \\ &= 2 \int_{\Gamma} w(\xi, t) \frac{\partial w}{\partial \nu}(\xi, t) d\xi - 2 \int_{\Omega} (\nabla w(\xi, t))^2 d\xi \\ &\quad + 2 \int_{\Omega} w(\xi, t) (F(w(\xi, t)) + r(\xi)u(t)) d\xi. \end{aligned} \quad (123)$$

Apply the boundary conditions. Use Poincaré inequality and let c_{Ω} be its constant. Also, use Young's inequality for all $\epsilon > 0$

$$\dot{E}(t) \leq -2(c_{\Omega} - \epsilon)E(t) + \frac{2}{\epsilon}u^2(t) \|r\|_2^2. \quad (124)$$

Set $\epsilon = c_{\Omega}/2$. Taking the integral over $[0, \tau]$ of (124) then yields the desired inequality. \square

The nonlinear heat equation satisfies assumptions A1-A5, and thus, Corollary 12 can be applied to obtain the optimality conditions. The cost function to be optimized is

$$J(\mathbf{x}, \mathbf{u}, \mathbf{r}) = \int_0^{\tau} \int_{\Omega} w^2(\xi, t) d\xi dt + \int_0^{\tau} u^2(t) dt. \quad (125)$$

Letting $\mathbf{p}(t) = f(\cdot, t)$, The optimizer (u^o, r^o, w^o, f^o) with initial condition $w_0 \in H_{\Gamma_0}^1(\Omega)$ satisfies

$$\begin{cases} \frac{\partial w^o}{\partial t}(\xi, t) = \Delta w^o(\xi, t) + F(w^o(\xi, t)) + r^o(\xi)u^o(t), & (\xi, t) \in \Omega \times (0, \tau], \\ w^o(\xi, t) = 0, & (\xi, t) \in \Gamma_0 \times [0, \tau], \\ \frac{\partial w^o}{\partial \nu}(\xi, t) = 0, & (\xi, t) \in \Gamma_1 \times [0, \tau], \\ w^o(\xi, 0) = w_0(\xi), & \xi \in \Omega. \end{cases}$$

$$\begin{cases} \frac{\partial f^o}{\partial t}(\xi, t) = -\Delta f^o(\xi, t) - F'(w(\xi, t))f^o(\xi, t) - w^o(\xi, t), & (\xi, t) \in \Omega \times (0, \tau], \\ f^o(\xi, t) = 0, & (\xi, t) \in \Gamma_0 \times [0, \tau], \\ \frac{\partial f^o}{\partial \nu}(\xi, t) = 0, & (\xi, t) \in \Gamma_1 \times [0, \tau], \\ f^o(\xi, \tau) = 0, & \xi \in \Omega, \end{cases}$$

$$\begin{cases} u^o(t) = - \int_{\Omega} r^o(\xi) f^o(\xi, t) d\xi, & t \in [0, \tau], \\ \int_0^{\tau} u^o(t) f^o(\xi, t) dt = 0, & \xi \in \Omega. \end{cases}$$

The worst initial condition over a unit ball satisfies

$$\begin{cases} \mu \left(\|w_0\|_{H_{\Gamma_0}^1(\Omega)} - 1 \right) = 0, \\ f^o(\xi, 0) + \mu w_0(\xi) = 0. \end{cases} \quad (126)$$

VIII. CONCLUSION

Optimal actuator design for quasi-linear infinite-dimensional systems with a parabolic linear part was considered in this paper. It was shown that the existence of an optimal control together with an optimal actuator design is guaranteed under natural assumptions. With additional assumptions of differentiability, first-order necessary optimality conditions were obtained. The theory

was illustrated by application to the Kuramoto-Sivashinsky (KS) equation and nonlinear heat equations.

Current work is concerned with developing numerical methods for solution of the optimality equations. Extension of these problems to situations where the input operator is not bounded on the state space is also of interest.

REFERENCES

- [1] K. Morris, S. Yang, Comparison of actuator placement criteria for control of structures, *Journal of Sound and Vibration* 353 (2015) 1–18.
- [2] M. I. Frecker, Recent advances in optimization of smart structures and actuators, *Journal of Intelligent Material Systems and Structures* 14 (4-5) (2003) 207–216.
- [3] M. Van De Wal, B. De Jager, A review of methods for input/output selection, *Automatica* 37 (4) (2001) 487–510.
- [4] K. Morris, Linear-quadratic optimal actuator location, *IEEE Transactions on Automatic Control* 56 (1) (2011) 113–124.
- [5] K. A. Morris, M. A. Demetriou, S. D. Yang, Using H_2 -control performance metrics for infinite-dimensional systems, *IEEE Transactions on Automatic Control* 60 (2) (2015) 450–462.
- [6] D. Kasinathan, K. Morris, H_{∞} -optimal actuator location, *IEEE Transactions on Automatic Control* 58 (10) (2013) 2522–2535.
- [7] A. Martínez, C. Rodríguez, M. E. Vázquez-Méndez, Theoretical and Numerical Analysis of an Optimal Control Problem Related to Wastewater Treatment, *SIAM Journal on Control and Optimization* 38 (5) (2000) 1534–1553.
- [8] A. Unger, F. Tröltzsch, Fast solution of optimal control problems in the selective cooling of steel, *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik* 81 (7) (2001) 447–456.
- [9] C. Li, E. Feng, J. Liu, Optimal control of systems of parabolic PDEs in exploitation of oil, *Journal of Applied Mathematics and Computing* 13 (1) (2003) 247.
- [10] J. L. Boldrini, B. M. C. Caretta, E. Fernández-Cara, Some optimal control problems for a two-phase field model of solidification, *Revista Matemática Complutense* 23 (1) (2009) 49.
- [11] D. Hömberg, C. Meyer, J. Rehberg, W. Ring, D. H. Omberg, Optimal control for the thermistor problem, *SIAM Journal on Control and Optimization* 48 (5) (2010) 3449–3481.
- [12] R. Buchholz, H. Engel, E. Kammann, F. Tröltzsch, On the optimal control of the Schlögl-model, *Computational Optimization and Applications* 56 (1) (2013) 153–185.
- [13] E. Casas, C. Ryll, F. Tröltzsch, Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems, *Computational Methods in Applied Mathematics* 13 (4) (2013) 415–442.
- [14] M. S. Edalatzaheh, A. Alasty, Boundary exponential stabilization of non-classical micro/nano beams subjected to nonlinear distributed forces, *Applied Mathematical Modelling* 40 (3) (2016) 2223–2241.
- [15] J. C. de los Reyes, R. Herzog, C. Meyer, Optimal control of static elastoplasticity in primal formulation, *SIAM Journal on Control and Optimization* 54 (6) (2016) 3016–3039.
- [16] I. Yousept, Optimal control of non-smooth hyperbolic evolution Maxwell equations in type-II superconductivity, *SIAM Journal on Control and Optimization* 55 (4) (2017) 2305–2332.
- [17] A. Fleig, R. Guglielmi, Optimal control of the Fokker-Planck equation with space-dependent controls, *Journal of Optimization Theory and Applications* 174 (2) (2017) 408–427.
- [18] G. Ciaramella, A. Borzi, Quantum optimal control problems with a sparsity cost functional, *Numerical Functional Analysis and Optimization* 37 (8) (2016) 938–965.
- [19] M. Hintermüller, T. Keil, D. Wegner, Optimal control of a semidiscrete Cahn-Hilliard-Navier-Stokes system with nonmatched fluid densities, *SIAM Journal on Control and Optimization* 55 (3) (2017) 1954–1989.
- [20] J. Merger, A. Borzi, R. Herzog, Optimal control of a system of reaction-diffusion equations modeling the wine fermentation process, *Optimal Control Applications and Methods* 38 (1) (2017) 112–132.
- [21] M. Sprengel, G. Ciaramella, A. Borzi, Investigation of optimal control problems governed by a time-dependent kohn-sham model, *Journal of Dynamical and Control Systems* 24 (4) (2018) 657–679.
- [22] S.-J. Kimmerle, M. Gerds, R. Herzog, Optimal control of an elastic crane-trolley-load system—a case study for optimal control of coupled ODE-PDE systems, *Mathematical and Computer Modelling of Dynamical Systems* 24 (2) (2018) 182–206.

- [23] M. S. Edalat zadeh, K. A. Morris, Stability and well-posedness of a nonlinear railway track model, *IEEE Control Systems Letters* 3 (1) (2019) 162–167.
- [24] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich, *Optimization with PDE constraints*, Vol. 23, Springer Science & Business Media, 2008.
- [25] G. Leugering, S. Engell, A. Griewank, M. Hinze, R. Rannacher, V. Schulz, M. Ulbrich, S. Ulbrich, *Constrained optimization and optimal control for partial differential equations*, Vol. 160, Springer Science & Business Media, 2012.
- [26] F. Tröltzsch, *Optimal control of partial differential equations: theory, methods, and applications*, Graduate studies in mathematics, American Mathematical Society, 2010.
- [27] M. Bergounioux, K. Kunisch, On the structure of Lagrange multipliers for state-constrained optimal control problems, *Systems & control letters* 48 (3-4) (2003) 169–176.
- [28] E. Casas, Pontryagin’s principle for state-constrained boundary control problems of semilinear parabolic equations, *SIAM Journal on Control and Optimization* 35 (4) (1997) 1297–1327.
- [29] J. P. Raymond, H. Zidani, Hamiltonian Pontryagin’s principles for control problems governed by semilinear parabolic equations, *Applied mathematics & optimization* 39 (2) (1999) 143–177.
- [30] C. Meyer, L. M. Susu, Optimal control of nonsmooth, semilinear parabolic equations, *SIAM Journal on Control and Optimization* 55 (4) (2017) 2206–2234.
- [31] C. Antoniadis, P. D. Christofides, Integrating nonlinear output feedback control and optimal actuator/sensor placement for transport-reaction processes, *Chemical Engineering Science* 56 (15) (2001) 4517–4535.
- [32] Y. Lou, P. D. Christofides, Optimal actuator/sensor placement for nonlinear control of the Kuramoto-Sivashinsky equation, *IEEE Transactions on Control Systems Technology* 11 (5) (2003) 737–745.
- [33] A. Armaou, M. A. Demetriou, Robust detection and accommodation of incipient component and actuator faults in nonlinear distributed processes, *AIChE journal* 54 (10) (2008) 2651–2662.
- [34] S. H. Moon, Finite element analysis and design of control system with feedback output using piezoelectric sensor/actuator for panel flutter suppression, *Finite Elements in Analysis and Design* 42 (12) (2006) 1071–1078.
- [35] M. R. Saviz, An optimal approach to active damping of nonlinear vibrations in composite plates using piezoelectric patches, *Smart Materials and Structures* 24 (11) (2015) 115024.
- [36] Y. Kuramoto, T. Tsuzuki, On the formation of dissipative structures in reaction-diffusion systems: Reductive perturbation approach, *Progress of Theoretical Physics* 54 (3) (1975) 687–699.
- [37] G. Sivashinsky, Nonlinear analysis of hydrodynamic instability in laminar flames—i. derivation of basic equations, *Acta astronautica* 4 (1977) 1177–1206.
- [38] R. V. Craster, O. K. Matar, Dynamics and stability of thin liquid films, *Reviews of modern physics* 81 (3) (2009) 1131.
- [39] A. Novick-Cohen, Interfacial instabilities in directional solidification of dilute binary alloys: The Kuramoto-Sivashinsky equation, *Physica D: Nonlinear Phenomena* 26 (1-3) (1987) 403–410.
- [40] R. E. LaQuey, S. M. Mahajan, P. H. Rutherford, W. M. Tang, Nonlinear saturation of the trapped-ion mode, *Phys. Rev. Lett.* 34 (7) (1975) 391–394.
- [41] I. Bena, C. Misbah, A. Valance, Nonlinear evolution of a terrace edge during step-flow growth, *Phys. Rev. B* 47 (12) (1993) 7408–7419.
- [42] P. D. Christofides, A. Armaou, Global stabilization of the Kuramoto-Sivashinsky equation via distributed output feedback control, *Systems and Control Letters* 39 (4) (2000) 283–294.
- [43] S. N. Gomes, D. T. Papageorgiou, G. A. Pavliotis, Stabilizing non-trivial solutions of the generalized Kuramoto-Sivashinsky equation using feedback and optimal control, *IMA Journal of Applied Mathematics (Institute of Mathematics and Its Applications)* 82 (1) (2017) 158–194.
- [44] R. J. Tomlin, S. N. Gomes, Point-actuated feedback control of multidimensional interfaces, *arXiv preprint arXiv:1901.09223* (2019).
- [45] E. Cerpa, Null controllability and stabilization of the linear Kuramoto-Sivashinsky equation, *Communications on Pure and Applied Analysis* 9 (1) (2010) 91–102.
- [46] E. Cerpa, A. Mercado, Local exact controllability to the trajectories of the 1-D Kuramoto-Sivashinsky equation, *Journal of Differential Equations* 250 (4) (2011) 2024–2044.
- [47] B. Sun, Maximum principle for optimal boundary control of the Kuramoto-Sivashinsky equation, *Journal of the Franklin Institute* 347 (2) (2010) 467–482.
- [48] P. Gao, Optimal distributed control of the Kuramoto-Sivashinsky equation with pointwise state and mixed control-state constraints, *IMA Journal of Mathematical Control and Information* 33 (3) (2016) 791–811.
- [49] W.-j. Liu, M. Krstic, Stability enhancement by boundary control in the Kuramoto-Sivashinsky equation, *Nonlinear Analysis: Theory, Methods & Applications* 46 (4) (2001) 485 – 507.
- [50] R. Al Jamal, K. Morris, Linearized stability of partial differential equations with application to stabilization of the Kuramoto-Sivashinsky equation, *SIAM Journal on Control and Optimization* 56 (1) (2018) 120–147.
- [51] A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems*, Springer Science & Business Media, 2012.
- [52] A. Bensoussan, G. Da Prato, M. C. Delfour, S. K. Mitter, *Representation and Control of Infinite Dimensional Systems*, Vol. 1, 2015.
- [53] G. Dore, Lp regularity for abstract differential equations, in: *Functional Analysis and Related Topics*, 1991, Springer, 1993, pp. 25–38.
- [54] G. R. Sell, Y. You, *Dynamics of evolutionary equations*, Vol. 143, Springer Science & Business Media, 2013.
- [55] P. Clément, S. Li, Abstract parabolic quasilinear equations and applications to a groundwater flow problem, *Advances in mathematical sciences and applications* 3 (1993) 17–32.
- [56] H. Amann, Compact embeddings of vector-valued sobolev and Besov spaces, *Glasnik matematički* 35 (55) (2000) 161–177.
- [57] M. S. Edalat zadeh, K. A. Morris, Optimal actuator design for semilinear systems, *SIAM Journal on Control and Optimization* 57 (4) (2019) 2992–3020.
- [58] S. Lang, *Real and functional analysis*, Vol. 142, Springer Science & Business Media, 2012.
- [59] D. Dier, Non-autonomous maximal regularity for forms of bounded variation, *Journal of Mathematical Analysis and Applications* 425 (1) (2015) 33–54.
- [60] A. Wouk, *A course of applied functional analysis* (1979).
- [61] R. F. Curtain, H. Zwart, *An introduction to infinite-dimensional linear systems theory*, Texts in Applied Mathematics, Springer New York, 2012.
- [62] I. Lasiecka, R. Triggiani, *Control theory for partial differential equations: continuous and approximation theories*, Vol. 1, Cambridge University Press Cambridge, 2000.
- [63] R. A. Adams, J. J. F. Fournier, *Sobolev spaces*, Pure and Applied Mathematics, Elsevier Science, 2003.
- [64] S. N. Chandler-Wilde, D. P. Hewett, A. Moiola, Interpolation of Hilbert and Sobolev spaces: quantitative estimates and counterexamples, *Mathematika* 61 (2) (2015) 414–443. [arXiv:1404.3599](https://arxiv.org/abs/1404.3599).
- [65] G. Fantuzzi, A. Wynn, Semidefinite relaxation of a class of quadratic integral inequalities, in: *Decision and Control (CDC)*, 2016 IEEE 55th Conference on, IEEE, 2016, pp. 6192–6197.
- [66] M. Ahmadi, G. Valmorbidia, A. Papachristodoulou, Dissipation inequalities for the analysis of a class of PDEs, *Automatica* 66 (2016) 163–171.
- [67] F. Mazenc, C. Prieur, Strict Lyapunov functions for semilinear parabolic partial differential equations, *Mathematical Control and Related Fields* 1 (2) (2011) 231–250.